Estimation of extreme value dependence: application to Australian spot electricity prices

by

Nadezda Frolova

A thesis submitted in partial fulfillment of the requirements for the degree of

Masters of Science

in

Statistics

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

There is an increasing interest in extreme value analysis for financial and climate data. Various statistical methods have been developed for estimating extreme value dependence in time series data sets and the field continues to grow. In this work we consider four statistical methods for estimating extreme value dependence: the extremogram and cross extremogram, quantile regression, the cross-quantilogram and the upper tail dependence coefficient estimated using Gumbel copulas. We consider within series dependence but also cross serial dependence which may be of more interest. We compare the four methods using a data set of spot electricity prices from Australian states included in the National Electricity Market. In addition we discuss the advantages and disadvantages of each method. Finally, a freeware R package, *extremogram*, is made available that implements the extremogram methods.

Acknowledgements

First of all, I would like to express my gratitude to my supervisor, Dr. Ivor Cribben for all his support and motivation during my Masters program. His guidance helped me in every step of my research. It was a great pleasure to work with him. I would also like to thank my supervisor Dr. Douglas Wiens for his academic advise and his valuable comments to my thesis paper.

I would also like to thank the Department of Mathematical and Statistical Sciences for the University of Alberta for their help and encouragement, and MITACS and ATB Financial for giving me an opportunuty to apply my new knowlende to an interesting project.

And last but not least, I want to thank my family, my husband Roman and my close friends for their tremendous support and for always being there for me for the past two years. It would have been impossible for me to complete my thesis without them.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Extreme events are those events that occur far from a centre of a distribution or, in other words, in the tails of a distribution. Data sets with extreme events have more observations that deviate from the mean compared to the normal distribution. While the study of extreme events is interesting in its own right, the study of the dependence between extreme events has recently received a great deal of attention. For example, in finance, we are interested in understanding the effect of contagion on the market while in climate science the objective is to understand the behaviour between extreme events in order to plan for emergencies.

A vast literature exists on extreme event analysis (see [28] for a summary of the work). Extreme events in time series have been studied in [31], [2], [6] and [23]. Most of the work focusses on extreme value distributions and the analysis of the upper tail dependence coefficients, which is a method for estimating extreme value dependence. A large number of classical statistical methods have been applied to the quantiles of a distribution rather than to the moments, which allows for the possibility to apply these methods to high quantiles of the data to study extreme events. For example, the periodogram, which is a commonly used tool in spectral analysis of time series was extended to quantiles in [29] and [30]. Value at Risk (VaR), a traditional measure of financial risk, was extended to time series data sets in [12] to study how quantiles that are estimated by VaR change over time. In addition, modelling extreme events in climate data was addressed in [9] and [18]. And finally, [1] and [7] considered the analysis of heavy tailed distributions.

Recently, the focus of the research has shifted to statistical methods for estimating the dependence between extreme values within and between time series. The first method, the extremogram was introduced in [4]. The concept was further extended in [5] and [3]. The extremogram is a measure of dependence between events that exceed a certain threshold. It estimates the probability that an extreme event will occur in a time series at time $(t+h)$ given there is an extreme event at time t. In simple terms it measures the effect which a large value of the time series, or extreme value, has on a future of the same time series or another time series, h - time - lags ahead. The extremogram and its derivatives can easily describe graphically and quantitatively - the size and persistence of extreme value clusters. The cross extremogram allows for the estimation of the extreme value dependence between two time series. Using various cutoffs for the definition of extreme events reveals interesting dependencies that cannot be revealed using other methods.

The second method is quantile regression, which was first introduced in [25] and then further extended to time series data sets in [26]. It is a robust alternative to the well-known Least Squares estimation for regression and a generalization of the Least Absolute Deviation regression model. The main benefit of quantile regression is its ability to be used for various quantiles of the data set in order to understand the association between a response variable and a set of covariates. Setting the quantiles to the tail of the data set allows us to use quantile regression as a tool for measuring extreme value dependence.

The third method is the cross-quantilogram. It was first discovered in [22] and has recently been extended to the bivariate case in [15]. The cross-quantilogram is a robust alternative to a widely used method in time series analysis to measure dependencies between two data sets – the cross correlation function. The cross-quantilogram is simply the correlation between two regression quantiles of two time series. Similar to quantile regression, the cross-quantilogram can be applied to high quantiles to measure correlations between extreme events in two time series.

Finally, the last method, which is the most popular method in extreme value theory, the tail dependence coefficient, was used to compared to three other methods. Here, estimation is carried out using copulas [19] which is a joint distribution function with uniform marginals. It is widely used to estimate the joint distribution of two data sets when they are not normally distributed. Also, it is commonly used to study the joint tail behavior with a tail dependence coefficient – a limiting probability of one data set exceeding a certain threshold given that another data set exceeds it too.

In this work we apply all four methods to spot electricity prices of Australia (for states that are included in the National Electricity Market). We compare and contrast the extreme value dependence estimated between high electricity prices in different regions. In this work, we focus on the estimation of general dependence between extreme events in the data set without emphasizing taking into account each individual extreme event. A detailed comparison of all methods is made and a discussion of advantages and disadvantages of each method is included. Finally, we created a freeware R package, *extremogram*, that implements the extremogram methods.

The rest of the work is organized as follows: Chapter 2 describes the data set and Chapter 3 focusses on the extremogram analysis. In Chapter 3 the quantile regression method is applied and the results are discussed. Chapters 4 and 5 describe results that are obtained using the cross-quantilogram and the Gumbel copula, respectively. We conclude with a discussion.

Chapter 2

Data: The National Electricity Market of Australia

The National Electricity Market (NEM) of Australia was first established in the year 1998 and it covers the following jurisdictions: New South Wales (NSW, including the Australian Capital Territory), South Australia (SA), Queensland (QLD), Victoria (VIC), and Tasmania (TAS). NEM is a wholesale market, and most of the electricity is generated by coal, wind, gas and hydro power plants.

The Australian Energy Market Operator (AEMO) manages the wholesale electricity market by setting a spot price and the amount of electricity to be produced based on supply and demand on the market. Electricity generators from all states in NEM submit their offers to AEMO once every five minutes. A spot price is determined every half an hour for each of the regions by averaging the submitted prices and setting up quantities according to the market demand in the trading period. Every two years, AEMO also sets the highest and lowest possible spot prices, which are called the Market Price Cap and the Market Floor Price respectively.

All regions are connected by interconnectors that transport electricity between them. Australia has the world's largest system of high-voltage interconnectors that includes five regulated interconnectors (two between New South Wales and Queensland, two between South Australia and Victoria, and one between Victoria and New South Wales) and one unregulated interconnector under water between Victoria and Tasmania. Figure 2.1 shows a map of Australia with marked interconnectors. The whole system runs a distance of about 5000 kilometers in total. Unregulated interconnectors earn a fixed revenue each year. For a regulated interconnector, trading on a spot market is the main source of income. Interconnectors allow for the transport of electricity from regions with low electricity price to connected regions with higher price and, therefore, enhance market competition.

The Australian Energy Regulator (AER) is the government organization that monitors the work of electricity market to ensure that NEM follows market laws and regulations. According to AER's 2009 Annual Report, coal was the most widely used source of energy in Australia. Three states that imported more electricity than they exported were South Australia, New South Wales and Tasmania. The report stated that electricity prices in South Australia depended on market conditions the most compared to other states. In 1998-1999, the state imported up to 25% of its consumed electricity. As gas power was the main source of electricity within the region, high fuel costs resulted in higher costs of electricity generation. During periods when prices spiked, New South Wales had a higher price compared to competitive alter-

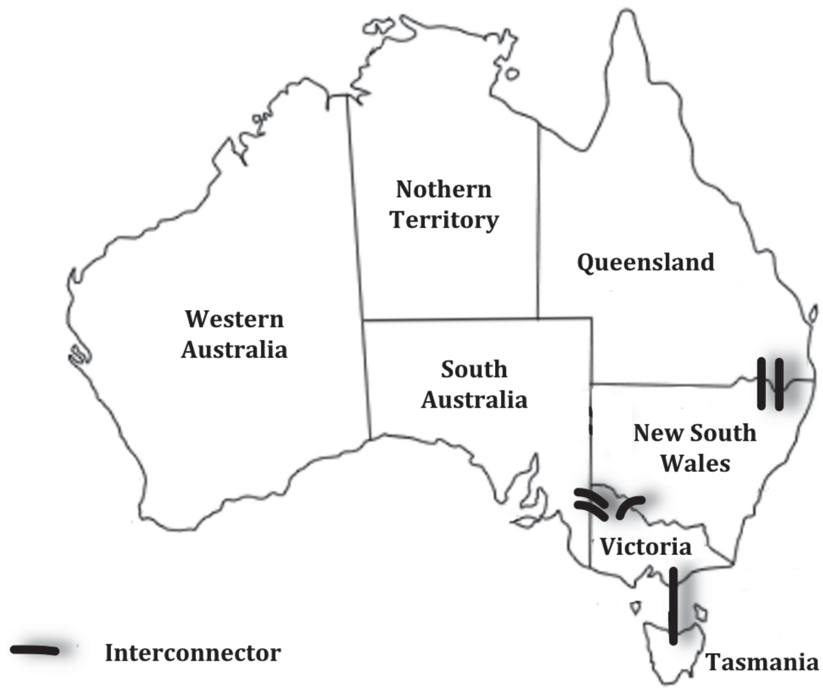

Figure 2.1: *Map of Australia with interconnectors. source: foundingdocs.gov.au*

natives in adjacent states. Before joining NEM in 2006, Tasmania had been able to satisfy all its internal electricity demand by hydroelectric power plants that had produced electricity at a relatively low price. However, after joining NEM and building an interconnector under water between itself and Victoria, a great amount of electricity began to be imported from the region with a lower price.

Victoria was a major exporting state because it had comparatively low production costs and was connected to three other states. Victoria used more hydroelectric power plans to produce energy compared to other mainland states. Electric power produced by hydroelectric power plants required lower costs compared to coal energy used as a primary source in New South Wales and South Australia. Queensland's production capacity was higher than the peak demand observed till the year 2009, which allowed the state to export a significant amount of electricity produced by its generators.

As discussed in AER's 2011 and 2014 Annual Reports, the Government of Australia introduced several climate change policies that included gradual closure of 2000 coal power plans among the country and achievement of 20% share of renewable energy by 2020. Regulations changed the main source of energy in Australia from coal to wind (up to 60% of produced energy) by the end of 2014. Another source of increase in usage of renewable energy was the installment of Rooftop Solar Photovoltaic (PV) power stations that was subsidized by the government. Power generation from sun energy was not traded through NEM. Instead, consumers that installed Rooftop PVs received a deduction from their electricity bills which caused the electricity demand on NEM to go down, and forced prices to decrease.

Climate change policies affected the distribution of importing and exporting states, making Tasmania one of the major exporting states. After new carbon pricing was introduced in 2012 as a part of the policy, it raised the competitiveness of electricity produced by hydroelectric power plants in Tasmania.

Trading between states introduces many dependencies between electricity prices because of the competitiveness of the market. Despite the ease of electricity transmission, when prices in two connected regions are low, most of the electricity consumed within each region is produced there, and electricity is not traded on the market. However, if prices are significantly different, trading between two regions becomes more important, and electricity starts to be transported more from a region with a lower price to a region with a higher price. An increase in electricity demand in the state with lower price makes the price rise in the other region within a certain amount of time. This introduces a dependence between high prices in two regions. Also, the more trading that is carried out between two regions the more dependent two regional markets become. Studying such dependencies gives a better understanding of the market and is important to create trading policies for market participants: electricity producers and consumers.

In this work, spot electricity prices data is collected for the time period starting from the year 2009 to the year 2014 every half hour. Figures 2.2, 2.3, 2.4, 2.5, and 2.6 show the time series plots for New South Wales, Queensland, South Australia, Tasmania and Victoria. Each series was split into two parts: from the year 2009 to the year 2011 (top plot) and from the year 2012 to the year 2014 (bottom plot). One can see that the majority of spot prices for all of the states of Australia stay at a low price range (e.g., for New South Wales 90% of observation are smaller than AUS \$ 50), but several spikes are evident during both of the periods for all of the states. These spikes and in particular the dependence between these spikes is of the most interest to us.

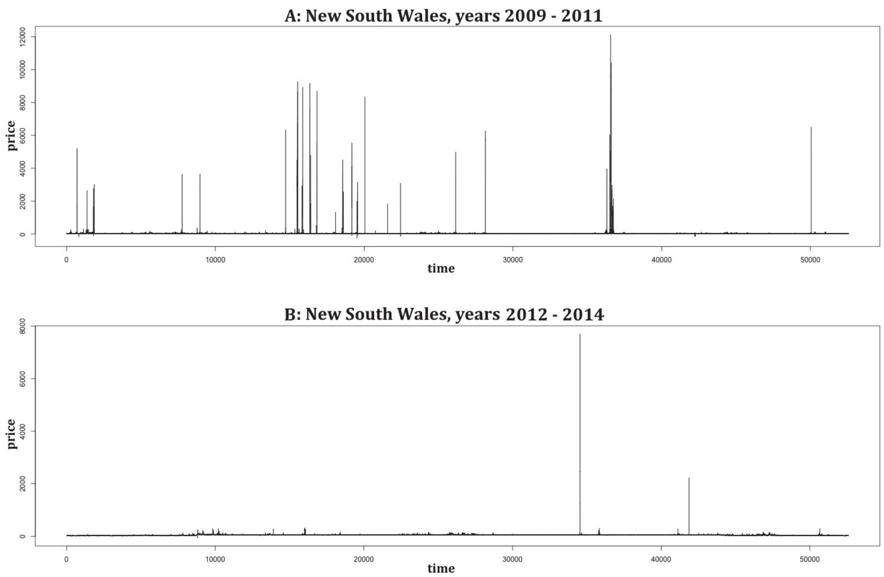

Figure 2.2: *A: Spot electricity prices for New South Wales from A: 2009 – 2011; B: 2012 – 2014.*

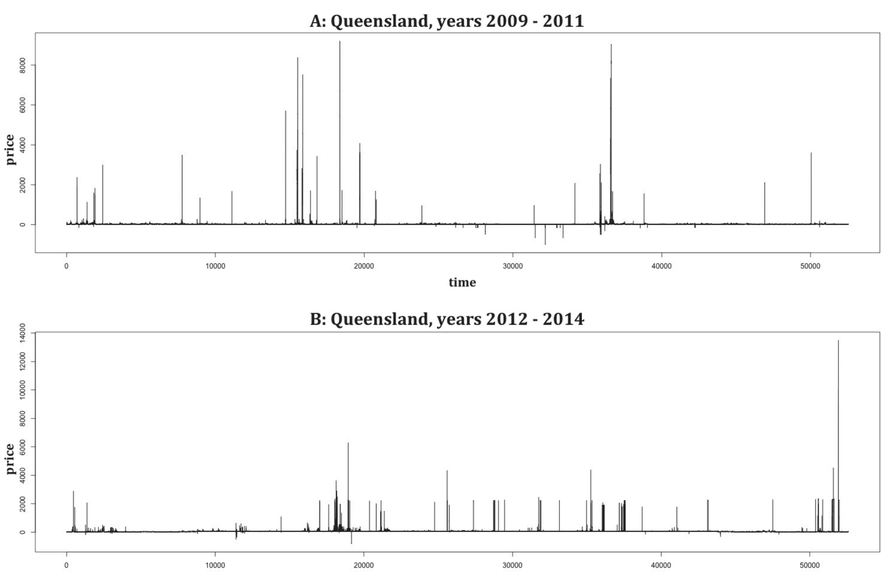

Figure 2.3: *A: Spot electricity prices for Queensland from A: 2009 – 2011; B: 2012 – 2014.*

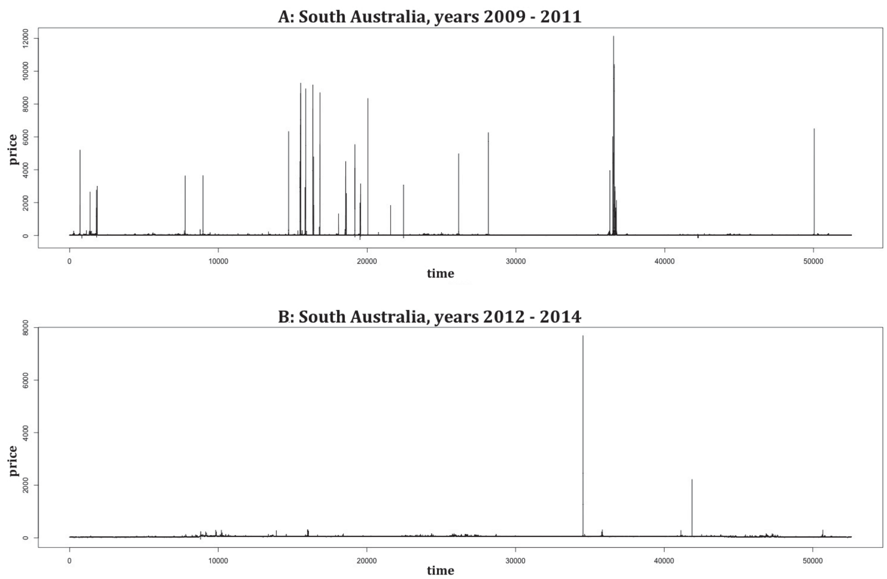

Figure 2.4: *A: Spot electricity prices for South Australia from A: 2009 – 2011; B: 2012 – 2014.*

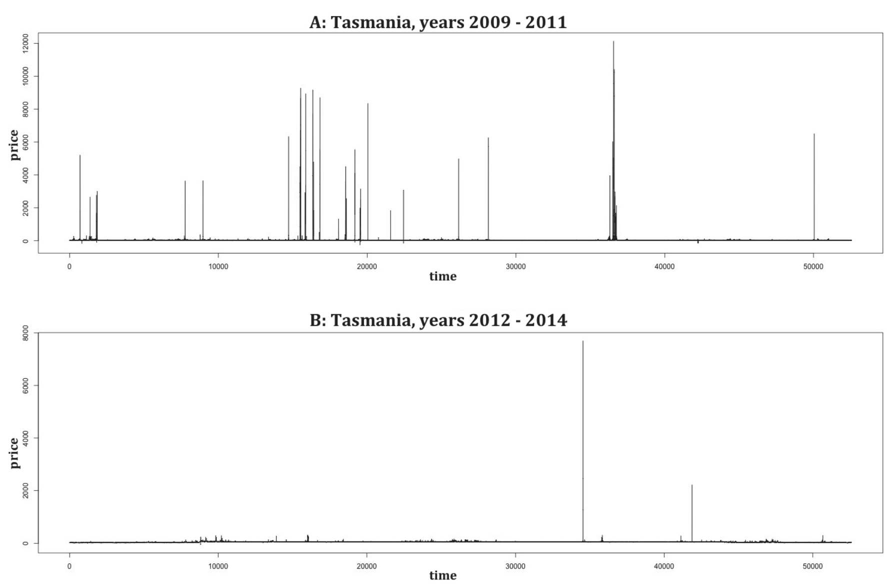

Figure 2.5: *A: Spot electricity prices for Tasmania from A: 2009 – 2011; B: 2012 – 2014.*

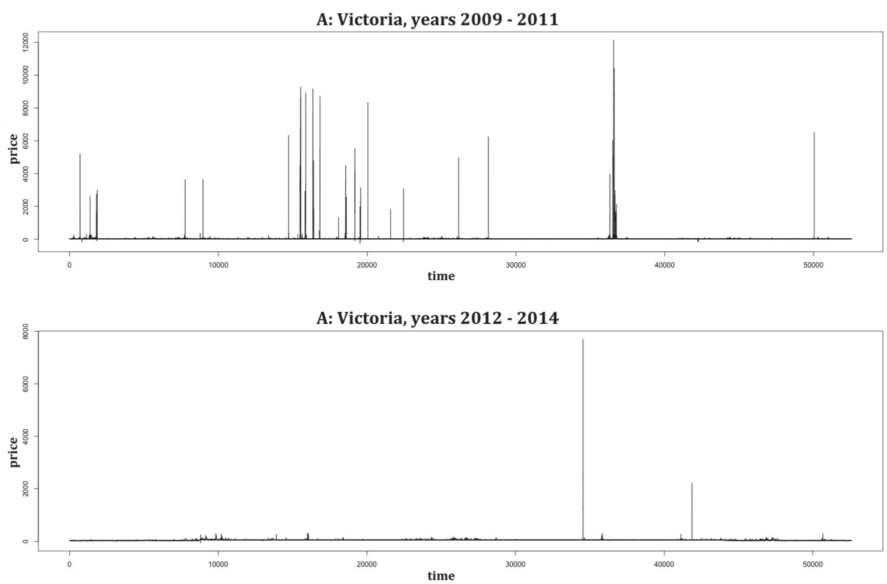

Figure 2.6: *A: Spot electricity prices for Victoria from A: 2009 – 2011; B: 2012 – 2014.*

Chapter 3

Extremogram analysis of spot electricity prices of Australia

3.1 Methodology

3.1.1 Measures of serial dependence

The autoregression function (ACF) is a well known and one of the most commonly used tools for time series analysis. It is used to study dependence within one time series. For a time series $\{X_t\}_{t=1\ldots T}$, it is a measure of linear dependence between the time series at time t and the same series at time $t+h$. The ACF is defined as:

$$\rho(h) = \frac{cov(X_t, X_{t-h})}{\sqrt{cov(X_t, X_t)cov(X_{t-h}, X_{t-h})}} \tag{3.1}$$

The ACF is used to find correlated lags within a time series, periodic trends and to identify white noise. The ACF in 3.1 is estimated by the sample

ACF as follows:

$$\widehat{\rho(h)} = \frac{\sum_{t=h+1}^{T}(X_t - \bar{X})(X_{t-h} - \bar{X})}{\sqrt{\sum_{t=1}^{T}(X_t - \bar{X})^2}\sqrt{\sum_{t=h+1}^{T}(X_{t-h} - \bar{X})^2}} \tag{3.2}$$

Another important measure of serial dependence, which is widely used in time series analysis, is the partial autocorrelation function (PACF). It is a measure of linear dependence between a time series at time t and the same series at time $t+h$ when the effect of the time series at times $(t+1)\ldots(t+h-1)$ is removed. The PACF is defined as:

$$\phi_{hh} = corr(X_h - \hat{X}_h, X_0 - \hat{X}_0) \tag{3.3}$$

where $\hat{X}_h$ and $\hat{X}_0$ are the regressions of X_h and $\hat{X}_0$ on $\{X_1, X_2, ..., X_{h-1}\}$, respectively. Similar to the ACF, the PACF is estimated with a sample partial autocorrelation function.

The two previous functions measure the linear dependence within a time series. However, sometimes we have several time series and we would like to measure linear dependence between them. The cross-correlation function (CCF) is a well-known method used in this case. Cross correlation function is a measure of linear dependence between two time series X_t and Y_{t+h}. It is defined as:

$$\rho_{X,Y}(h) = \frac{cov(X_{t+h}, Y_t)}{\sqrt{cov(X_{t+h}, X_{t+h})cov(Y_t, Y_t)}} \tag{3.4}$$

The CCF in 3.4 is estimated using the sample CCF using:

$$\widehat{\rho(h)} = \frac{\sum_{t=h+1}^{T}(X_{t+h} - \bar{X})(Y_t - \bar{Y})}{\sqrt{\sum_{t=h+1}^{T}(X_{t+h} - \bar{X})^2}\sqrt{\sum_{t=1}^{T}(Y_t - \bar{Y})^2} \tag{3.5}$$

3.1.2 Time series modelling

In statistics, several time series models are available to represent the serial dependence within a time series, its periodicity, conditional heteroskedasticity, and other features. Two most commonly used time series models are Autoregressive Moving Average (ARMA) and Integrated Autoregressive Moving Average (ARIMA) introduced in [10]. A time series is ARMA(p, q) if it is stationary and it can be represented in the following way:

$$X_t = \phi_{0+} \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=0}^{q} \theta_j \epsilon_j + \omega_t \tag{3.6}$$

with $\phi_p \neq 0$ and $\theta_q \neq 0$ and ω_t is an uncorrelated white noise term with variance of $\sigma_\omega > 0$. We can assume that X_t has mean 0. The parameters p and q are the autoregressive and the moving average orders respectively. The parameter q is estimated as a number of significant correlated lags in the ACF. Similarly, the parameter p is estimated as a number of significant correlated lags in the PACF. A time series is said to be ARIMA(p, d, q) if $\bigtriangledown^d X_t = (1 - L)^d X_t$ is ARMA(p, q), where L is a lag operator (i.e., $LX_t = X_{t-1}$).

To choose the order of the ARIMA model or perform model selection, we can use information criteria. Generally, an information criterion is a measure of goodness-of-fit relative to the number of parameters used. It aims to select a model with relatively a good fit while using a minimum number of parameters. The Akaike Information Criterion (AIC) is defined as:

$$AIC = 2k - 2ln(L) \tag{3.7}$$

where k is a number of estimated parameters in the model and $ln(L)$ is a log-likelihood function. A model with minimum AIC is considered the best compared to all other models.

Sometimes it may be interesting to model not only the series itself, but also its variance. This is widely carried out in economics and in finance. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced in [11], is used for modeling the variance. X_t follows a GARCH(m, s) model if its conditional variance can be represented as follows:

$$\epsilon_t = \sigma_t z_t, \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma_{t-i}^2 + \sum_{j=1}^{s} \beta_j \epsilon_{t-j}^2 \tag{3.8}$$

where ϵ_t is a white noise term, z_t is a N(0,1) random variable and σ_t is a time-dependent standard deviation. Essentially, the GARCH(m, s) model is the ARMA(p, q) model for squared series and the lag order of the model is estimated similarly to ARMA(p, q): by the number of correlated lags in the ACF and PACF for squared series.

3.1.3 The extremogram

The ACF, PACF and CCF described in Section 3.1.1 focus on the centre of the distribution and do not focus on the dependence in the tails. The extremogram overcomes this limitation for a stationary time series and focusses only on extreme event dependence within one or between two time series. Nowadays, extreme value analysis is becoming ever more popular. Hence, the extremogram is a useful addition to existing methods of extreme value analysis.

Consider a stationary time series X_t. As introduced in [4], the univariate extremogram as a measure of dependence between extreme values is defined as the limit (provided it exists) of the following conditional probability:

$$\rho_{A,B}(h) = \lim_{x\to\infty} P(x^{-1}X_h \in B | x^{-1}X_0 \in A), \quad h = 0, 1, 2... \tag{3.9}$$

where A, B are sets that contain extreme values (usually taken as $(1;\infty)$).

A natural estimate of the limit above is the sample extremogram:

$$\hat{\rho}_{A,B}(h) = \frac{\sum_{t=1}^{n-h} I_{(a_m^{-1}X_{t+h}\in B, a_m^{-1}X_t\in A)}}{\sum_{t=1}^{n} I_{(a_m^{-1}X_t\in A)}} \tag{3.10}$$

where A, B are sets that contain extreme values and are bounded away from 0 (usually taken as $(1;\infty)$) and a_m is a high empirical quantile. Observations above a_m are considered extreme events.

The meaning of the extremogram is as follows: the extremogram is the probability that given an extreme event at some time t there will be another one at time $t+h$. Here, sets $\{x^{-1}X_0 \in A\}$ and $\{x^{-1}X_h \in B\}$ become extreme in the limit, which means that probabilities of their occurance are converging to 0. One can study interesting dependencies between extreme events in time series with this tool by changing the sets A and B, that is setting them as upper or lower quantiles, etc. It can be particularly useful in studies of time series with heavy-tailed marginal distributions where the occurrence of extreme events (shocks) is of interest. As was already mentioned above, financial or climate data oftentimes have this feature.

The univariate extremogram allows to estimate the extreme event dependence within only one time series. However, in many cases, another time series may affect the the series that is being analyzed. In this case, it is useful to estimate the extreme event dependence between those two time series. The cross extremogram is used for this purpose. For two time series $\{X_t\}_{t=1...T}$ and $\{Y_t\}_{t=1...T}$, the cross extremogram is a measure of the conditional dependence between extreme events in bivariate time series. It is defined as a limit of the following conditional probability (provided it exists):

$$\rho_{A,B}(h) = \lim_{x\to\infty} P(x^{-1}Y_h \in B | x^{-1}X_0 \in A), \quad h = 0, 1, 2... \tag{3.11}$$

Similar to the univariate case, the natural estimate of the limiting conditional probability is the sample cross-extremogram that is calculated as follows:

$$\hat{\rho}_{A,B}(h) = \frac{\sum_{t=1}^{n-h} I_{(a_{m,Y}^{-1} Y_{t+h} \in B, a_{m,X}^{-1} X_{t+h} \in A)}}{\sum_{t=1}^{n} I_{(a_m^{-1} X_{t+h} \in A)}} \tag{3.12}$$

where $a_{m,X}$ and $a_{m,Y}$ are empirical quantiles that indicate extreme events in time series X_t and Y_t, respectively.

In simple terms, the cross extremogram values show the probability of an extreme event occurring in one time series (Y_t) at some time $t + h$ given the occurrence of an extreme event in another time series (X_t) at time t.

In practice, not all time series data sets are stationary. Using the extremogram on a nonstationary data set may lead to unreliable results. In this case one of the time series models described in Section 3.1.2 can be fit to the time series. And the extremogram can be estimated on the weakly stationary residuals.

3.1.4 Inference on the extremogram

To infer significance of the sample extremogram value at some lag h, a permutation test is used to estimate the extremogram values under the assumption of independence. To carry out the permutation test, the time series is permuted multiple times and the extremogram is re-estimated for each of the permuted samples. In the bivariate case the permutation is carried out on pairs of data (X_i, Y_i). To estimate significance bands for a desired level α, empirical quantiles $\alpha/2$ and $(1 - \alpha/2)$ are calculated for all of the extremogram values. Extremogram values calculated on the original data set that lie outside of estimated quantiles indicate significant extreme value dependence within one or between the time series.

As discussed in [5], confidence interval estimation for the extremogram is another important part of the analysis. This estimation is based on the stationary bootstrap procedure proposed in [21], which allows for the sampling of blocks of random size of the same time series with replacement and re-estimating the extremogram for each of the bootstrap samples. The stationary bootstrap differs from a fixed block bootstrap procedure in that the blocks used in the stationary bootstrap have a random size that follows the geometric distribution. Its distribution mass function is $P(X = x) = p(1 - p)^{x-1}$, where $p \in (0; 1)$. A scheme of one stationary bootstrap iteration is presented in Figure 3.1. This procedure accounts for specific time series features such as autocorrelation, periodicity, etc. by choosing an appropriate mean block size. To keep the dependence between the two series, resampling is carried out in pairwise fashion for the cross extremogram.

After applying the stationary bootstrap to the time series multiple times

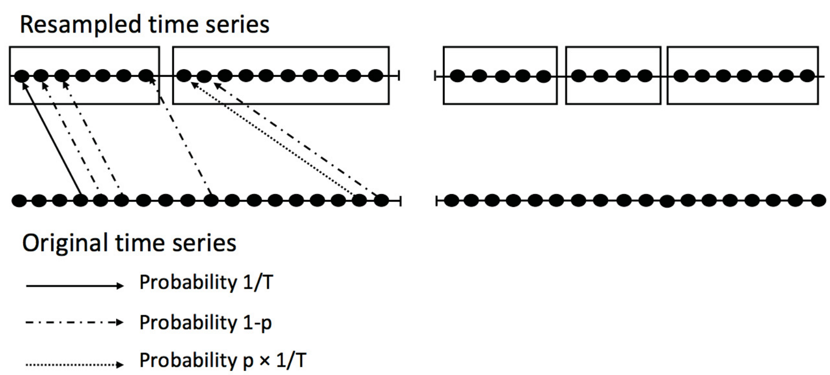

Figure 3.1: *Stationaty bootstrap procedure for a time series* $\{X_t\}_{t=1...T}$.

and re-estimating the extremogram for each of the bootstrap samples, we can calculate the confidence intervals for each of the extremogram values using the $(\alpha/2)^{th}$ and $(1-\alpha/2)^{th}$ empirical quantiles.

An R package [20] was created for extremogram method. It includes functions for the estimation of both univariate and cross extremograms, permutation confidence intervals, and stationary bootstrap confidence intervals. The package documentation file can be found in the Appendix.

3.2 Univariate extremogram analysis

The methodology described in the Section 3.1 is applied to the spot electricity prices data set described in Section 2 above. The first part of the analysis is the estimation of the univariate extremogram. The 99.5^{th} quantile was chosen to study the extreme event dependence in the data set as this quantile both significantly extreme but still has approximately 500 observations above it. As the number of the observations is sufficiently large, 99.5 can be used as a threshold for the analysis.

We first consider the first half of the spot electricity prices data set, from 2009 to 2011. For this period, most of the extremograms show a vivid periodic trend with a period 48. As the data was collected once every half hour, this means that there is a high probability given an extreme event at time t to observe another one in the next 24, (48, 96, etc.) hours. Initial lags ($h = (1, 2, 3)$) also show significant probabilities of high prices for electricity. To perform the permutation test to check for significance of the extremogram values, 1000 permutations were used. An example of a periodic trend of an extremogram can be seen in Figure 3.2 – the univariate extremogram for New South Wales. It has the clearest periodicity compared to other states. The rest of the univariate extremograms for the data set 2009 – 2011 can be found in the Appendix (p. 84, 85).

A possible reason for the periodicity in the given time series is that only 25% of the electricity was consumed by residents for individual use. Another 25% is consumed for commercial use (retail, services, etc.), and the remaining is consumed by various industries. Industries usually have high electricity

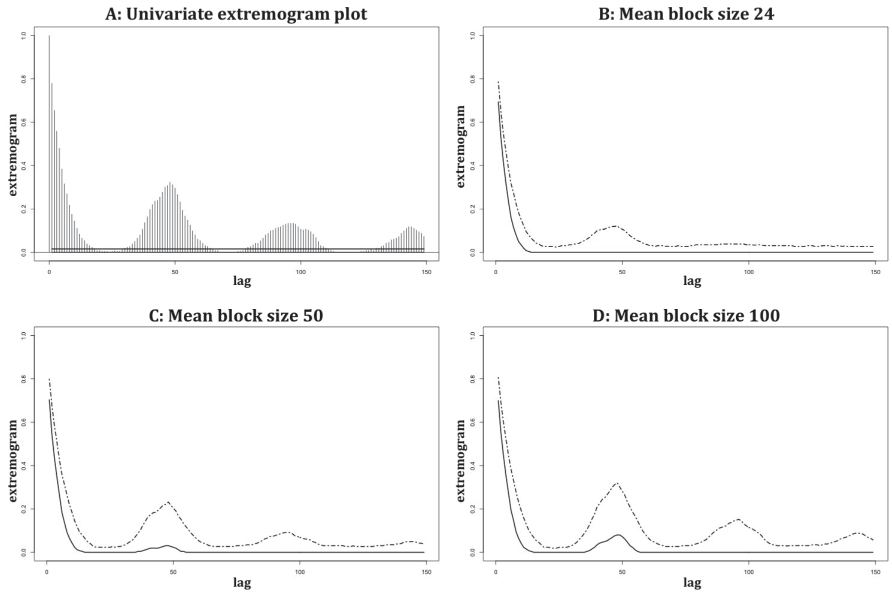

Figure 3.2: *A: Univariate extremogram plot for NSW (2009 – 2011); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

demand at the start of the working day and low demand at nighttime. As electricity prices depend on the demand, this distribution of consumption may introduce a periodic trend. This may also explain the absence of periodicity in Tasmania. This region does not have any major industries located in it and hence the electricity demand is distributed more evenly throughout the day.

The stationary bootstrap procedure was applied to estimate confidence intervals for each of the extremograms (with 1000 replications). It is evident that a mean block size of 100 best captures the periodic dependence compared to other mean block sizes that were tried - 24 and 50. Smaller block sizes allow us to only capture high probabilities for extreme event dependence at lag 48 (which corresponds to 1 day ahead) and lag 96 (two days ahead), respectively. The small number of significant values of the extremogram at lag 144 (three days after time t) is missed in both cases as the mean block size in too small to capture the serial dependence at higher lag orders.

As mentioned above, the univariate extremogram for Tasmania does not exhibit a periodic trend (Figure 3.3). The plot decays slowly starting from probability values of about 70% at lag 1 to almost 0 at the lag 140. There is some evidence of an increase in probability values at lag 48, but it does not result in a clear periodic trend that the rest of the states exhibit.

Again, block size for the stationary bootstrap appears to work best for

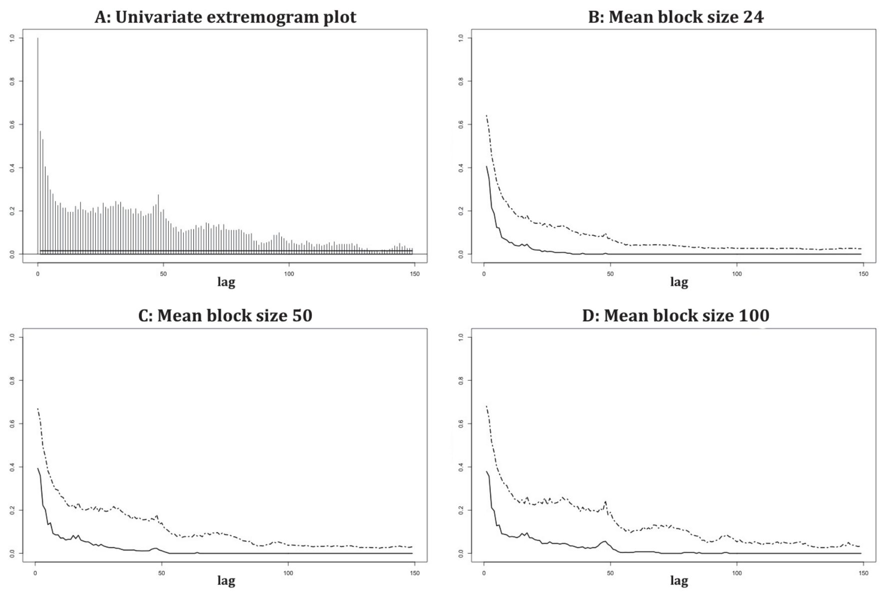

Figure 3.3: *A: Univariate extremogram plot for TAS (2009 – 2011); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

Tasmania. A block size 24 captures only the main decreasing trend without putting any emphasis on the lag 48. The reason for this is that when using a mean block size of 24 the probability of obtaining a block size 48 is very small. Using a block size of 50 provides a better idea of the specific features of the extreme dependence in Tasmania, but it decays quicker than the true extremogram. Hence, to capture all of features of the data a mean block size of 100 is used in the rest of the extremogram analysis. It is used to estimate confidence intervals if no other mean block size is.

As the next step, a time series model was fitted to the time series explain the dependence between extreme events. The extremogram procedure was applied to the model's residuals to estimate the remaining dependence. The goal of this step was to find a model that eliminates the majority of the extreme value dependence within each of the time series in order to provide a better understanding of the underlying process.

We first used as ARIMA model (see Section 3.1.2). Models were estimated using the software MATLAB. The best ARIMA model was chosen by minimizing the AIC for each of the states. ARIMA models with one seasonal difference explained more extreme event dependence compared to other ARIMA models. Nevertheless, these models still showed some visible signs of dependence. In addition, although the seasonal difference was applied to the data set, the seasonal trend was still preserved in the extremogram plots. The greatest portion of the extreme event dependence within a series model was explained by ARIMA(2,0,7)$(3,1,1)_{48}$ model for New South Wales. The extremogram for the residuals after fitting this ARIMA model is presented in Figure 3.4. Extremogram plots for the rest of the states are listed in the Appendix (p. 90, 92).

As all of the series have multiple spikes in electricity prices, a GARCH(1,1) model was fitted to the data. The model does not give any visible improvement on explaining extreme event dependence as extremogram values estimated on its residuals are close to those of the original data set. The periodic trend observed in most of the series is still preserved, as GARCH by definition does not account for periodicity. For that reason, the state where GARCH(1,1)provides an improvement is Tasmania (Figure 3.5) as it does not have any clear peri-

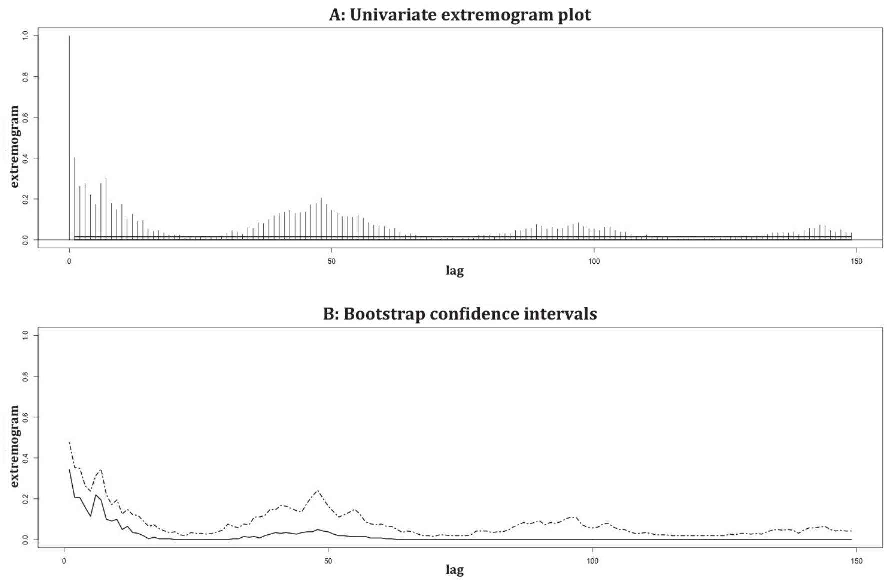

Figure 3.4: *A: Univariate extremogram plot after fitting an ARIMA(2,0,7)(3,1,1)$_{48}$ to NSW (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

odic trend. Hence, for this state the use of the GARCH(1,1) model to explain majority of extreme value dependence can be considered satisfactory.

For the period 2009 – 2011, the time series model that explained most of the extreme event dependence for all of the states was a combination of ARIMA and GARCH(1,1) models. First, an ARIMA model was fitted to account for the autocorrelation and seasonality in the data. Then, a GARCH(1,1) was fitted to the ARIMA residuals to explain the spikes in electricity prices. The univariate extremogram was estimated on the resulting residuals for each of the series. For almost all of the states, a seasonal ARIMA was essential before fitting the GARCH model, excluding Victoria. For Victoria, an ARIMA$(0,1,0)(0,1,0)_{48}$ (first difference and a seasonal difference) was sufficient (the extremogram plot can be found in Figure 3.6).

Next we consider the data from 2012 – 2014. For this period, all of

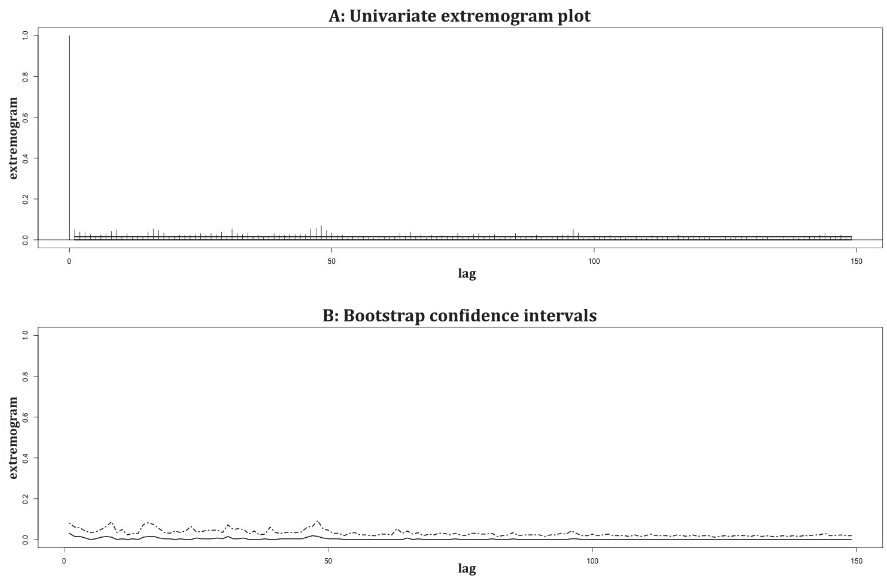

Figure 3.5: *A: Univariate extremogram plot after fitting a GARCH(1,1) to residuals for TAS (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

the states show that the extreme event dependence decreased compared to the earlier period. For those states that had a periodic trend (all of the mainland states), the period becomes less visible but still can be observed. The state where the periodicity remains the most noticeable is Victoria (Figure 3.7: a comparison of the univariate extremograms for period 2009 – 2011 and 2012 – 2014. Among the rest of the states, the one that exhibits the most visible changes in the dependence of extreme events compared to the earlier period is Queensland. Having a very clear periodic trend in the earlier period of the analysis, the state has almost no dependence left for the later period. A comparison of univariate extremograms for period 2009 – 2011 and 2012 – 2014 is in Figure 3.7. The rest of the univariate extremogram figures for the period 2012 – 2014 are presented in the Appendix (p. 86 – 88).

A possible reason for the decreasing dependence within the series could

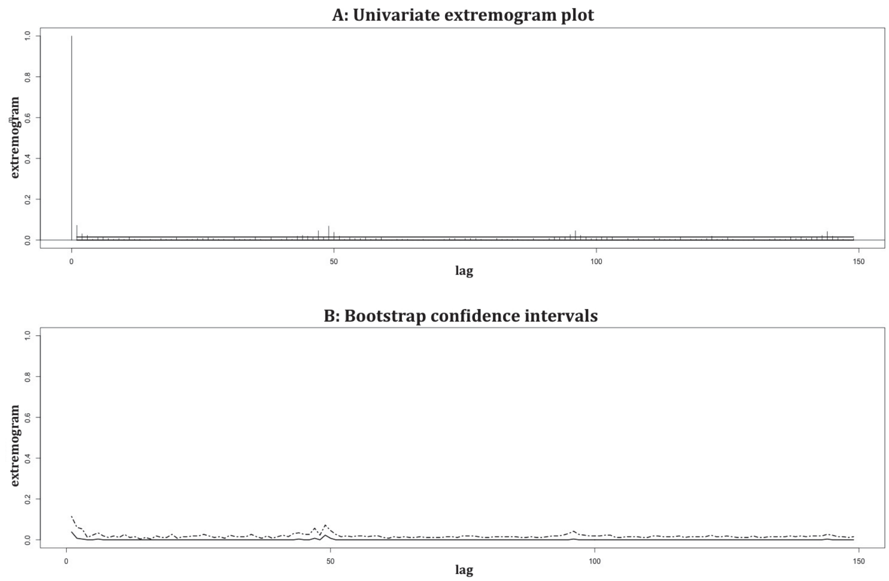

Figure 3.6: *A: Univariate extremogram plot after fitting an ARIMA and GARCH(1,1) to residuals for VIC (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

be due to the climate change policy of Australian Government described in the Section 2. Some of the electricity was not traded through NEM from the year 2012. Fewer spikes can be observed in time series Figures 2.2 – 2.6. Hence, as mentioned at the beginning of this section, the series exhibit less dependence for higher quantiles.

Similar to the period of 2009 – 2011, a mean block size of 100 for the stationary bootstrap procedure gives the most accurate estimation of the confidence intervals for the extremogram. One can see that the majority of the features of the extremogram plots for this period are captured with this mean block size.

Following the same scheme of analysis that was carried out for the 2009 – 2011 period, ARIMA models were fitted to each of the series to partially explain some of the extreme value dependence. For the later period, ARIMA

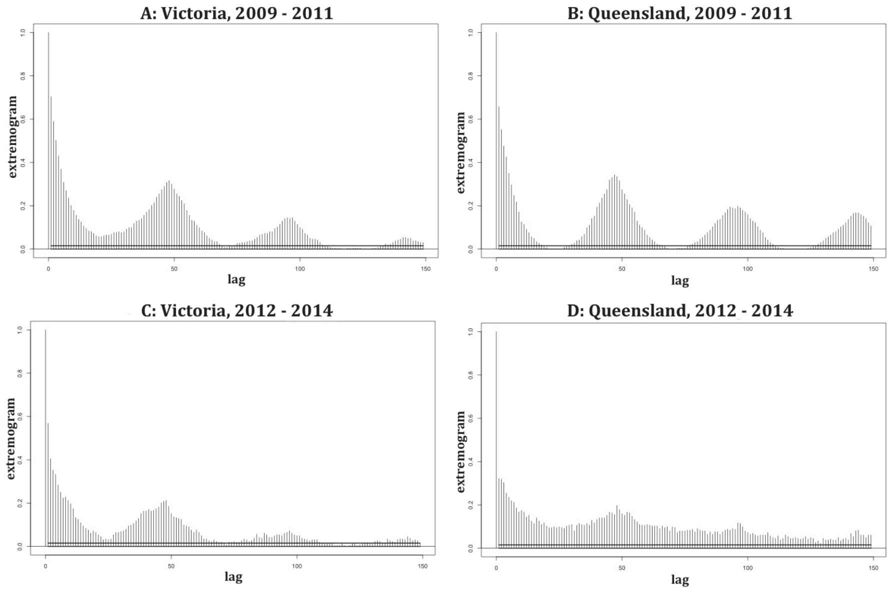

Figure 3.7: *A: Univariate extremogram plot for A: VIC (2009 – 2011); B: QLD (2009 – 2011); C:VIC (2012 – 20114); D: QLD (2012 – 2014);*

models work marginally better than for the earlier period as more extreme event dependence was explained by each of the models. Again, the ARIMA model was able to explain more dependence for New South Wales (Figure 3.8) compared to all other states (ARIMA model that was fitted to New South Wales data set is ARIMA(3,0,7)(4,1,2)$_{48}$). For the rest of the states more dependence still remains in the extremogram of the ARIMA residuals.

The next time series model that was fitted is the GARCH(1,1) model. For the later period, when the extreme event dependence was less evident for the majority of the states, and most of the periodicity was gone, GARCH(1,1) model performs fairly well. The majority of the dependence can be explained by it and the univariate extremogram estimated on its residuals results in the plot with the probability values mostly falling inside the 95% confidence bands produced by the permutation procedure. Notice that the majority of extremogram values are close to 0. Bootstrap confidence intervals support the

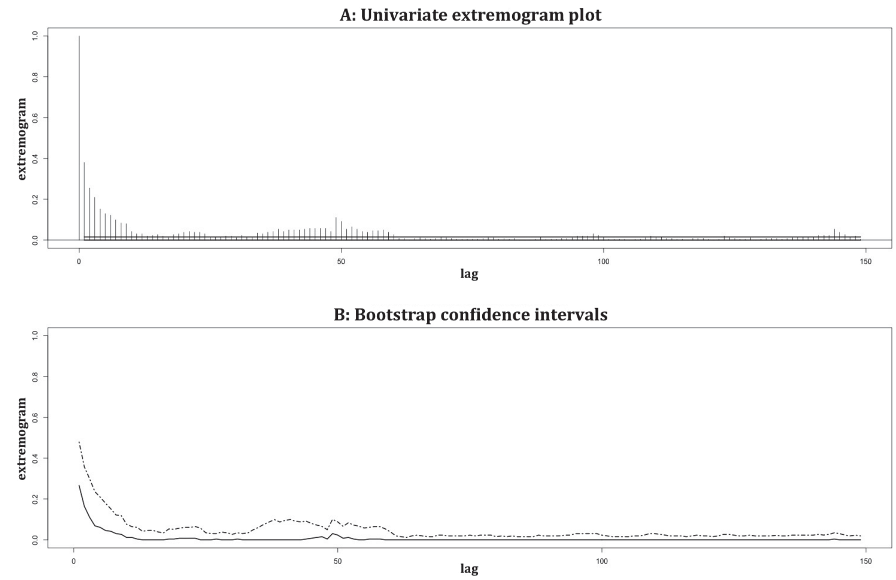

Figure 3.8: *A: Univariate extremogram plot after fitting an ARIMA(3,0,7)(4,1,2)$_{48}$ to NSW (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

statement, as for most of the lags, two confidence bands are close to the value of 0. This means that the vast majority of lags show an insignificant (if any) dependence between extreme events. As an example, an extremogram for the residuals after fitting a GARCH(1,1) to Victoria is presented in Figure 3.9. Extremogram results for the rest of the states can be found in the Appendix (p. 99, 101).

Univariate extremogram analysis was also performed on the whole data set (2009 – 2014) for each of the states. We can observe two general tendencies among the univariate extremogram plots: univariate extremograms for the whole data set look either identical to those for the period 2009 – 2011 (e.g., the extremogram for New South Wales in Figure 3.10, Tasmania and Victoria in the Appendix (p. 89, 90) or have more significant values at lags 0, 48, 96, and 144 but the peak values for those lags are smaller than those

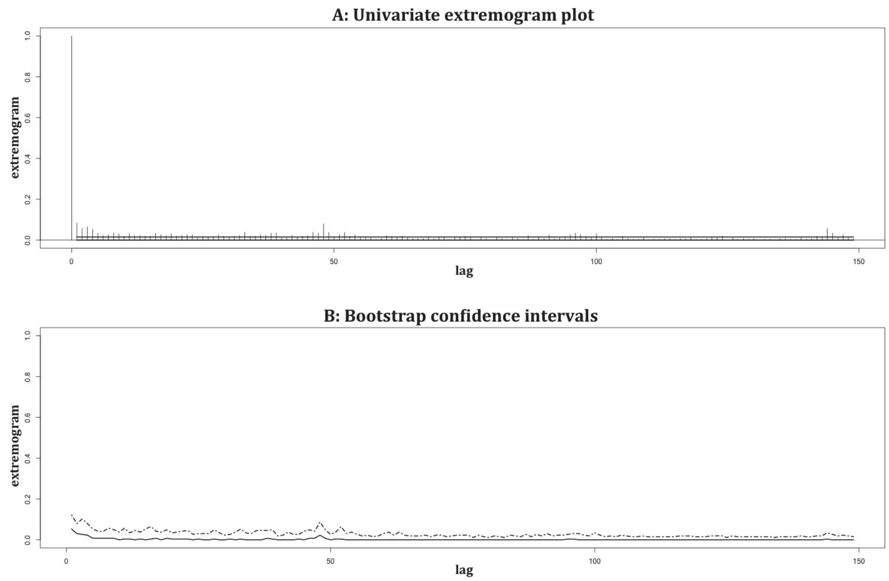

Figure 3.9: *A: Univariate extremogram plot after fitting a GARCH(1,1) to residuals for VIC (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

for the period 2009 – 2011. Two examples for this are Queensland and South Australia, which are presented in the Appendix, (p. 88).

The parts B,C and D (Figure 3.10) of the univariate extremogram plots for the whole data set suggest that the mean block size in the stationary bootstrap for the confidence intervals calculation should still be 100.

For the whole data set, the first attempt to explain the extreme event dependence with ARIMA models was not successful as most of the existing dependence was left in ARIMA residuals. Univariate extremogram plots can be seen in Figure 3.11 and in the Appendix (p. 95, 96). We observe unusual features that do not appear in any of the subsets of the data. Several states have spikes in their extremogram values every second lag (New South Wales, Queensland, South Australia). Tasmania also has spikes that occur every sixth lag. Neither the earlier or later half of the data set have patterns similar to this in the extremogram plots for the residuals after fitting the ARIMA model.

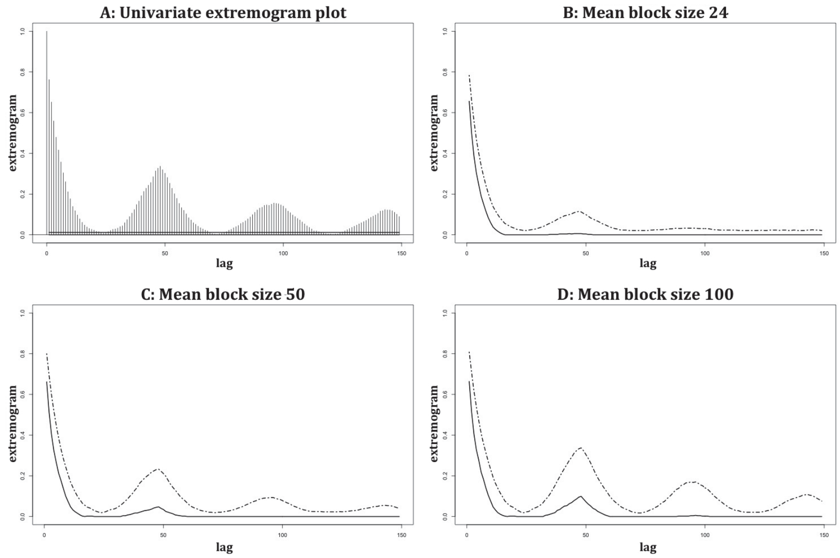

Figure 3.10: *A: Univariate extremogram plot for NSW (2009 – 2014); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

Similar to the first half of the data set, a GARCH(1,1) model was insufficient to explain the majority of the extreme event dependence because of the seasonality present in the data. The combination of a ARIMA(0,1,0) model and a GARCH(1,1) model was found to be optimal in eliminating serial dependence within the series. An example of the univariate extremogram estimated on the residuals after fitting a ARIMA and a GARCH model to South Australia can be seen in Figure 3.12. The rest of the extremogram plots can be found in the Appendix (p. 104, 106).

To conclude: the period 2012 – 2014 has less extreme event dependence than the earlier one, and most of the dependence is concentrated in the immediate lags (1 – 3). Less periodic trend and smaller extremogram values for the period 2012 – 2014 indicate a smaller probability of another extreme event occurrence in the future. Models selected for the second half of the data set are easier and contain less (or even no) periodic dependence in them. Models

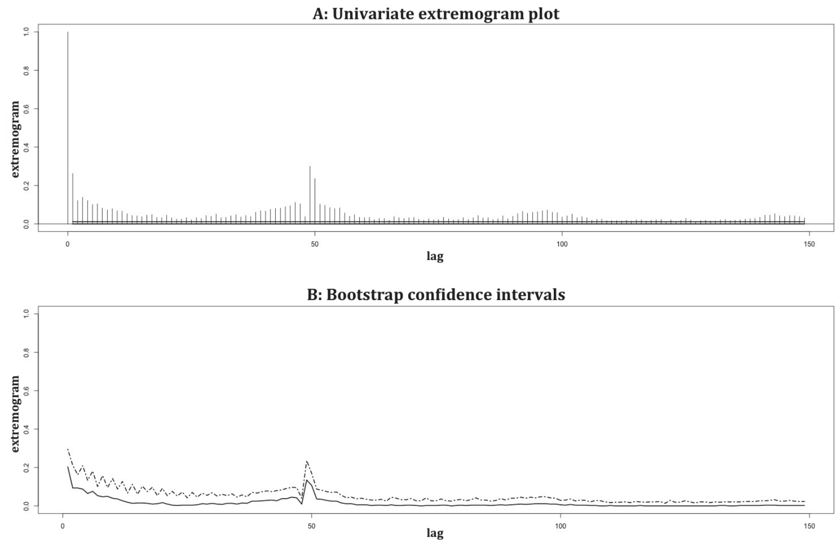

Figure 3.11: *A: Univariate extremogram plot after fitting an ARIMA(2,0,1)(1,1,1)$_{48}$ to QLD (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

for the whole data set exhibit dependence that is a mixture of the extremal dependence between the first and second periods.

As a last step in the univariate extremogram analysis, a simulation of the chosen time series models for all data sets (ARIMA and GARCH for the period 2009 – 2011, GARCH for 2012 – 2014, ARIMA(0,1,0) and GARCH for 2009 –2014) was performed to re-create a time series with features similar to those of the original data set. Unfortunately, this this analysis failed to produce reasonable results, as the estimated models were non-stationary (the sum of the coefficients is greater than 1 both for GARCH and ARIMA models). The reason for such behavior is that the vast majority of observed electricity prices (roughly 90%) are less that 50 Australians dollars. At the same time, large occasional spikes can be found in all of the data sets, and at times electricity prices can reach a value more than 8000 AUS$ which causes non-stationarity in the time series.

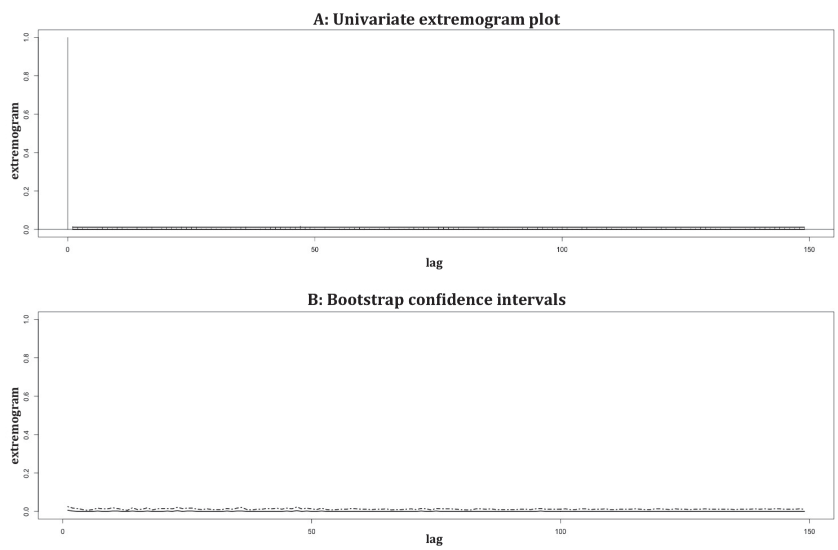

Figure 3.12: *A: Univariate extremogram plot after fitting a GARCH(1,1) to residuals estimated for first differenced series for South Australia (period: 2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

3.3 Cross extremogram analysis

The univariate extremogram provides information about the extremal dependence within one time series. As some of the states in the National Electricity Market have interconnectors between them, it allows electricity to be transported from one state to another and be sold there. This kind of market dependence could possibly influence price dependence between states (if a price in New South Wales becomes higher than in Queensland, NSW begins to increase the demand by buying its electricity which makes the price in Queensland higher). For this reason, examining cross extremogram plots will give additional information about the dependence of extreme events between two time series and indicate some similarities in two states' electricity market trends.

An example of the similarities in two states' electricity market trends is the dependence between extreme events in South Australia and New South Wales for the period 2009 – 2011 in Figure 3.13. The figure shows probabilities of high electricity prices in South Australia given high electricity prices in New South Wales for lags 0 to 150. We can see a strong periodic trend in the plot (similar to those seen in univariate extremogram plots) but here the two states are not connected. However, both states have a strong dependence on electricity import. Hence, it is plausible for them to have similar market trends.

The main trend that is observed in all the cross extremogram plots is the same as for the univariate extremograms: the second half of the data set has less dependence than the first half. A possible reason of a decrease of the

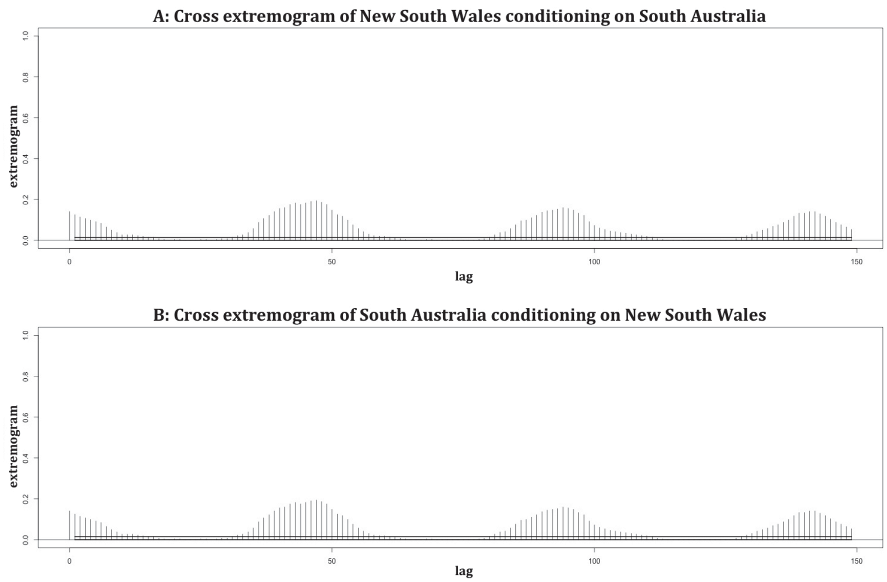

Figure 3.13: *A: Cross extremogram for NSW conditioning on SA (2009 – 2011); B: Cross extremogram for SA conditioning on NSW (2009 – 2011)*

dependence between the series as due to the climate chage policy. It forces wind and water power plants, which produce cheaper electricity, to become more popular. Hence, less trading between regions was required in the second half of the period. An example is in Figure 3.14: cross extremograms for Queensland and New South Wales for two separate parts of the data set. We can see that the extreme event dependence between two series significantly decreased, and two states almpst have no dependence in the second half of the period, even though New South Wales in the only source of electricity export for Queensland.

Another finding was that for all states the highest dependence may be found with states that have the greater number of interconnectors between them compared to states connected by the smaller number of interconnectors or not connected at all. An example of this statement is electricity prices in New South Wales conditioning on Queensland and Victoria (Figure 3.15). Only

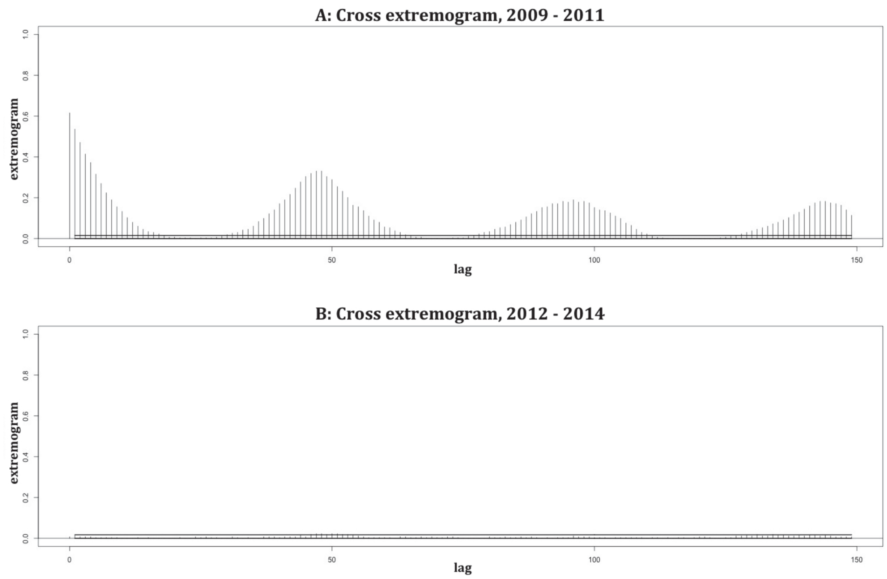

Figure 3.14: *A: Cross extremogram for QLD conditioning on NSW (2009 – 2011); B: Cross extremogram for QLD conditioning on NSW (2012 – 2014);*

one interconnector connects Victoria and New South Wales and the probability of observing a high price in New South Wales if there is a high price in Victoria is smaller then if there is a high price in Queenland. At the same time, Queensland and New South Wales are connected by two interconnectors. This is true for the rest of the states, and another example can be found in the Appendix (p. 106).

As we already discussed above, for 2012 to 2014, the extreme event dependence between states becomes significantly lower than for years 2009 – 2011. The only state that still remains dependent with the rest of the states is Victoria. Although extremogram values themselves are close to 0 and fall within 95% significance bands, bootstrap confidence intervals still show a probability of them being significant. Less periodicity can be observed, but we can still see a spike in the extremogram values at lag 48 in the bootstarp confidence intervals. All of the cross extremograms are listed in the Appendix (p. 107,

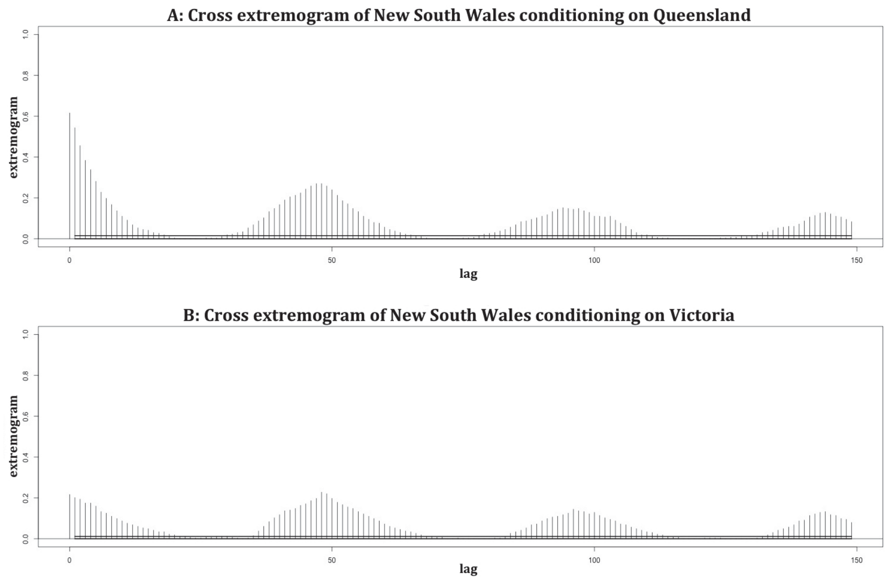

Figure 3.15: *A: Cross extremogram for NSW conditioning on QLD (2009 – 2011); B: Cross extremogram for NSW conditioning on VIC (2009 – 2011);*

108). The rest of the cross extremograms for this period are uncorrelated at all lags from 0 to 150.

The whole data set includes states whose cross extremogram shapes mimic those for the first half (years 2009 – 2011). But the cross extremogram values are affected by the second half of the data set that did not have a much dependence between the series. Hence, the cross extremogram values are lower for the whole data set compared to the first half but the relationship is still preserved. States that follow this described shape of the cross extremogram plot are New South Wales and Queensland. An example of this is in the Appendix(p. 109). Tasmania has almost no extreme event dependence for the whole data set. The only state that in connected to it is Victoria, and a negligible amount of dependence is observed between high electricity prices in these states. The cross extremogram plot can be found in the Appendix (p. 108).

Generally, all cross extremograms estimated for the whole period of

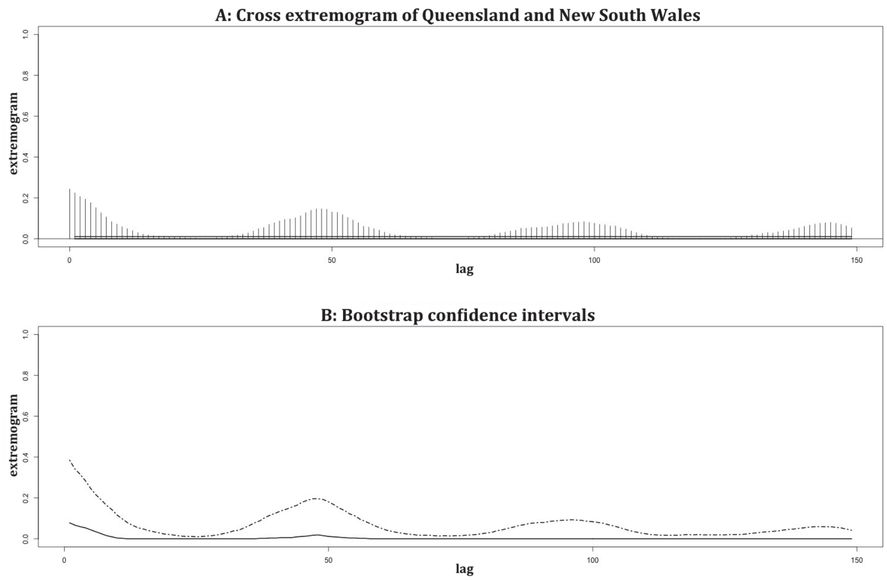

Figure 3.16: *A: Cross extremogram for QLD conditioning on NSW (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

analysis also show all expected relationship as well as the first half of the data set, i.e., states that are connected by a greater number of interconnectors tend to have stronger extreme value relationship than those that have less interconnectors or those that are not connected at all. A good example of this is South Australia. Figure 3.17 shows the cross extremogram with Victoria. The two states are connected by two interconnectors and, hence, transportation of electricity is easiest compared to all other states.

A time series model that explains most of the extreme event dependence within each of the series was estimated in the previous section. Univariate extremograms estimated for models' residuals show almost no remaining dependence. Now, cross extremograms were also estimated for the residuals after fitting time series models to see whether any dependence remains between any two series. The results show that the vast majority of the dependence between the series is also due to the underlying ARIMA and GARCH processes and all

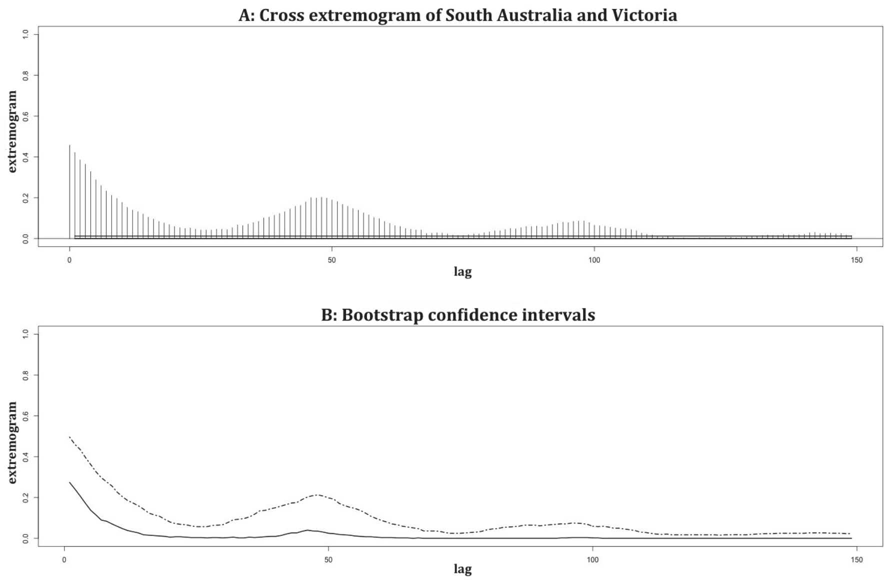

Figure 3.17: *A: Cross extremogram for SA conditioning on VIC (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

cross extremograms show almost no remaining dependence. Figure 3.18 shows the cross extremograms of Victoria and each of the states after a model was estimated for all of them are shown. The same results were obtained for all of the periods that were analyzed for all of the states.

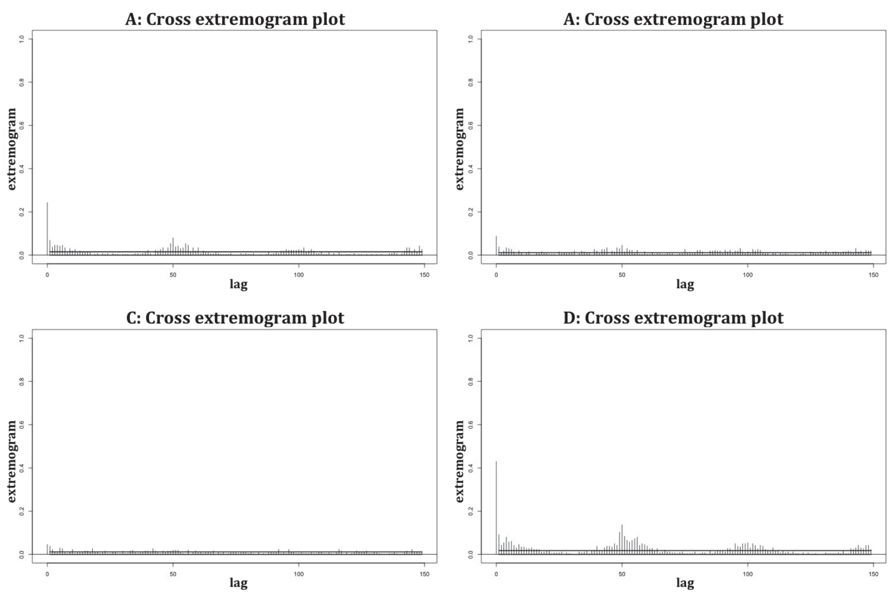

Figure 3.18: *Cross extremograms of VIC (2009 – 2014) after time series models were fitted to all of the states A: Conditioning on QLD; B: Conditioning on SA; C: Conditioning on TAS; D: Conditioning on NSW.*

Chapter 4

Quantile regression analysis of spot electricity prices of Australia

4.1 Methodology

Quantile regression (QR) was first introduced in [25]. It was proposed as an extension of Least Absolute Deviations model, a robust alternative of the Least Squares Estimate, the most widely used regression method in statistics. Quantile regression is widely used for data sets that do not meet the assumption of normally distributed error terms. An additional feature of the Quantile Regression is that it allows to estimate the dependence between two data sets not only for a median (as Least Absolute Deviations), but also for various other quantiles. Regression coefficients could change when estimated for different quantiles giving more information about the dependence between a response and predictors. In addition, [26] extends the idea of quantile regression to time series data sets. Together all of these reasons make quantile regression another useful tool to study extreme event dependence in time series data sets.

The setup of quantile regression is as follows, let $X_{T\times p} = (X_1, X_2,\ ...\ X_p)$ be a p-dimentional matrix. Also, $X_i = (X_{1i}, X_{2i},\ ...\ , X_{iT}), \quad i = 1, ... \, p$ where p the number of predictors and T is the sample size. This matrix is a matrix of predictors for regression, and is also known as a design matrix. In time series analysis lags of the predictors or the response variable can also be included in the design matrix. Also, let $\{y_t : t = 1, ..., T\}$ be a sample from a time series random variable.

As already mentioned, the most commonly used regression model in statistics is Least Squared Estimation (LSE). Its solution minimizes the sum of squares of regression error terms. The regression estimate, $\hat{\theta}$, this defined by:

$$\hat{\theta} = arg \min_{\theta} ||Y_t - X_t^{\top}\theta||^2 \tag{4.1}$$

The solution for the least squares problem is $\hat{\theta} = (X^{\top}X)^{-1}XY$. One downside of LSE is that it requires error terms to be normally distributed for statistical inference ($\epsilon_t\ N(0, \sigma^2 I)$). However, this assumption is often violated with the distribution of error terms often skewed. As LSE estimates a mean trend, it is highly affected by outliers. Another model, Least Absolute Deviations, is an L^1-norm regression models, and it is able to avoid this problem. Similar to the Least Squares regression problem, the LAD estimate, $\hat{\theta}$, satisfies the following:

$$\hat{\theta} = arg \min_{\theta} |Y_t - X_t^\top \theta| \tag{4.2}$$

The LAD estimates the median trend which is more robust compared to LSE. Quantile regression is a generalization of the LAD regression problem that estimates not only a median, but any other regression quantiles. If $\hat{\theta}(\tau^{th})$, for $0 < \tau < 1$, is a regression quantile which is defined as a solution to the following minimization problem:

$$\min_{\theta \in \mathbb{R}^{p+1}} \sum_{t=1}^{T} \rho_\tau (Y_t - X_t^\top \theta) \tag{4.3}$$

where $\rho(u) = u(\tau - I(u < 0))$ is a quantile loss function, then $\hat{\theta}(\tau^{th})$ is called the conditional quantile function of Y_t. Equation 4.3 can also be rewritten as:

$$\min_{\theta \in \mathbb{R}^{p+1}} [\sum_{t \in \{t: Y_t \geq X_t^\top \theta\}} \tau |Y_t - X_t^\top \theta| + \sum_{t \in \{t: Y_t < X_t^\top \theta\}} (1 - \tau)|Y_t - X_t^\top \theta|] \tag{4.4}$$

The solution to quantile regression, $\hat{Q}(\tau | x_t) = x_t^\top \hat{\theta}(\tau^{th})$, is called a conditional quantile function. It obtains a line of fit that has $100\tau\%$ of observations above it and $100(1 - \tau)\%$ of observations below it. Least Absolute Deviations regression is a special case of the quantile regression problem with $\tau = 0.5$. As we are focussed on extreme value dependence among high electricity prices, we consider only very high quantiles of the time series data set. Estimation of the quantile regression problem is performed using a linear programming algorithm applied to the minimization problem (4.3) above.

A model with many unimportant predictors complicates its interpretation and may decrease its prediction accuracy. [32] summarizes several variable selection methods that can be used for the quantile regression. Two of the algorithms were used to perform variable selection in this work are Least Absolute Shrinkage and Selection Operator (LASSO), proposed in [27], and Smoothly Clipped Absolute Deviation (SCAD), introduced in [16]. Both of these methods are examples of penalized regression that simultaneously perform variable selection and coefficient estimation.

The LASSO is an L^1 penalty which was originally introduced for the multiple regression model. It was first applied to quantile regression in [24]. Consider data sets $X_{T\times p} = (X_1, X_2, \ ... \ X_p)$ and $\{Y_t : t = 1, ..., T\}$ as in the previous section. Then LASSO is defined as:

$$\min_{\theta\in\mathbb{R}^{p+1}} \sum_{t=1}^{T} \rho_\tau(Y_t - X_t^\top\theta) \quad \text{subject to} \quad \sum_{j=1}^{p} |\theta_j| \leq \lambda \tag{4.5}$$

where $\lambda \geq 0$ is called a tuning parameter that controls the amount of regularization. Increasing λ leads to more shrinkage of the regression coefficients to 0 or shrinks them exactly to 0.

The SCAD penalty is usually defined as using its first derivative:

$$p'(\theta) = \lambda\Big\{I(\theta \leq \lambda) + \frac{(a\lambda - \theta)_+}{(a-1)\lambda} I(\theta > \lambda)\Big\} \tag{4.6}$$

for $\theta > 0$, where $a > 2$ and $\lambda > 0$ are two tuning parameters that control the amount of regularization, and $p(\theta)$ is the SCAD penalty function. Here $(u)_+ = \min(0, u)$ denotes the positive part of u.

Both of the penalties are symmetric around the origin. LASSO is convex, and its penalty increases linearly with a magnitude of the coefficient. This means that larger coefficients receive more shrinkage. The SCAD penalty function is also non-convex and it has the same shape as LASSO around the origin, but it becomes parallel to x-axis as it moves away from the origin. This allows SCAD to penalize large and small coefficients equally, which results in unbiased estimates for larger coefficients.

4.2 Quantile regression analysis

The method described in the previous section is now applied to the spot electricity prices of Australia. The whole data set of spot electricity prices was used (compared to Sections 3.2 and 3.3 where it was split into two halves). To estimate the quantile regression relationship, $\tau = 0.995$ was used, and a $\alpha = 0.05$ significance level was chosen as a general rule to classify a variable as a significant predictor.

4.2.1 Quantile Regression analysis of electricity prices in New South Wales

The first part of the analysis involves fitting a quantile regression model to each of the variables using only one other variable as a predictor, to find variables and lags that are significantly associated with the response. This was performed to compare whether variables that were estimated to have a significant extreme event relationship using the cross extremogram (Section 3.3) will have a significant association using an alternative method. To narrow down the choice of lags, only lags 1, 48 and 96 were used as they were found to be most dependent in the previous chapter.

Overall we found that estimated dependences disagree with what we would expect from the relationships between states' electricity markets. For example, New South Wales, was found to have a strong relationship with states South Australia at lags 1 and 96, and Victoria at lags 48 and 96. Results from the quantile regressions are summarized in the Appendix (p. 127). All of these variables are significant at the $\alpha = 0.05$ level. As was discussed in the Section 2, New South Wales and Victoria are connected by one interconnector. New South Wales is connected by two interconnectors with Queensland, but none of the mentioned lags are significant predictors. Also, New South Wales is not connected to South Australia, but lags 1 and 48 are significant predictors, which might come from similar market trends. Tasmania was found to be an insignificant predictor. Both models that only include a lag of Tasmania as a predictor would provide a similar fit to a model that only includes an intercept term.

In the next step, lags of New South Wales were added as predictors in each of the models in Table A.1. The models fit of these are given in the Appendix (p. 128). Here, only Victoria at lag 1 is still significant at the $\alpha = 0.05$ level. New South Wales at lag 1 is significant in the presence of each state (Queensland, South Australia, Tasmania, and Victoria) at lag 1. New South Wales at lag 96 is significant together with Victoria. South Australia remains significant (at $\alpha = 0.1$ level) only at lag 96. This might be due to the possibility for electricity to be resold between two states through Victoria.

Then lags 1, 48, and 96 of each state (including New South Wales itself) were included as predictors and quantile regression models were estimated for each of these. The results of models fit are presented in Tables 4.1 and 4.2. None of the models have all three significant predictors. Victoria at lag 96 is significant in the presence of the rest of the Victoria lags and South Australia at lag 96 is significant in the presence of lags of South Australia. Hence, these two variables are more useful in predicting the 0.995^{th} quantile of New South Wales than the electricity price at lag 48 in New South Wales itself. In the model that includes lags 1, 48 and 96 of New South Wales all predictors are insignificant.

Finally, lags 1, 48, and 96 of each of the states were included with lags

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
8.43743	2.03110	0.36750	0.08534
(0.89577)	*(0.09340)*	*(0.81481)*	*(0.26461)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
78.69076	0.33151	0.56527	0.65166
(0.00007)	*(0.10141)*	*(0.01240)*	*(0.00000)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
-7.76183	0.82192	1.69678	1.06722
(0.86678)	*(0.35083)*	*(0.08523)*	*(0.02199)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
126.26276	0.00526	-0.00824	0.05014
(0.00000)	*(0.94502)*	*(0.90980)*	*(0.83152)*

Table 4.1: *Estimated QR models for NSW using each of the other states at lags 1, 48 and 96 as predictors, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
-35.86229	3.34828	0.55775	-0.00404
(0.55957)	*(0.14895)*	*(0.27581)*	*(0.90427)*

Table 4.2: *Estimated QR models for NSW using NSW at lags 1, 48 and 96 as predictors, $\tau = 0.995$. p-values given in parenthesis.*

1, 48 and 96 of New South Wales. The results are presented in the Appendix (p. 129). It is evident that the combination of all variables makes all of the predictors included in a model insignificant, except New South Wales. Lag 1 of New South Wales is significant in the presence of Victoria and Tasmania, and lag 48 is significant in the presence of Queensland and Victoria. When lags 1, 48 and 96 of New South Wales are combined together with the same lags of South Australia, New South Wales is significant at lag 96. None of the predictors are significant in this model, not even South Australia at lag 96 that was significant in the previous models.

We can see that estimated coefficients for lags of New South Wales are consistent among all of the models. However, the other predictor's coefficients are notably different, that is, some change their signs when New South Wales is added to the model. An example of such predictors are Queensland at lags 1 and 48.

All of the states, excluding Tasmania, strongly depend on the periodic market trend, which might affect the models estimated above. To eliminate this general tendency of the electricity market, all of the series were deseasonalized. Usually, a Least Squares model with indicator variables for each period is used to deseasonalized a data set, but, as was mentioned in Section 4.1, the model is highly affected by outliers, and all data sets involved in the analysis have multiple spikes. Hence, to increase efficiency, a quantile regression model with $\tau = 0.5$ was used instead (with indicator variables for each 30 minute interval as predictors) to eliminate the daily periodic trend. Then, all the QR models were re-estimated on the resulting residuals. Several new results were found.

When the periodic trend is removed from all of the variables, for models with one predictor, Queensland at lag 48, South Australia as lags 1 and 48 and Victoria at lag 96 become important predictors. Lags 1 and 48 of Victoria are no longer useful predictors for electricity prices in New South Wales when used alone. The results are summarized in the Appendix (p. 130).

For models with one predictor and a corresponding lag of New South Wales the latter is significant at lags 1 and 48 when combined with South Australia at lags 1 and 48, respectively. None of the predictors are significant in this type of models. The results are summarized in the Appendix (p. 131)

As for models with lags 1, 48 and 96 from each state, the model with lags of Victoria and Queensland finds that two predictors (lags 1 and 96 of Queensland and lags 1 and 48 of Victoria) are significant for the deseasonalized data set in each of the models. It is surprising that Queensland at these lags is not significant in any other combinations of predictors. The results are shown in Table 4.3. Also, none of the New South Wales lags become significant when the periodic trend is removed (see Table 4.4).

For models containing lags 1, 48 and 96 of each state and the same lags

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
47.89946	2.02937	0.36881	0.08624
(0.00000)	*(0.00001)*	*(0.55244)*	*(0.00000)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
98.74169	0.33173	0.56544	0.65199
(0.00000)	*(0.00000)*	*(0.21630)*	*(0.00005)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
75.05567	0.82311	1.69772	1.06755
(0.00000)	*(0.00000)*	*(0.00000)*	*(0.36499)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
91.77145	0.00433	-0.00799	0.04386
(0.00743)	*(0.95153*	*(0.97277)*	*(0.91649)*

Table 4.3: *Estimated QR models for NSW using each of the other states at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	NSW (lag 1)	NSW (lag 48)	NSW (lag 96)
58.88936	3.34964	0.55901	-0.00503
(0.00000)	*(0.22688)*	*(0.20571)*	*(0.95275)*

Table 4.4: *Estimated QR models for NSW using NSW at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

of New South Wales the results, which can be found in the Appendix (p. 132), are similar to those of the original data set. The only difference is that lag 1 of New South Wales becomes significant for the model containing Victoria, and New South Wales at lag 48 becomes significant in the presence of Tasmania lags.

4.2.2 Quantile Regression analysis of electricity prices in Queensland

We next analyzed Queensland. The 99.5^{th} quantile was used as the τ parameter for quantile regression. Queensland is connected only to New South Wales by two interconnectors, and these states' markets are highly dependent. Nevertheless, these two states do not have a strong association between them. Other variables are significant predictors for electricity prices in Queensland: South Australia at lags 1 and 96, and Victoria at lags 48 and 96. The results are in teh Appendix (p.133). Some of these dependencies may be as a result of a general market trend that affects most of the states. This was examined separately using deseasonalized data set later in this section. Following the same scheme of analysis as was used for New South Wales, lags of electricity prices in Queensland were added to the QR models above. The results are summarized in the Appendix (p. 134). We can see that Queensland at lag

1 is significant in every model where it was used. The model that contains Victoria and Queensland at lag 96 has Queensland significant at lag 96. For all of the models, none of the predictor variables is significant. Hence, similar to New South Wales, past electricity prices of Queensland are more important in estimating a current price compared to prices in other states.

Models in Table 4.5 contain lags 1, 48 and 96 of each of the other states and Queensland itself in Table 4.6. From the Table 4.6 it is evident that none of the lags of Queensland are significant when combined together in one model. Among models containing other states as predictors (Table 4.5) lag 1 of South Australia is not significant in the presence of other lags of South Australia. However, lag 48 becomes significant when used together with lags 1 and 96 of South Australia. In the model with New South Wales none of the variables is significant at the $\alpha = 0.05$ level. Only lag 48 of Victoria shows significance at the $\alpha = 0.1$, which is surprising as Queensland is not connected to Victoria. In the presence of all the lags of Victoria, lag 96 was the only significant lag in the model. None of the Tasmania lags are significant when they are combined in the same model.

The combination of lags 1, 48 and 96 of Queensland and each of the other states in NEM does not reveal any new dependencies. With the presence of all six predictors, lags 1 of Queensland are significant in all of the models except the one containing South Australia. Lag 48 of Queensland is significant only in the presence of New South Wales. These results are summarized in teh Appendix (p. 135).

Several other relationships between Queensland and the rest of the states were found after the periodic trend was removed from all of the time series.

Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
8.43743	2.03110	0.36750	0.08534
(0.89577)	*(0.09340)*	*(0.81481)*	*(0.26461)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
78.69076	0.33151	0.56527	0.65166
(0.00007)	*(0.10141)*	*(0.01240)*	*(0.00000)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
-7.76183	0.82192	1.69678	1.06722
(0.86678)	*(0.35083)*	*(0.08523)*	*(0.02199)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
126.26276	0.00526	-0.00824	0.05014
(0.00000)	*(0.94502)*	*(0.90980)*	*(0.83152)*

Table 4.5: *Estimated QR models for QLD using each of the other states at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
-35.86229	3.34828	0.55775	-0.00404
(0.55957)	*(0.14895)*	*(0.27581)*	*(0.90427)*

Table 4.6: *Estimated QR models for QLD using QLD at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Three new variables become significant for the deseasonalized data set when used as a predictor alone: New South Wales and South Australia each at lag 48 and Tasmania at lag 96. These relationships are relatively weak compared to the general market trend affecting all of the series. Thus, the two predictors are not significant for the original data set. One other variable (Victoria at lag 96) is no longer a valuable predictor. Hence, the significance of these variables for the original data set is due to the seasonal market trend. The results for the quantile regression fit for each on the variables and lags for a deseasonalized data set are presented in the Appendix (p. 136).

Depersonalizing the data set also shows that when lags 1, 48 and 96 of each state are combined together, New South Wales and South Australia are significant at lags 1 and 96. Victoria is significant at lags 1 and 48. Apart from South Australia, which is significant at each lag when used alone as a predictor, lags of New South Wales and Victoria, which are not significant when used alone, now show significance when combined together. The opposite is also true, i.e. lags 96 of New South Wales and Victoria are significant when used alone but become insignificant when combined together with other lags of the same state.

For the remaining analysis performed on the deseasonalized data set the

Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
47.89946	2.02937	0.36881	0.08624
(0.00000)	*(0.00001)*	*(0.55244)*	*(0.00000)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
98.74169	0.33173	0.56544	0.65199
(0.00000)	*(0.00000)*	*(0.21630)*	*(0.00005)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
75.05567	0.82311	1.69772	1.06755
(0.00000)	*(0.00000)*	*(0.00000)*	*(0.36499)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
91.77145	0.00433	-0.00799	0.04386
(0.00743)	*(0.95153)*	*(0.97277)*	*(0.91649)*

Table 4.7: *Estimated QR models for QLD using each of the other states at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

results are similar to those of the original data set. The quantile regression models fit are presented in the Appendix (p. 137 – 139).

4.2.3 Quantile Regression analysis of electricity prices in South Australia

South Australia is the state that is heavily dependent on market conditions as it imports up to 25% of energy consumed within the state. The major source of import to the state is electricity produced in Victoria. For the quantile regression with a single predictor (Table 4.8), Victoria is significant is lags 1 and 48. These are, as expected, the only two significant predictors for South Australia.

The addition of corresponding lags of South Australia reveals results to

Intercept	**QLD (lag 1)**	**Intercept**	**NSW (lag 1)**
135.47807	2.11332	-68.05331	7.74540
(0.12345)	*(0.17252)*	*(0.94861)*	*(0.80880)*
Intercept	**QLD (lag 48)**	**Intercept**	**NSW (lag 48)**
210.32955	0.43123	28.45065	3.97437
(0.00000)	*(0.43123)*	*(0.89325)*	*(0.49000)*
Intercept	**QLD (lag 96)**	**Intercept**	**NSW (lag 96)**
218.47606	0.32723	165.36115	1.28745
(0.03322)	*(0.80739)*	*(0.64301)*	*(0.86280)*
Intercept	**TAS (lag 1)**	**Intercept**	**VIC (lag 1)**
234.79405	0.02628	-27.70381	5.90046
(0.00783)	*(0.98666)*	*(0.58966)*	*(0.00000)*
Intercept	**TAS (lag 48)**	**Intercept**	**VIC (lag 48)**
237.73996	0.01787	9.59829	4.67917
(0.00086)	*(0.98619)*	*(0.76088)*	*(0.00000)*
Intercept	**TAS (lag 96)**	**Intercept**	**VIC (lag 96)**
238.74236	0.01755	179.12949	1.03707
(0.07047)	*(0.99396)*	*(0.35350)*	*(0.84696)*

Table 4.8: *Estimated QR models for SA using one predictor at lag h (specified separately for each model),* $\tau = 0.995$. *p-values given in parenthesis.*

the two previous states. For the models that contain two predictors, only lags of South Australia itself are significant. South Australia at lag 1 is significant in the presence of Queensland, New South Wales or Tasmania. Lags 48 and 96 of South Australia are significant when used together with New South Wales or Tasmania. The results can be found in the Appendix (p. 140).

In the presence of other lags of the same variable (lags 1, 48 and 96 of each state) only Victoria at lag 1 is significant. Even past values of South Australia itself are not useful predictors when they all are used in the same model. Results of models fit are can be found in the Appendix (p. 141). Nevertheless, when lags of each state are combined with lags in South Australia, lag 1 of South Australia is significant in each on the models except the model with Tasmania. Estimated coefficients are summarized in the Appendix (p. 142).

As was already discussed, South Australia is heavily dependent on market conditions. For this reason, when the daily periodic trend is removed, the only two significant predictors are still Victoria at lags 1 and 48, which is the same result as the original data set. This result can be found the the Appendix (p. 143). Among models that contain lags 1, 48 and 96 of each state, none of the predictors are significant for the deseasonalized data set (Table 4.9). Only the combination of past prices of South Australia at lags 1, 48 and 96 shows a significant association with lag 1 (Table 4.10). The rest of the models are shown in the Appendix.

4.2.4 Quantile Regression analysis of electricity prices in Victoria

The last state that is located on Australian mainland and that has not yet been discussed is Victoria. It fulfills the majority of its electricity demand with generators within the state, but is depend on electricity proves in other states due the exportation. This can be shown using quantile regression models that

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
171.61551	1.54439	0.22975	0.16860
(0.01312)	*(0.40672)*	*(0.81737)*	*(0.51621)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
137.83825	6.99014	-0.00951	-0.00974
(0.88119)	*(0.85805)*	*(0.96888)*	*(0.97038)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
124.64341	5.46550	0.22955	0.17430
(0.01305)	*(0.05088)*	*(0.93432)*	*(0.23944)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
197.78284	0.02502	0.00861	0.02059
(0.00000)	*(0.86025)*	*(0.98126)*	*(0.96594)*

Table 4.9: *Estimated QR models for SA using each of the other states at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	SA (lag 1)	SA (lag 48)	SA (lag 96)
61.44341	1.71259	0.28232	0.06824
(0.04554)	*(0.00073)*	*(0.60840)*	*(0.81864)*

Table 4.10: *Estimated QR models for SA using SA at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

were estimated using a single predictor (Table 4.11). Victoria is connected to New South Wales and South Australia, and lag 48 of New South Wales and lags 1 and 48 of South Australia are significant. Even though Queensland is not connected to Victoria, lags 1 and 48 are also significant. But for models that include also corresponding lags of Victoria, only Victoria at lag 96 is significant in the presence of New South Wales or South Australia, and for models with lags 1, 48 and 96 of each state only South Australia at lag 1 is a significant predictor. The results of the models are in the Appendix (p. 146).

For the next step, lags 1, 48 and 96 of each state were combined with Victoria, and models with six predictors were fitted to the data set containing electricity prices of Victoria. No novel dependencies are revealed in these mod-

Intercept	**QLD (lag 1)**	**Intercept**	**SA (lag 1)**
81.25383	0.70449	38.07286	1.20510
(0.00000)	*(0.01628)*	*(0.03223)*	*(0.00000)*
Intercept	**QLD (lag 48)**	**Intercept**	**SA (lag 48)**
103.89164	0.34708	69.77077	0.83138
(0.00000)	*(0.00000)*	*(0.00000)*	*(0.00000)*
Intercept	**QLD (lag 96)**	**Intercept**	**SA (lag 96)**
112.62040	0.18149	112.55899	0.14706
(0.00000)	*(0.12408)*	*(0.00000)*	*(0.66760)*
Intercept	**TAS (lag 1)**	**Intercept**	**NSW (lag 1)**
112.57349	0.08745	-17.63467	2.98033
(0.00000)	*(0.63701)*	*(0.97010)*	*(0.81735)*
Intercept	**TAS (lag 48)**	**Intercept**	**NSW (lag 48)**
117.73516	0.04697	8.18424	2.42424
(0.00002)	*(0.89053)*	*(0.83829)*	*(0.01176)*
Intercept	**TAS (lag 96)**	**Intercept**	**NSW (lag 96)**
118.51601	0.03591	58.12300	1.38170
(0.82620)	*(0.98298)*	*(0.13663)*	*(0.11161)*

Table 4.11: *Estimated QR models for VIC using one predictor at lag h (specified separately for each model), $\tau = 0.995$. p-values given in parenthesis.*

els. The only significant predictor is Victoria at lag 1 in the model containing Queensland. The results can be found in the Appendix.

After the periodic trend is removed from the data set, the electricity price in New South Wales is no longer a significant predictor for the deseasonalized data set when used as the only predictor. Overall, the dependence structure between Victoria and the rest of the states for the deseasonalized data set stays similar to the original data set. This can bee seen in Tables 4.12 and 4.13 below and in the Appendix (p. 149).

Intercept	QLD (lag 1)	QLD (lag 48)	QLD (lag 96)
66.35211	0.70468	0.14516	0.07424
(0.00000)	*(0.00000)*	*(0.45688)*	*(0.45567)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
46.75473	1.20408	0.20897	0.01167
(0.00236)	*(0.00000)*	*(0.15906)*	*(0.90369)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
45.64943	2.80656	-0.00184	-0.00316
(0.12160)	*(0.73085)*	*(0.99035)*	*(0.96917)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
80.48913	0.08289	0.01988	0.01280
(0.00000)	*(0.69790)*	*(0.43768)*	*(0.87134)*

Table 4.12: *Estimated QR models for VIC using each of the other states at lags 1, 48 and 96 as predictors for a deseasonalized data set, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	VIC (lag 1)	VIC (lag 48)	VIC (lag 96)
39.93229	2.38398	0.20175	0.03972
(0.00000)	*(0.12993)*	*(0.83911)*	*(0.83593)*

Table 4.13: *Estimated QR models for VIC using VIC at lags 1, 48 and 96 as predictors for a deseasonalized data set, $\tau = 0.995$. p-values given in parenthesis.*

4.2.5 Quantile Regression analysis of electricity prices in Tasmania

The last state in Australia that is a part of NEM is Tasmania. As it is located on an island and has only one interconnector with Victoria, the state is not expected to have strong dependence with the rest of the states. This is supported by quantile regression results with only one predictor. None of the variables were found to be significant. These results are presented in the Appendix (p. 152). Nevertheless, Victoria at lag 1 is significant in the presence of lags 48 and 96 of the same state (Table 4.14). Lag 1 of Tasmania itself is significant in the combination of lags 48 and 96 of the same state (Table 4.15) and when these three lags are combined with the same lags of each other state.

These results are summarized in the Appendix (p. 154). None of the states are significant in any other models. This confirms the assumption that the state's electricity market is not highly dependent with other states' markets.

The vast majority of combinations of variables that were created by

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
143.19302	0.12374	-0.00833	-0.00684
(0.00000)	*(0.70899)*	*(0.94939)*	*(0.97168)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
138.65708	0.13377	0.16817	-0.00122
(0.97571)	*(0.95074)*	*(0.94549)*	*(0.99967)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
96.47898	0.62663	0.81331	-0.01065
(0.00027)	*(0.34224)*	*(0.00000)*	*(0.90975)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW(lag 96)**
122.64117	0.52743	-0.01252	0.04312
(0.10209)	*(0.79350)*	*(0.96092)*	*(0.38582)*

Table 4.14: *Estimated QR models for TAS using each of the other states at lags 1, 48 and 96 as predictors, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
41.70902	0.91080	0.32790	0.05093
(0.00092)	*(0.00000)*	*(0.00485)*	*(0.75137)*

Table 4.15: *Estimated QR models for TAS using TAS at lags 1, 48 and 96 as predictors, $\tau = 0.995$. p-values given in parenthesis.*

grouping variables do not give significantly good results. It is possible that depending on the market conditions a state can be dependent on prices in several different states at several different lags. To carry out variable selection on the predictors, SCAD and LASSO penalties were introduced to shrink the least important variables' coefficients down to 0. Both of the regularization methods did not perform well on any of the data sets leaving all of the coefficients non-zero and shrinking all of them equally for higher λ. Hence, we do not include their results.

Chapter 5

Cross-quantilogram analysis of spot electricity prices of Australia

5.1 Methodology

The cross-correlation function (in Section 3.1.1) is an important tool in time series analysis. It estimates the dependence between two time series. It can be used to find predictors in a regression model. However, similar to the autocorrelation function, cross-correlation estimates the dependence for the center of distributions of both of the series. It is useful when the distributions of both of the series are normal. When distributions are not normal and (or) when the correlation between conditional quantiles is of the most interest, the cross-quantilogram is a useful tool to estimate dependence between the series and for selecting predictors for quantile regression.

The quantilogram introduced in [22] is a quantile analogue of the autocorrelation function described in Section 3.1.1. The cross-quantilogram focusses on measuring dependencies between two time series $\{Y_{1,t}\}_{t=1...T}$ and $\{Y_{2,t}\}_{t=1...T}$, at two conditional quantiles τ_1 and τ_2. It was introduced in [15] where it is defined as a measure of the dependence between different parts of the distributions of two strictly stationary time series (cross-correlation between different conditional quantiles). It is defined as:

$$\rho_{\tau_1,\tau_2}(h) = \frac{E\big[\psi_{\tau_1}(Y_{1t} - q_{1,t}(\tau_1))\psi_{\tau_2}(Y_{2,t-h} - q_{2,t-h}(\tau_2))\big]}{\sqrt{E\big[\psi^2_{\tau_1}(Y_{1t} - q_{1,t}(\tau_1))\big]E\big[\psi^2_{\tau_2}(Y_{2,t-h} - q_{2,t-h}(\tau_2))\big]}} \tag{5.1}$$

for $\tau_1, \tau_2 \in (0,1)$; $h \in \{..., -1, 0, 1, ...\}$. Here $q_{1,t}(\tau_1)$ and $q_{2,t-h}(\tau_2)$ are conditional quantile functions as in the Chapter 4 for $\{Y_{1,t}\}_{t=1...T}$ and $\{Y_{2,t}\}_{t=1...T}$, respectively, and $\psi_\tau(u) = I[u < 0] - \tau$. As τ_1 and τ_2 are not necessarily equal, hence, $\hat{\rho}_{\tau_1,\tau_2}(h) \neq \hat{\rho}_{\tau_2,\tau_1}(h)$.

An inferential procedure for the cross-quantilogram values is described in [15]. The paper gives an overview of the stationary bootstrap procedure used to estimate confidence intervals to infer significance of the cross-quantilogram values (unlike the cross extremogram where stationary bootstrap was used to construct confidence intervals for the cross extremogram values). In this case two time series are re-sampled separately from each other. This procedure allows to keep the serial dependence within each series but breaks the dependence between the two series. A cross-quantilogram value that falls outside of the stationary bootstrap confidence bands is considered to be statistically significant.

The interpretation of the cross-quantilogram is similar to the interpretation of the cross extremogram for upper quantiles. For a positive quantilogram value, which falls outside of the stationary bootstrap confidence bands, if at time t a time series $Y_{2,t}$ is larger than its τ_2^{th} conditional quantile, the chance of a time series $Y_{1,t}$ to be above its τ_1^{th} conditional quantile in h time periods is significantly high. Similarly, for a negative cross-quantilogram value, two time series are likely to be in different parts of their distributions (i.e., one time series above its conditional quantile, and another one below).

5.2 Cross-quantilogram analysis

5.2.1 Cross-quantilogram of New South Wales

The cross-quantilogram was estimated for New South Wales and each of the other states for the whole data set: from 2009 – 2014. The 99.5^{th} quantile was used again in the analysis.

Cross-quantilograms for lags 0 to 150 are shown in Figures 5.1 conditioning on Queensland for New South Wales and in the Appendix (p. 159, 160) conditioning on the rest of the states. It is evident that the plots have similar shapes to the corresponding cross extremograms. However, the cross-quantilogram represents the correlation between conditional quantiles of each of the series at lag h. Hence, it allows negative values unlike the cross extremogram that represents the probability given an extreme electricity price in New South Wales at time t to observe an extreme electricity price in each of the states at time $(t + h)$. Consequently, the cross extremogram values are all non-negative.

In general, most of the cross-quantilograms for New South Wales (ex-

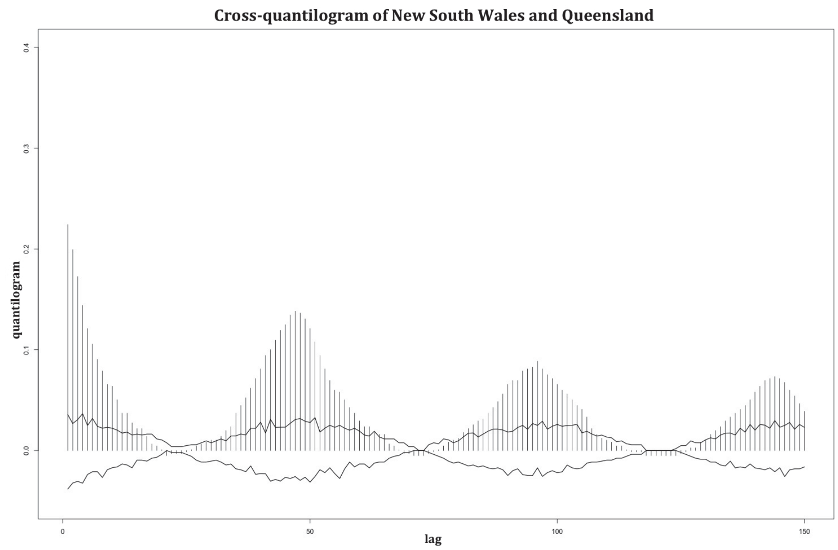

Figure 5.1: *A: Cross-quantilograms for NSW conditioning on QLD (2009 – 2014).*

cept conditioning on Tasmania) have a periodic trend with period 48, which means that lags 48, 96, 144 of these states are more dependent with a price in New South Wales at time t. Confidence intervals for the cross-quantilograms were estimated using the stationary bootstrap algorithm with 100 iterations for each lag and have the same periodic trend. This means lags 24, 72, 120 have significant negative correlation with New South Wales.

Similar to the cross extremograms, New South Wales and Queensland have the highest dependence between states. This is most visible for lags 0 to 5. At the same time, Victoria has the strongest dependence with New South Wales at lag 48 with the cross-quantilogram value above 0.2. Remember that two of these variables (Queensland at lag 1 and Victoria at lag 48) are the most significant predictors for New South Wales in quantile regression analysis. The same result was obtained in Chapter 4. South Australia has a similar relationship with New South Wales as Queensland for most of the lags, except the first few. As was already discussed in previous sections, most of the dependence between New South Wales and South Australia is caused by similar market conditions and the need to import electricity from Victoria when electricity prices are high within each of the states. Tasmania is mostly unrelated with New South Wales. The two lags that have highest significant cross-quantilogram values are lag 8 and lag 40.

5.2.2 Cross-quantilogram of Queensland

We analyzed Queensland with the cross-quantilogram. It also shows the same result that was obtained with the cross extremogram. The state that it has the strongest relationship with New South Wales at a 99.5 quantile (Figure 5.2). Queensland only has an interconnector to New South Wales. Thus, trading is available only with this state, which results in the extreme value dependence. The remaining cross-quantilograms are presented in the Appendix (p. 161, 162). It is evident that Queensland is also dependent on South Australia. The highest value in the cross-quantilogram between Queensland and South Australia is obtained at lag 48. As we saw in the previous chapters most of the mainland states are dependent, even when they not connected by an interconnector. The potential for electricity to be re-sold through several states might also introduce dependencies between regions that are not connected. It is clear that the dependence is not strong between Queensland and Victoria. The cross-quantilogram between these two states is roughly 0.06 for lags 0 and 48, and less for the rest of the lags. Similar to the results of New South Wales, Queensland at lag t is negatively dependent with other mainland states at lag 24, 72, and 120, and has also almost no dependence with Tasmania.

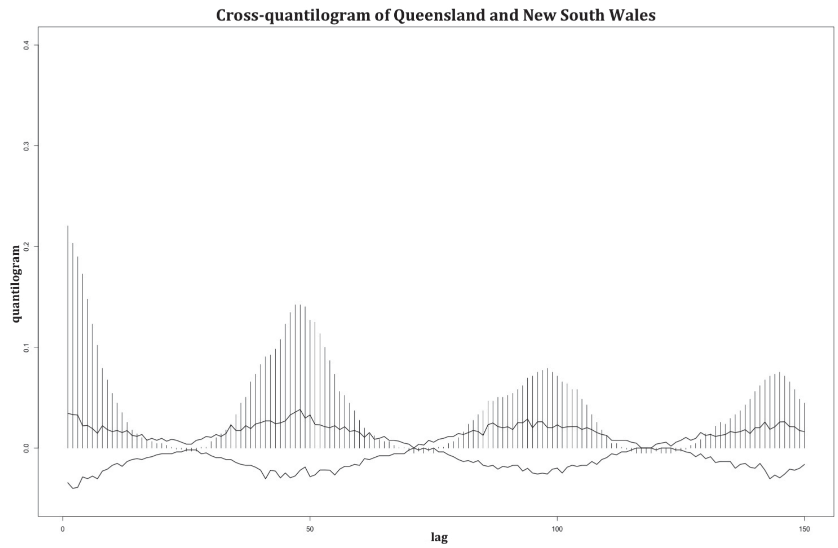

Figure 5.2: *A: Cross-quantilograms for QLD conditioning on NSW (2009 – 2014).*

5.2.3 Cross-quantilogram of South Australia

South Australia imports electricity form Victoria. Thus, it can be assumed that Victoria should have the strongest dependence with South Australia compared to other states. This fact is proved in Figure 5.3. Compared to the rest of the cross-quantilogram figures of South Australia in the Appendix (p. 162, 163), the cross-quantilogram of South Australia and Victoria has a clear periodic trend and the highest dependence for lags 1 to 150 compared to the rest for the states. Lags that have the highest correlation with South Australia at time t are the few initial lags (0 to 7) and lag 48. Also, South Australia at time t is not negatively dependent on any of the lags up to 117. Most of the lags included in the plot show significance. However, Victoria at lag 1 is not significant when used as a single predictor in a Quantile Regression model, and lag 48 is significant despite the lower correlation compared to lag 1.

Unlike Queensland and New South Wales, South Australia has some

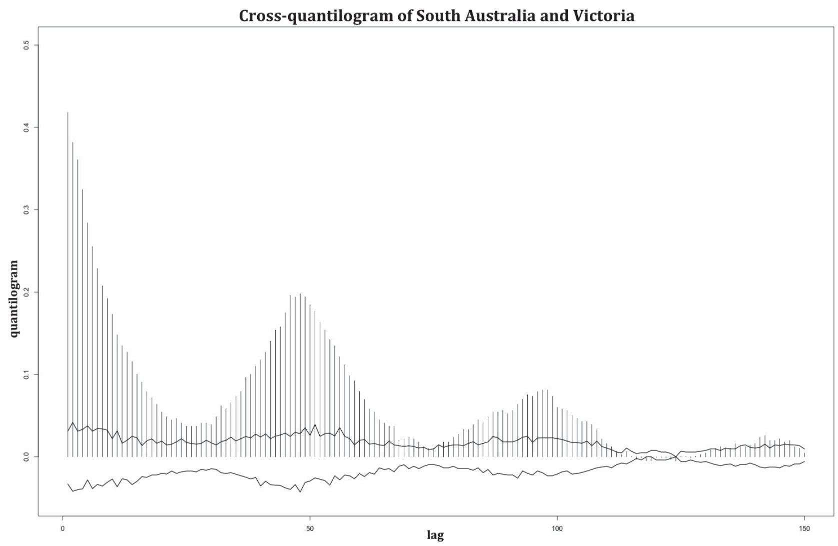

Figure 5.3: *A: Cross-quantilograms for SA conditioning on VIC (2009 – 2014).*

significant cross-quantilogram values with Tasmania. We can see that lags 0-3 and lag 48 have the highest cross-quantilogram values (roughly 0.06) that correspond to the highest extreme value between South Australia at time t and Tasmania. The cross-quantilograms of South Australia with Queensland and New South Wales do not have a great number of significant lags in cross-quantilogram the plots. The results are presented in the Appendix (p. 163).

5.2.4 Cross-quantilogram of Victoria

We look at Victoria next. Victoria is a major exporting state and is assumed to have the strongest dependence with those states that import electricity, i.e., New South Wales and South Australia.

We can see in Figure 5.4 that Victoria has a strong dependence with

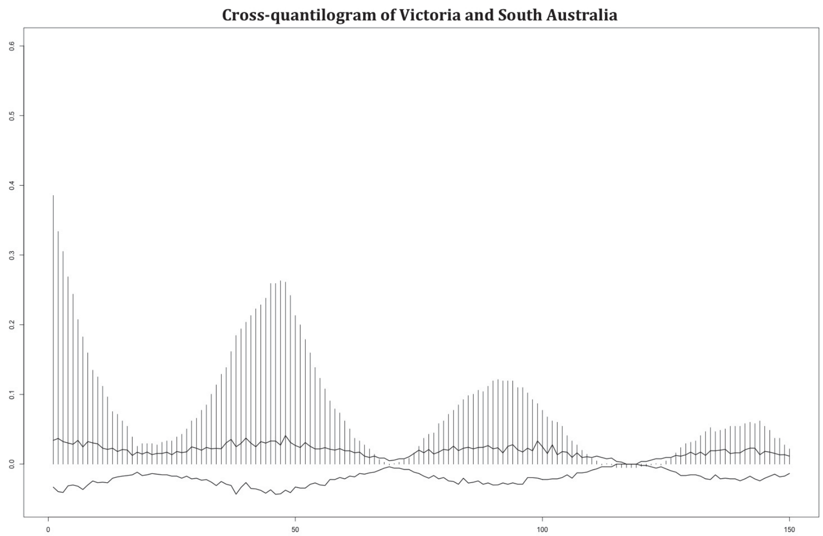

Figure 5.4: *A: Cross-quantilograms for VIC conditioning on SA (2009 – 2014).*

South Australia. SA is connected only with Victoria, hence, it heavily depends on the import of cheaper electricity generated by water power plants from Victoria. Unlike South Australia, New South Wales is also connected to Queensland and, hence, has a competitive option for importing. It is evident that the dependence between New South Wales and Victoria, which is presented in the Appendix (p. 164), is smaller compared to the the dependence between South Australia and Victoria (Figure 5.4).

Tasmania and Queensland have a weak extreme value dependence with Victoria compared to two other states. For these, only lags 1 and 48 show significant dependence. Queensland is not connected to Victoria and is not adjacent to it. Hence, the weak dependence is understandable. The plot can be found in the Appendix (p. 164). Despite being connected only with Victoria, Tasmania is able to satisfy most it's electricity demand with generators within the state. For this reason, Victoria is not significantly dependent on it. However, the extremogram analysis in Chapter 3 shows that electricity prices in Tasmania are more dependent with the other states than the rest of the states are dependent on electricity prices of Tasmania. We check this in the next section.

5.2.5 Cross-quantilogram of Tasmania

Finally, we consider the cross-quantilograms of Tasmania. The analysis shows more symmetric relationship between the states. Tasmania is more dependent with South Australia and Victoria compared to New South Wales and Queensland. The shapes of the plots and the significant lags are also similar to the cross-quantilograms conditioning on Tasmania discussed earlier in this section. All of the cross-quantilogram plots can be found in the Appendix.

In general, not all the states that are significant and highly dependent with one another are useful predictors for quantile regression. For example, Victoria at lag 48 is a has a smaller p-value for a quantile regression when New South Wales is the response variable then Victoria at lag 1 but it has a higher cross-quantilogram value. However, all the results obtained with the cross-quantilogram confirm the quantile regression results when only one predictor was used. Lags with the highest cross-quantilogram values have smaller p-values in the quantile regression. The same results were obtained with the cross extremogram analysis, as plots obtained using both methods have similar shapes. Hence, the selection of predictors made with the cross-extremogram is confirmed by the cross-quantilogram estimated for the same quantile.

The cross-quantilogram measures predictability. Hence, it is a more useful tool for selecting the most important predictors for a given time series than the extremogram. For the given data set, the cross extremogram and cross-quantilogram results with respect to finding the best predictors for quantile regression are equivalent because lags with the highest correlation between time t and $(t + h)$ are positive. Negative cross-quantilogram values result in a cross extremogram value that is close to 0. If the best predictors were negatively correlated with the response, the cross extremogram as it was used in Chapter 3 would miss them. However, another option is available for the extremogram in order to capture negative associations. We can find a probability of an extreme event occurring in a lower (upper) tail of one time series given another extreme event in an upper (lower) tail of another time series. This, however, is a more complicated way of selecting predictors.

Chapter 6

The copula method for estimating extreme value dependence

6.1 Methodology

This chapter introduces a non-parametric approach for the estimation of extreme value dependence based on bivariate copulas. Copula models are used to estimate dependence between distributions whose marginals are not normal. They are a commonly used tool in statistics for the estimation of joint and conditional distribution functions. Also, they are widely used to calculate the upper tail dependence coefficient, which is another method for estimating extreme value dependene. For this reason, copulas have become a popular method to estimate extreme value dependence.

Before introducing copulas, we first recall a non-parametric measure of association between two data sets – Kendall's rank correlation coefficient (Kendall's τ) [14], is defined as:

$$\tau = \frac{N_c - N_d}{n(n-1)/2} \tag{6.1}$$

where N_c is the number of concordant pairs (i.e., when $x_i > x_j$ and $y_i > y_j$ or $x_i < x_j$ and $y_i < y_j$), N_d is the number of discordant pairs (i.e., when $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$). Similar to the correlation coefficient, two data sets have positive association if $\tau > 0$ and negative association if $\tau < 0$. If $\tau = 0$ it is assumed that two data sets have no association.

Now, a bivariate copula is a bivariate distribution function with uniform margins. Copulas are widely used to the estimate joint distribution function using Sklar's theorem [19]:

Theorem 1. *Let F be a bivariate distribution function with two marginals - F_1 and F_2. Then there exists a bivariate copula such that for every $x \in \mathbb{R}$:*

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \tag{6.2}$$

If both margins are continuous, then the copula function is unique.

According to Theorem 1, every joint distribution function can be decomposed into two marginals and a copula function. A class of copulas called extreme-value copulas [17] models extreme event dependence. Following the definition in [17], a bivariate copula is an extreme-value copula if it satisfies the following:

$$C^*(u_1, u_2) = \lim_{n \to \infty} C(u_1^{1/n}, u_c^{1/n})^n \tag{6.3}$$

for $[u_1, u_2] \in [0, 1]^2$. Another important concept in copula theory is max-stable copulas. A bivariate copula is max-stable if the following holds:

$$C(u_1^{1/m}, u_c^{1/m})^m = C(u_1, u_2) \tag{6.4}$$

for $m \geq 1$ and for $[u_1, u_2] \in [0, 1]^2$.

A copula is an extreme-value copula if and only if it is max-stable.

Many types of copulas exist in statistical theory. We focus on the Gumbel copula [8]. A Gumbel copula is an extreme-event copula, and it requires the estimation of only one parameter. It is also widely used to study the upper tail dependence. In a bivariate case, the Gumbel copula for the bivariate distribution function $F(x_1, x_2) = \exp(-(\exp(-\theta x_1) + \exp(-\theta x_2))^{1/\theta})$ is given by:

$$C(u_1, u_2) = \exp(((-log\, u_1)^{\theta} + (-log\, u_2)^{\theta})^{1/\theta}) \tag{6.5}$$

for $[u_1, u_2] \in [0, 1]^2$ and where $\theta \in [1, \infty)$ is a parameter indicating the strength on the dependence (larger θ indicates stronger dependence). θ in usually estimated by the method of maximum likelihood.

Copulas play an important role in estimating the tail dependence in bivariate data sets using the tail dependence coefficient. The upper tail dependence coefficient [13] is a limiting conditional probability that one time series exceeds a certain cutoff given that another time series exceeds this cutoff and is defined as:

$$\lambda_{upper} = \lim_{u \to 1^-} P(X_1 > F_1^{(-1)}(u) | X_2 > F_2^{(-1)}(u)) \tag{6.6}$$

λ_{upper} can also be expressed in terms of a bivariate copula function as:

$$\lambda_{upper} = \lim_{u \to \to 1^-} \frac{1 - 2u + C(u, u)}{1 - u}. \tag{6.7}$$

if $\lambda_{upper} \in (0, 1]$, then the joint bivarite distribution of X_1 and X_2 has upper tail dependence. For the Gumbel copula, the limit in equation 6.7 is equal to $(2 - 2^{1/\theta})$.

6.2 Analysis of spot electricity prices of Australia using copula method

The methodology described in Section 6.1 is now applied to the spot electricity prices of Australia. Given the nature of the method, we focus on the estimation of overall dependence between time series assuming it does not change over time. First, to get a general idea about the dependence structure between any two series Kendall's τ coefficient was estimated. The results are summarized in Table 6.1. We can see that the two series that have the strongest dependence among all the pairs are Victoria and South Australia ($\tau = 0.8260$). The two states are connected by two interconnectors, and South Australia is strongly dependent on the import of electricity from Victoria as it has no other sources to import into the state. Hence, the strong dependence between the two states is reasonable given the market conditions. Victoria and New South Wales are another pair of states that has a strong dependence ($\tau = 0.8019$). New South Wales also depends on electricity imported from Victoria, but it has a competitive option to import electricity from Queensland if prices are lower there. Queensland itself is connected only with New South Wales. This explains the strong dependence with New South Wales for Queensland ($\tau = 0.7907$) compared to three other states. South Australia is strongly related with New South Wales ($\tau = 0.7428$). As discussed in previous chapters, the two states are both dependent on the import from other states. Hence, similar market trends make them strongly correlated without having any actual connection between them.

Tasmania is the state that has the lowest dependence with the other

	NSW	QLD	SA	TAS	VIC
NSW	1.0000	0.7907	0.7428	0.5086	0.8019
QLD	0.7907	1.0000	0.6571	0.4502	0.7047
SA	0.7428	0.6571	1.0000	0.5272	0.8260
TAS	0.5086	0.4502	0.5272	1.0000	0.5714
VIC	0.8019	0.7047	0.8260	0.5714	1.0000

Table 6.1: *Kendall's τ for spot electricity prices of Australia for each of the two series.*

states. It is not located on Australian mainland and is able to satisfy most of its electricity demand with water power plants that produce cheaper electricity within the state. This explains the low dependence with other states. Queensland and Tasmania have the lowest dependence compared to other states ($\tau = 0.4502$) as they are the furthest away from each other in distance.

Next, the parameters of the Gumbel copula were estimated by maximum likelihood for each of the two series in the data set. Then, these parameters were used to estimate the upper tail dependence coefficients. The results are summarized in Tables 6.2 and 6.3.

It is evident that the dependence structure estimated by the Gumbel

	NSW	QLD	SA	TAS	VIC
NSW	-	3.73083 *(0.00000)*	3.109764 *(0.00000)*	1.809224 *(0.00000)*	4.21157 *(0.00000)*
QLD	3.73083 *(0.00000)*	-	2.42848 *(0.00000)*	1.611878 *(0.00000)*	2.752440 *(0.00000)*
SA	3.109764 *(0.00000)*	2.42848 *(0.00000)*	-	1.860927 *(0.00000)*	4.42610 *(0.00000)*
TAS	.809224 *(0.00000)*	1.611878 *(0.00000)*	1.860927 *(0.00000)*	-	2.072000 *(0.00000)*
VIC	4.21157 *(0.00000)*	2.752440 *(0.00000)*	4.42610 *(0.00000)*	2.072000 *(0.00000)*	-

Table 6.2: *Dependence parameter for bivariate Gumbel copula. p-values are given in parenthesis)*

	NSW	QLD	SA	TAS	VIC
NSW	1.00	0.7958318	0.7503121	0.5331482	0.8211002
QLD	0.7958318	1.00	0.6696736	0.4627047	0.7136226
SA	0.7503121	0.6696736	1.00	0.5486791	0.839467
TAS	0.5331482	0.4627047	0.5486791	1.00	0.6027158
VIC	0.8211002	0.7136226	0.839467	0.6027158	1.00

Table 6.3: *Tail dependence coefficients for bivariate Gumbel copula*

copulas is exactly the same that was obtained by Kendall's τ correlation coefficient even though, extreme value dependence should not necessarily mimic the dependence in the centers of the distributions. Again, Victoria has the strongest dependence with South Australia and New South Wales compared to any of the other two states. These have the highest estimated bivariate Gumbel copula coefficient and the largest tail dependence coefficient. Tasmania is the state that has the lowest dependence with other states.

Overall, for the analysis of spot electricity prices of Australia, estimating bivariate copulas does not give any additional information on the strength of the dependence between series compared to Kendall's τ correlation coefficient.

Also, the upper tail dependence coefficients show the asymptotic association of two data sets and are not connected to any particular cutoffs or quantiles. In addition, both methods estimate a general dependence between the data sets and do not take into account possible lag dependence which is present in all of the data sets. All three methods used in previous chapters allows to account both for lag dependence and include a particular cutoff.

For reasons described above, estimating the extreme value dependence in time series data sets with copulas and tail dependence coefficients is considered unsatisfactory for the analysis of spot electricity prices of Australia. However, the two methods used in this section might be more useful for different data sets.

Chapter 7

Discussion

In this work, extreme value dependence for the spot electricity prices of Australia was estimated using four different methods: univariate and cross extremograms, quantile regression, cross-quantilogram and bivariate copula.

It was found that the extremogram and quantilogram are useful tools for exploratory data analysis. Both of the methods can be used to test for the presence of any dependence in the upper or lower quantiles between two series. In addition, the univariate extremogram allows us to study the dependence between extreme events within one the series. Nevertheless, the two methods provide marginally different information about the two series. The cross-quantilogram is a useful tool to measure extreme value dependence between conditional quantiles of two series at different lags. On the other hand, the cross extremogram gives information about the probability of extreme events occurring at a certain point in the future given an extreme event at time t. Also, the two methods treat negative associations differently. With the cross-quantilogram two events are considered negatively correlated if one event happens above a conditional quantile and another event happens below it. The cross extremogram focusses more on the tails of the distributions, and it does not take into account events happening in centres. For the extremogram, a negative association can be estimated as a probability of extreme events happening in different tails of the distributions. The differences in the two methods make the cross-quantilogram a useful tool to help select predictors in quantile regression. The cross extremogram is a valuable statistical method to explore extreme event dependence between two time series that gives additional information about the dependence, which cannot be found by any other method. Unlike the cross-quantilogram, it also can be extended to a multivariate case with 3 time series. However, this extension complicates the computation procedure. The cross extremogram can also be used as a way of selecting predictors for quantile regression but the necessity to estimate negative associations separately from positive ones makes it less effective compared to the cross-quantilogram. Finally, both the cross-quantilogram and the cross extremogram have a common limitation: the two methods are proven to work well for strictly stationary time series, which are rare in many applications.

Quantile regression is a method used to predict conditional quantiles of a data set using a set of predictors. The method does not require any strict assumptions about the data, and recent extensions of the method to the time series setting makes it a useful tool for estimating how extreme events in one time series data set depend on other time series data sets. Similar to the cross-quantilogram, quantile regression focusses on events occurring above a chosen conditional quantile and events below it without making any emphasis on actual numerical values of each observation. Nevertheless, the method gives additional information about the nature of extreme event occurrence compared to the cross extremogram. Another benefit of the quantile regression is that it allows us to estimate the dependence between one time series and multiple predictors, including lags of the same variable, in the same model giving more information as to which variables influence the occurrence of extreme events more than others. A disadvantage of the quantile regression is that it has a limited number of model selection procedures. Existing approaches work well for the median but show weak performance for higher quantiles. Also, a solution for a quantile regression problem may not be unique.

The copula method for estimating extreme value dependence, which was also used in the analysis, is a convenient method to estimate joint distributions whose marginal distributions not normal. Some copulas are also helpful in analysis of joint asymptotic tail behavior, but for our purposes of the copula method gives a limited amount of information. In general, the dependence estimated using copulas can only be compared to the other results as no strict rule for classifying the dependence as strong or weak exists. For this reason, estimating a five-dimensional copula for all of the series together does not give any additional information as there is no base of comparison for the results. Also, compared to thee other methods, copulas only estimate general dependence between data sets and treat the dependence as stable over time as lags cannot be included. The method does not take into account time series features (such as autocorrelation, seasonality, etc.) of data sets it is applied to.

In conclusion, all of the extreme value dependence methods have a common limitation: they require a significantly large time series data set to give reliable results. Nevertheless, when this condition is satisfied all the methods can be considered useful and, if used together, they provide a rich source of information about dependence between extreme events compared to a single method. Finally, all methods have dedicated R packages. This makes implementation of all methods very easy.

Chapter 8

Conclusion

The aim of this thesis was to apply various methods for estimating extreme value dependence in time series data sets for spot electricity prices of Australia and to compare their performance. The emphasis was made on the upper tail of each time series as understanding the dependence in high electricity price is generally of more interest compared to low electricity prices.

We first concentrated on the extremogram and the cross extremogram analysis. It was applied to study probabilities of extreme event occurrence at time $t + h$ given an extreme event at time t in the same or in another time series, respectively. The method focusses on the tails of distributions and gave additional information about the nature of extreme event dependence, which can be concluded by other methods used further in the thesis. We created an R package for application of the extremogram procedure that includes functions for estimating univariate and cross extremograms, permutation confidence intervals and stationary bootstrap confidence intervals for extremogram values.

The cross-quantilogram was used to estimate extreme value dependence between high upper conditional quantiles, and it was found to be a useful tool for selecting predictors for quantile regression. Quantile regression itself was found to be a convenient method for predicting future extreme conditional quantiles based on a single or several predictors. Finally, the copula method was used as a non-parametric method for estimating dependence in upper tails of distributions.

Appendix A

Appendix

A.0.1 Appendix 1

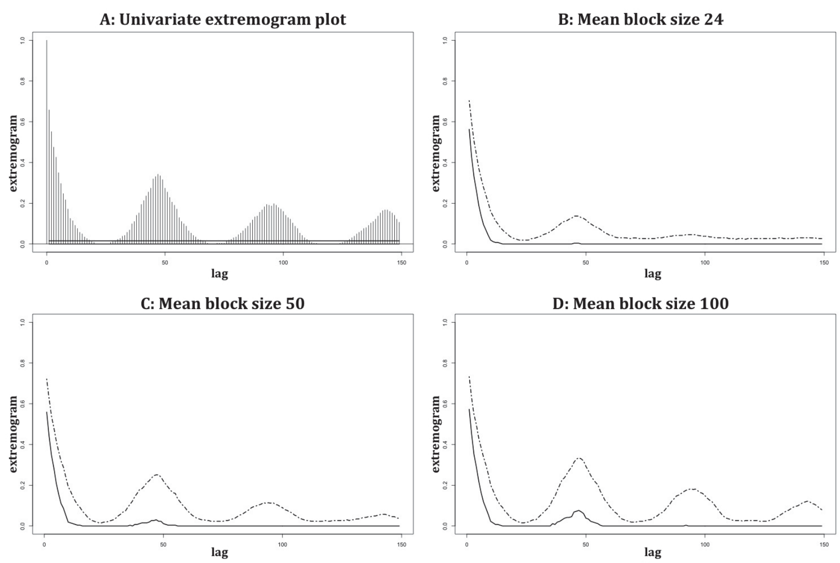

Figure A.0.1: *A: Univariate extremogram plot for QLD (2009 – 2011); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

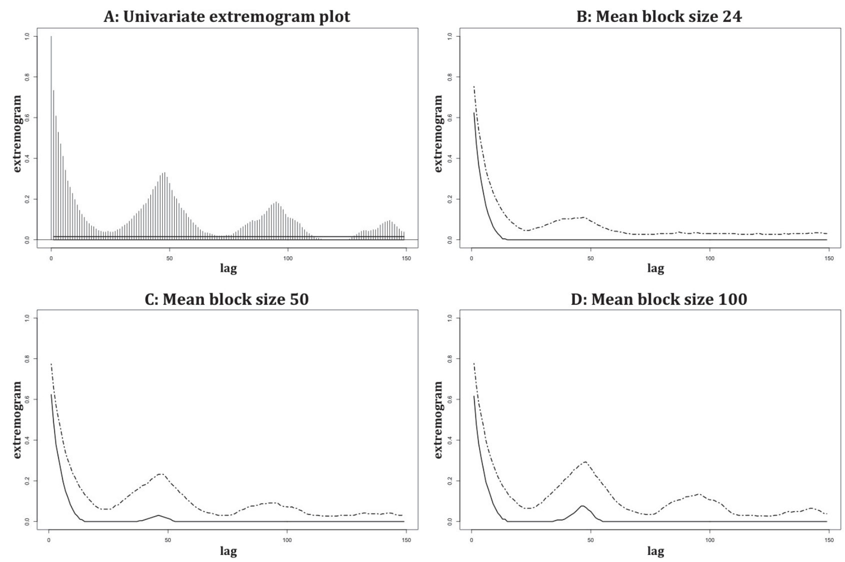

Figure A.0.2: *A: Univariate extremogram plot for SA (2009 – 2011); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

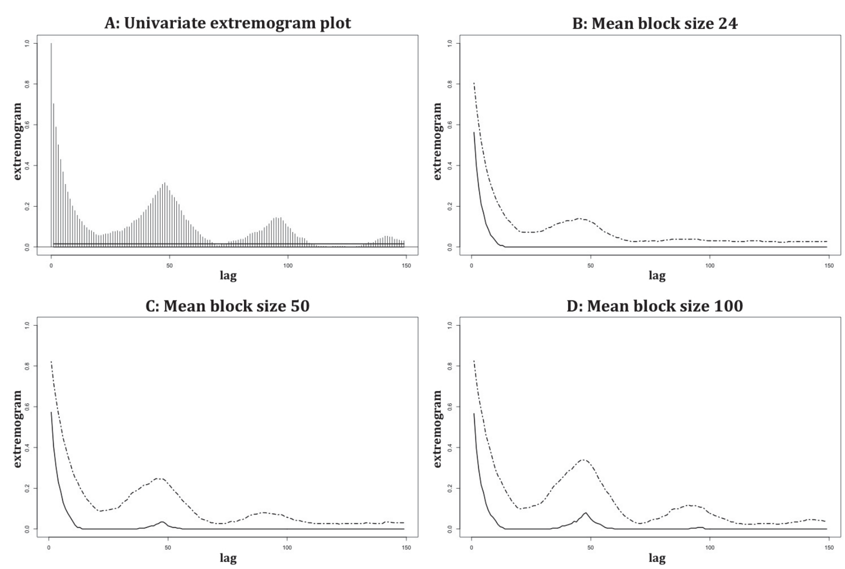

Figure A.0.3: *A: Univariate extremogram plot for VIC (2009 – 2011); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

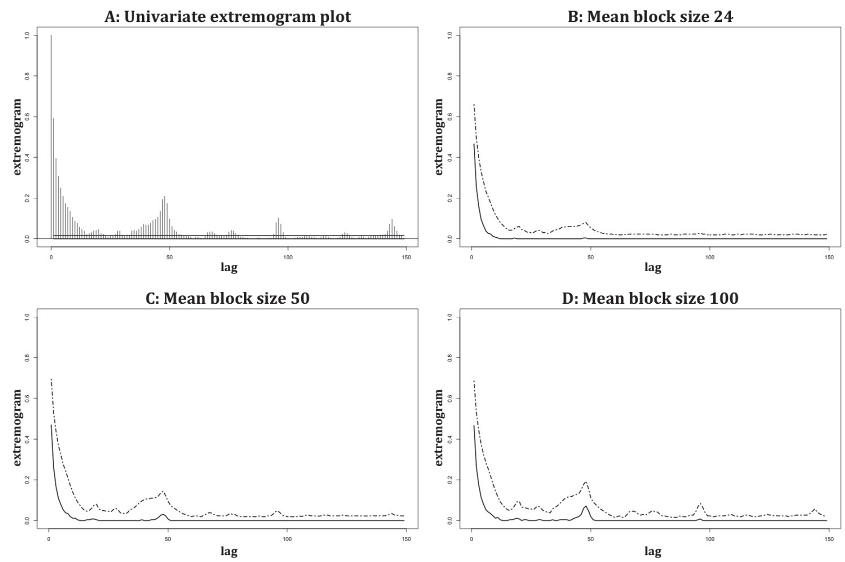

Figure A.0.4: *A: Univariate extremogram plot for NSW (2012 – 2014); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

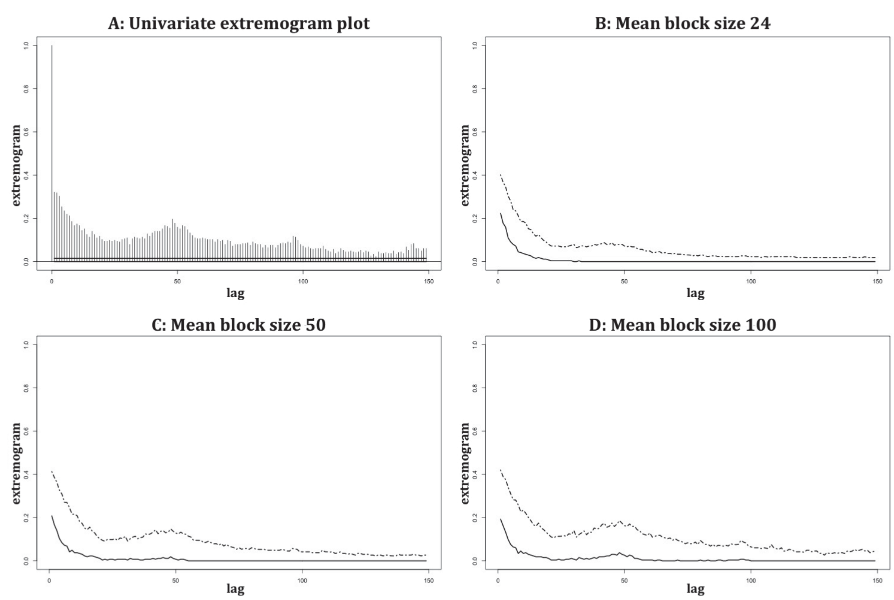

Figure A.0.5: *A: Univariate extremogram plot for QLD (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

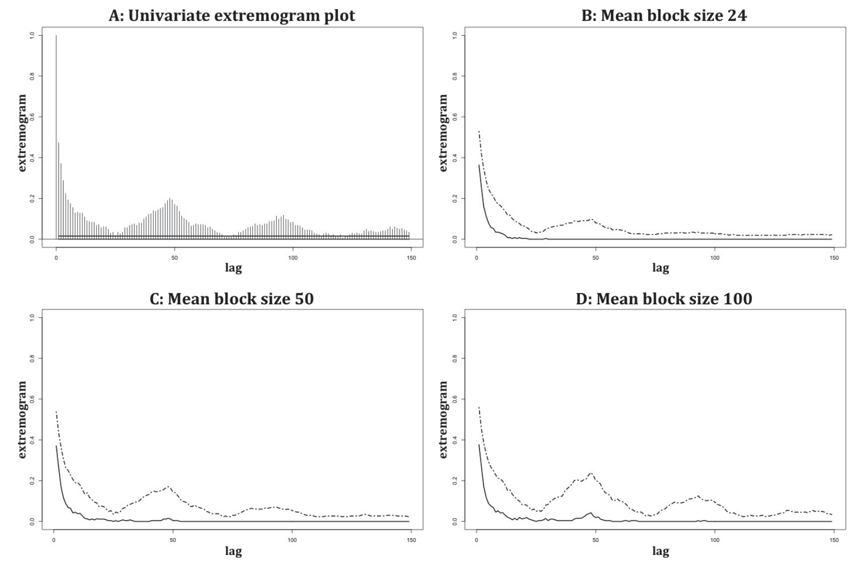

Figure A.0.6: *A: Univariate extremogram plot for SA (2012 – 2014); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

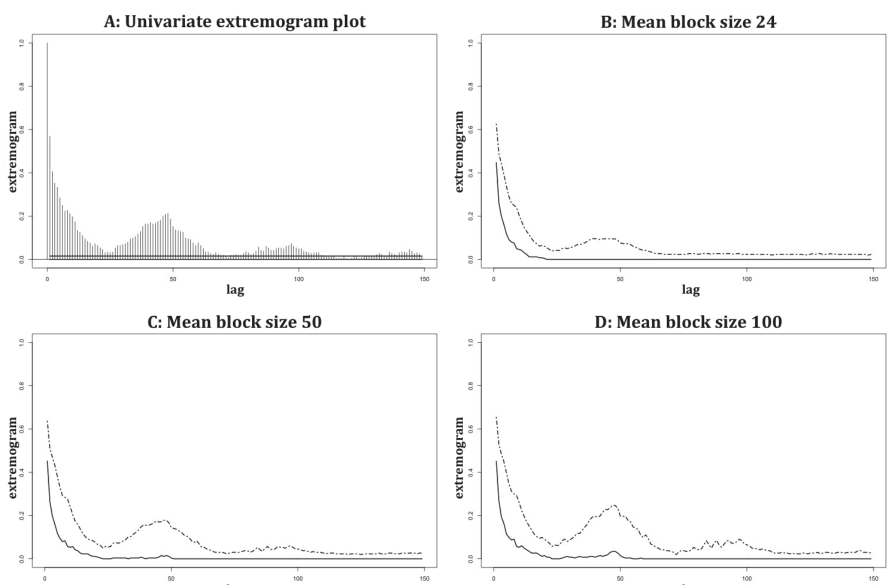

Figure A.0.7: *A: Univariate extremogram plot for VIC (2012 – 2014); Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

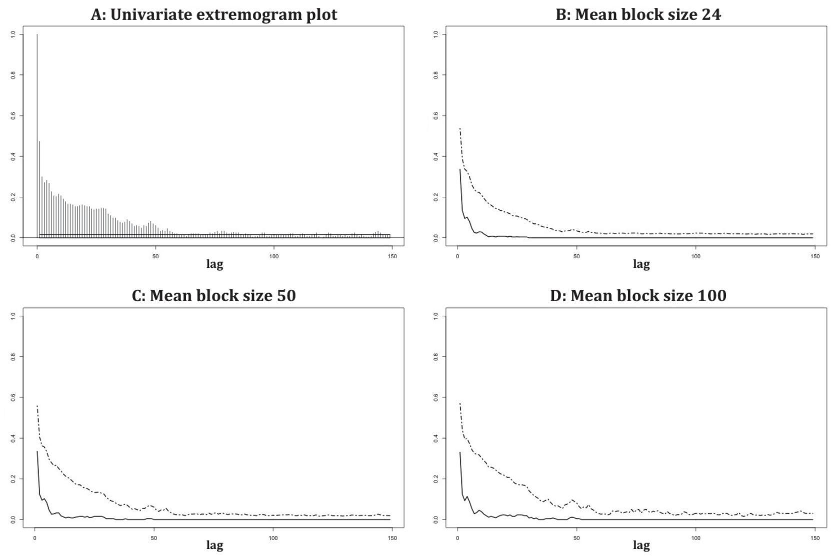

Figure A.0.8: *A: Univariate extremogram plot for TAS (2012 – 2014);Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

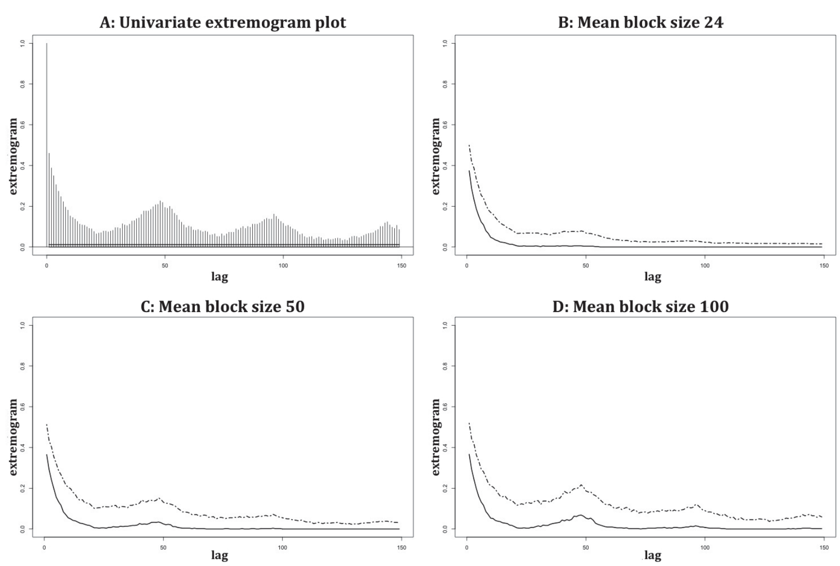

Figure A.0.9: *A: Univariate extremogram plot for QLD (2009 – 2014), Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

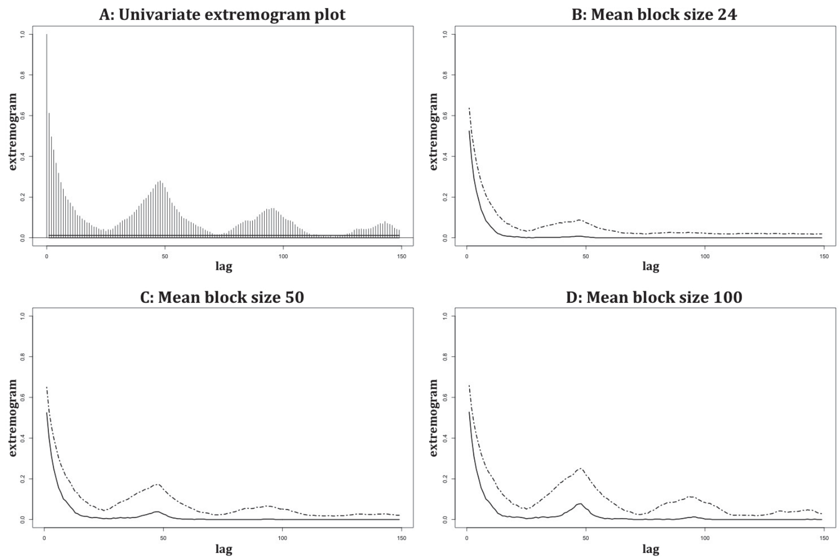

Figure A.0.10: *A: Univariate extremogram plot for SA (2009 – 2014), Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

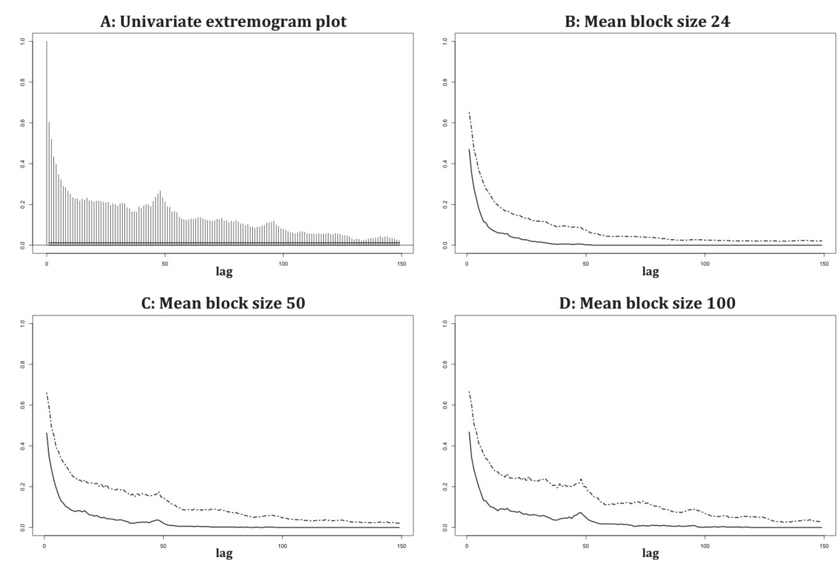

Figure A.0.11: *A: Univariate extremogram plot for TAS (2009 – 2014), Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

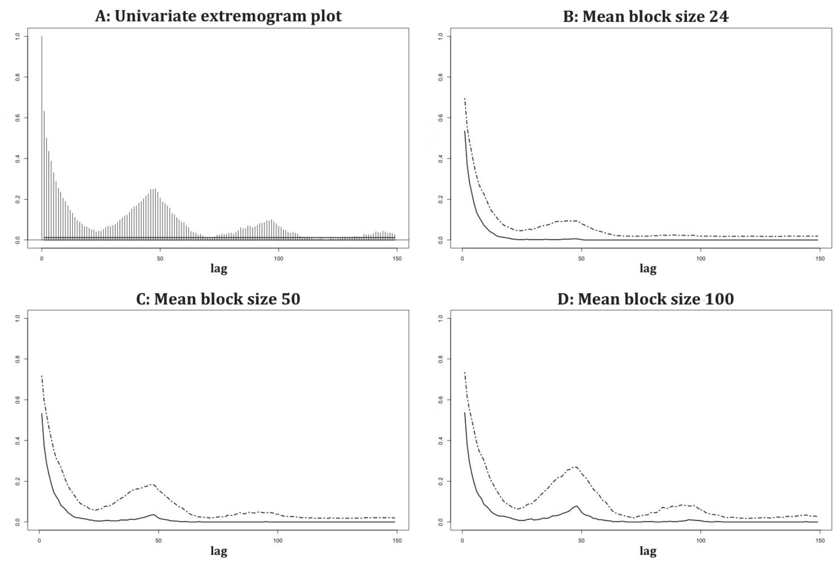

Figure A.0.12: *A: Univariate extremogram plot for VIC (2009 – 2014), Stationary bootstrap confidence intervals with mean block size B: 24; C: 50; D: 100.*

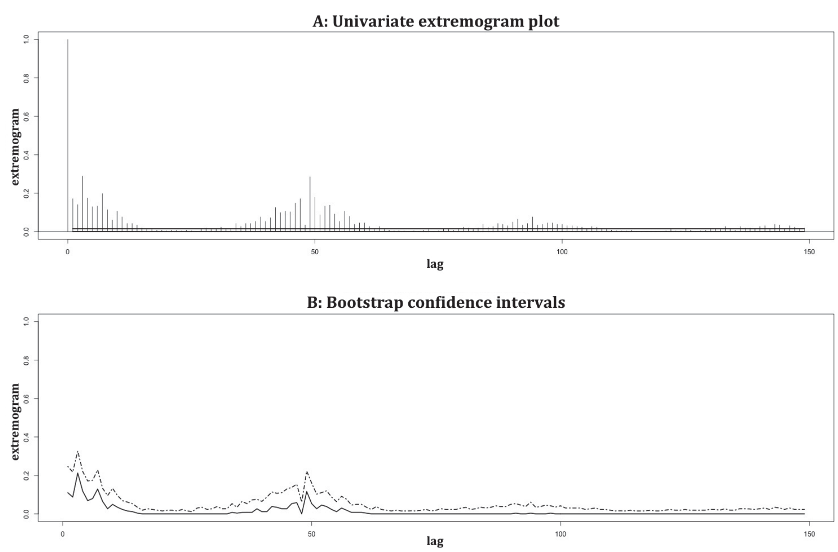

Figure A.0.13: *A: Univariate extremogram plot after fitting an ARIMA(3,0,5)(3,1,2)$_{48}$ to QLD (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

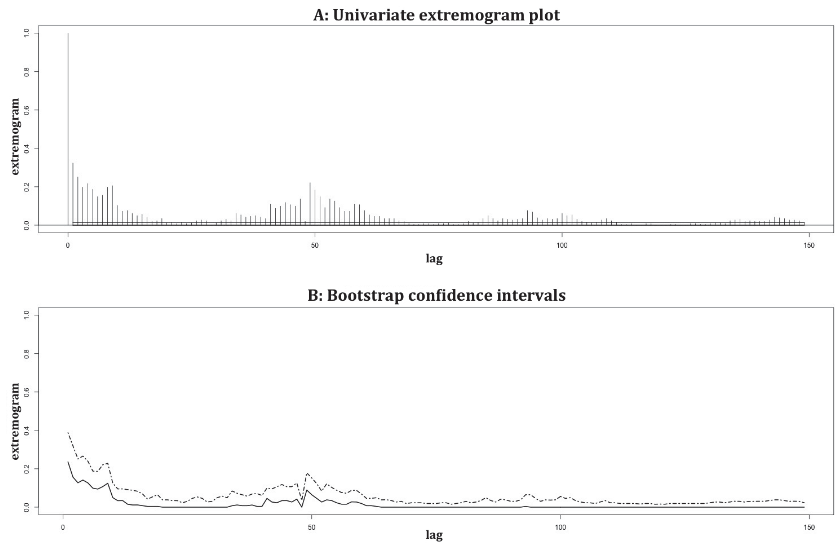

Figure A.0.14: *A: Univariate extremogram plot after fitting an ARIMA(5,0,8)(3,1,3)$_{48}$ to SA (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

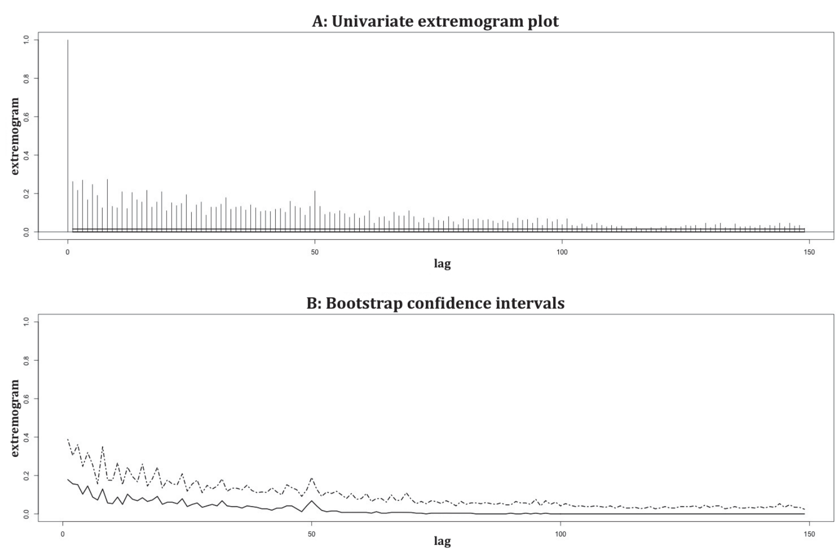

Figure A.0.15: *A: Univariate extremogram plot after fitting an ARIMA(4,0,4)(1,1,1)$_{48}$ to TAS (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

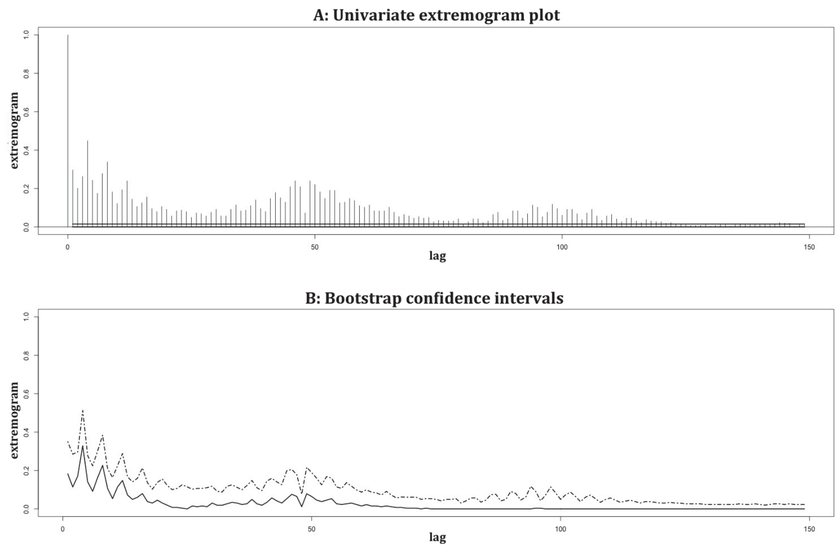

Figure A.0.16: *A: Univariate extremogram plot after fitting an ARIMA(3,0,4)(2,1,1)$_{48}$ to VIC (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

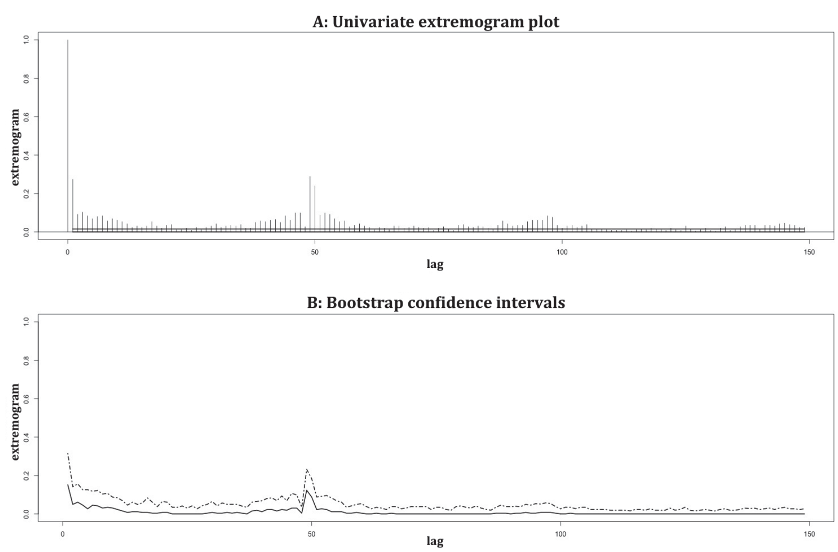

Figure A.0.17: *A: Univariate extremogram plot after fitting an ARIMA(3,0,5)(5,1,1)$_{48}$ to QLD (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

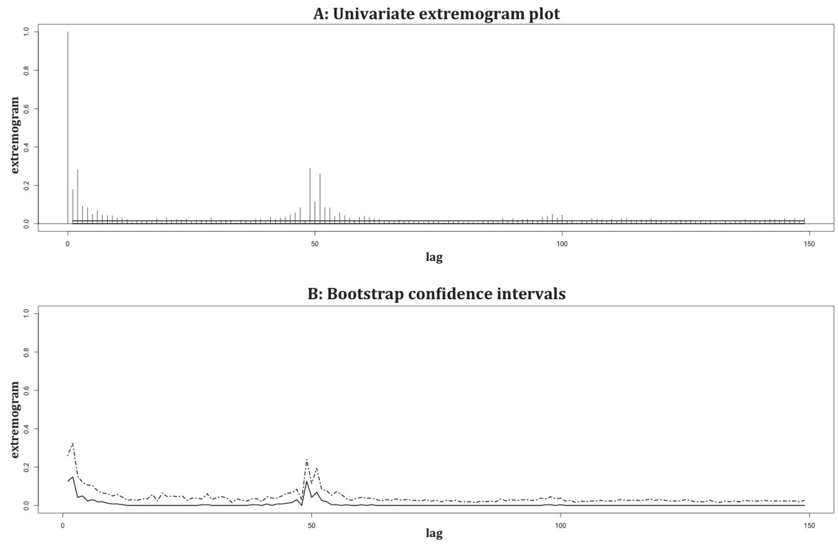

Figure A.0.18: *A: Univariate extremogram plot after fitting an ARIMA(5,0,8)(1,1,2)$_{48}$ to SA (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

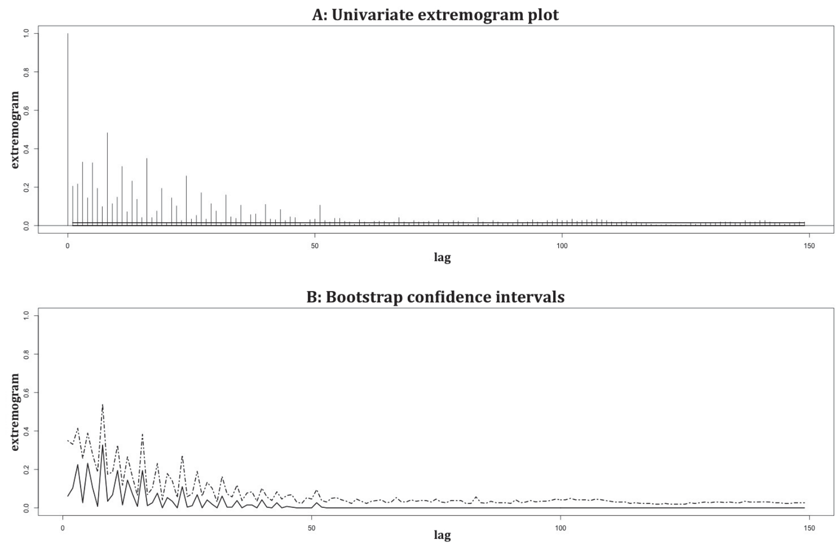

Figure A.0.19: *A: Univariate extremogram plot after fitting an ARIMA(3,0,4)(1,1,1)$_{48}$ to TAS (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

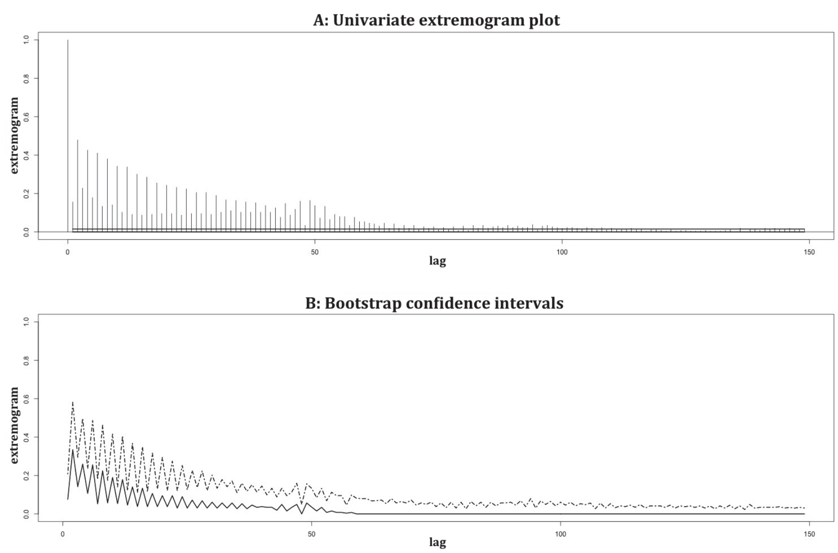

Figure A.0.20: *A: Univariate extremogram plot after fitting an ARIMA(4,0,4)(1,1,1)$_{48}$ to VIC (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

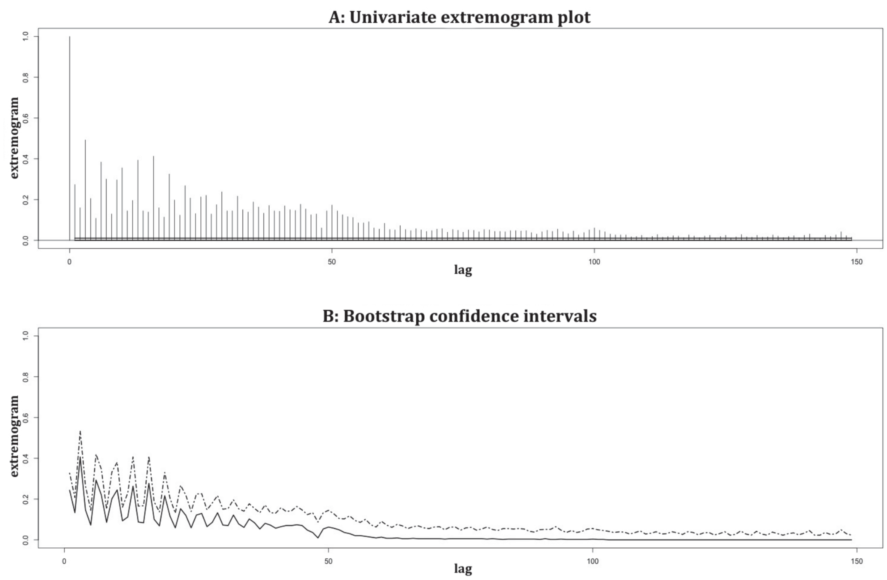

Figure A.0.21: *A: Univariate extremogram plot after fitting an ARIMA(5,0,6)(1,1,1)*$_{48}$ *to NSW (2011 - 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

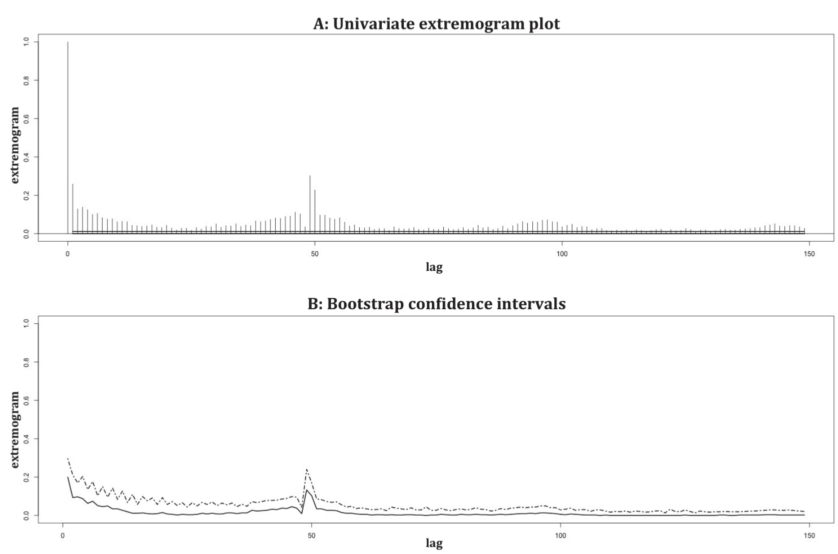

Figure A.0.22: *A: Univariate extremogram plot after ARIMA(4,0,1)(1,1,1)*$_{48}$ *to SA (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

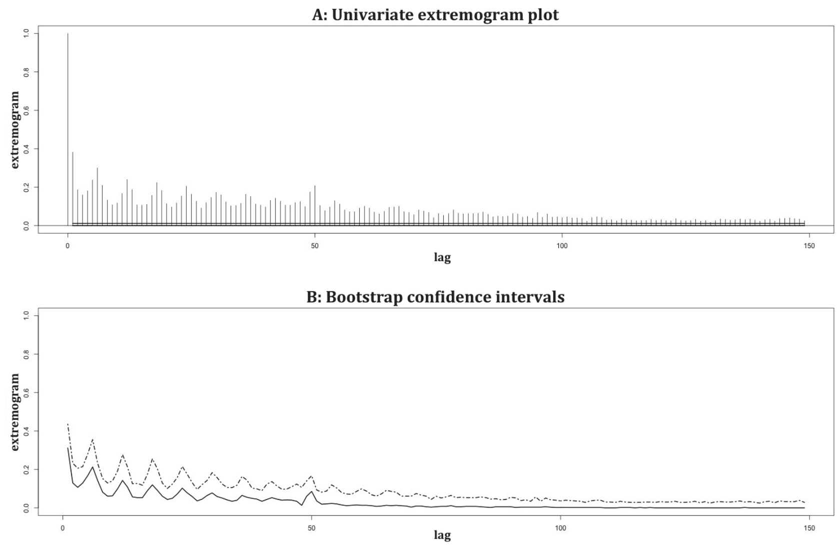

Figure A.0.23: *A: Univariate extremogram plot after fitting an ARIMA(3,0,2)(1,1,1)$_{48}$ to TAS (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

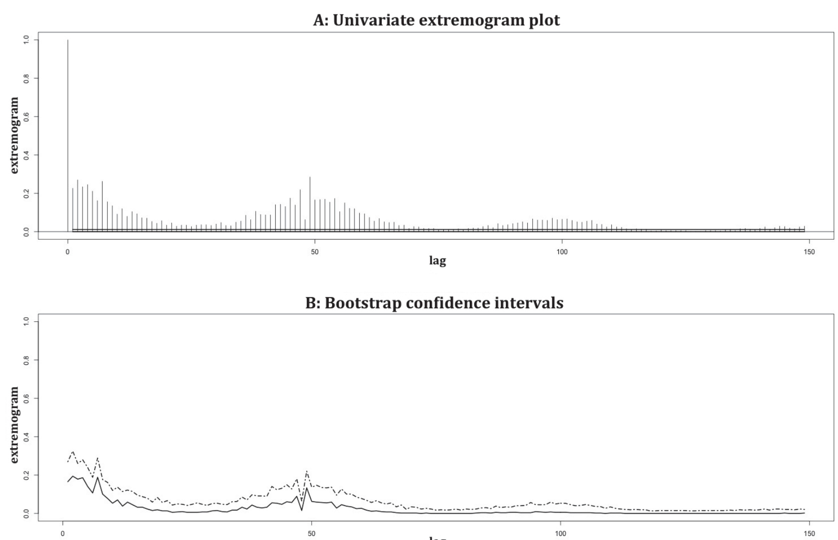

Figure A.0.24: *A: Univariate extremogram plot after fitting an ARIMA(3,0,2)(1,1,1)$_{48}$ to VIC (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

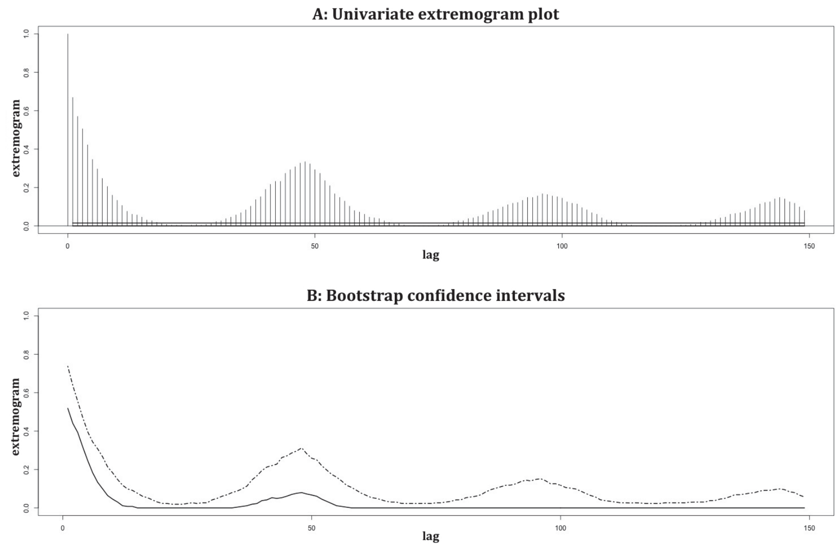

Figure A.0.25: *A: Univariate extremogram plot after fitting a GARCH(1,1) to NSW (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

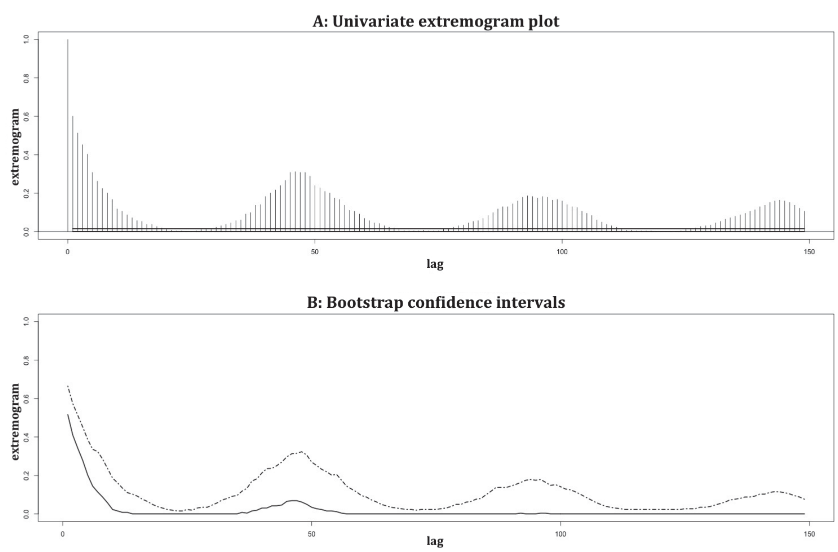

Figure A.0.26: *A: Univariate extremogram plot after fitting a GARCH(1,1) to QLD (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

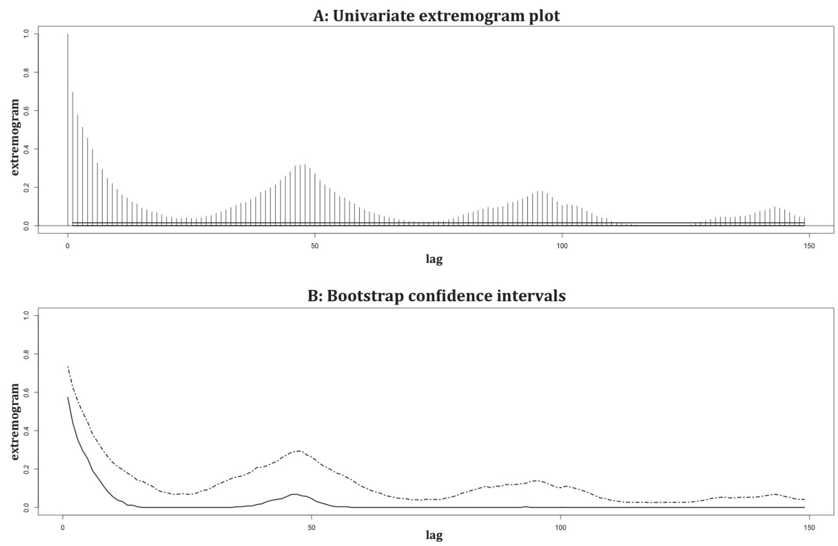

Figure A.0.27: *A: Univariate extremogram plot after fitting a GARCH(1,1) to SA (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

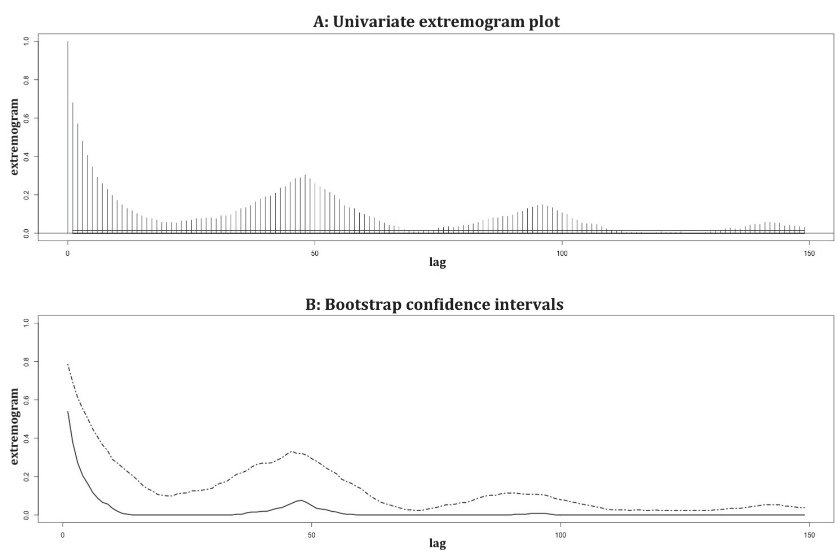

Figure A.0.28: *A: Univariate extremogram plot after fitting a GARCH(1,1) to VIC (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

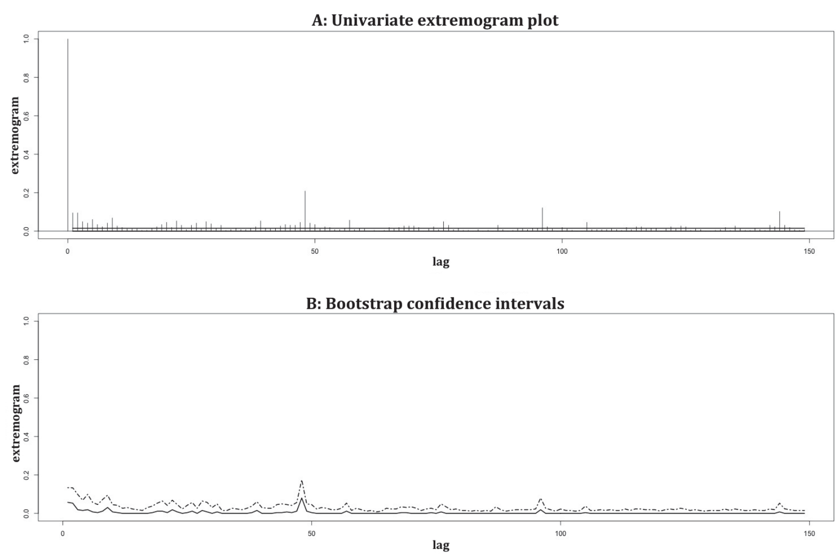

Figure A.0.29: *A: Univariate extremogram plot after fitting a GARCH(1,1) to NSW (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

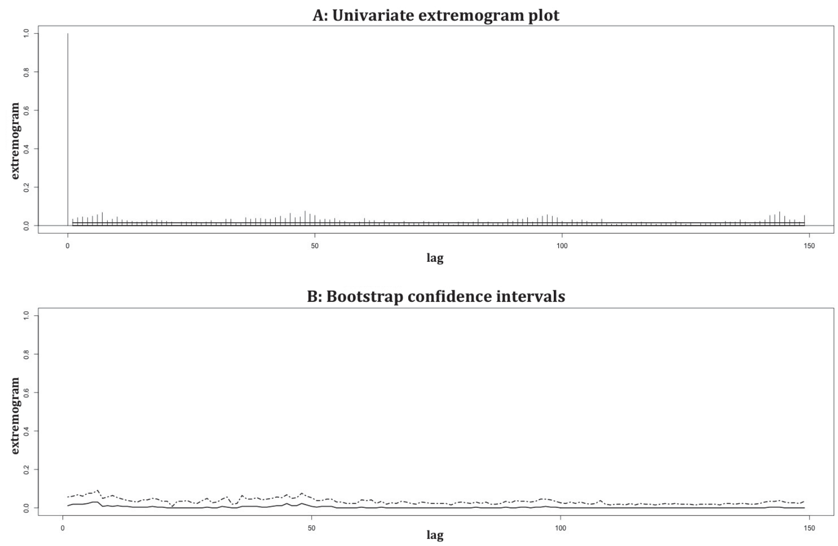

Figure A.0.30: *A: Univariate extremogram plot after fitting a GARCH(1,1) to QLD (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

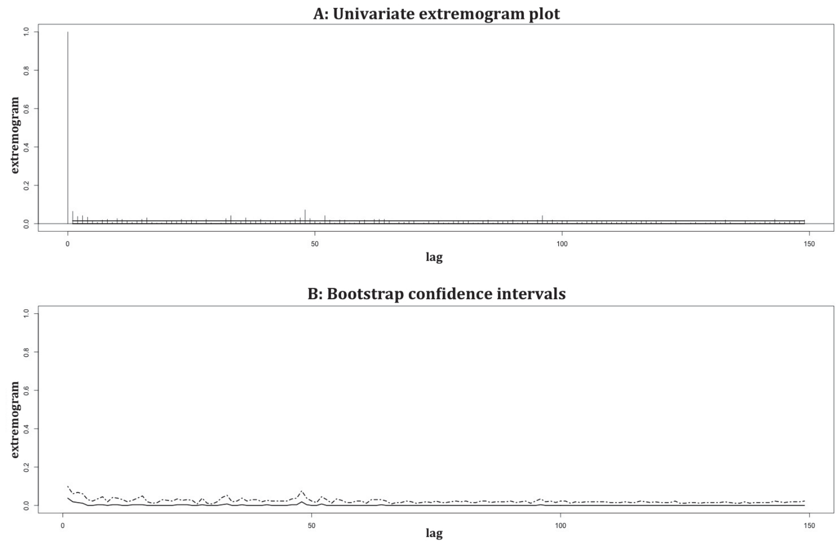

Figure A.0.31: *A: Univariate extremogram plot after fitting a GARCH(1,1) to SA (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

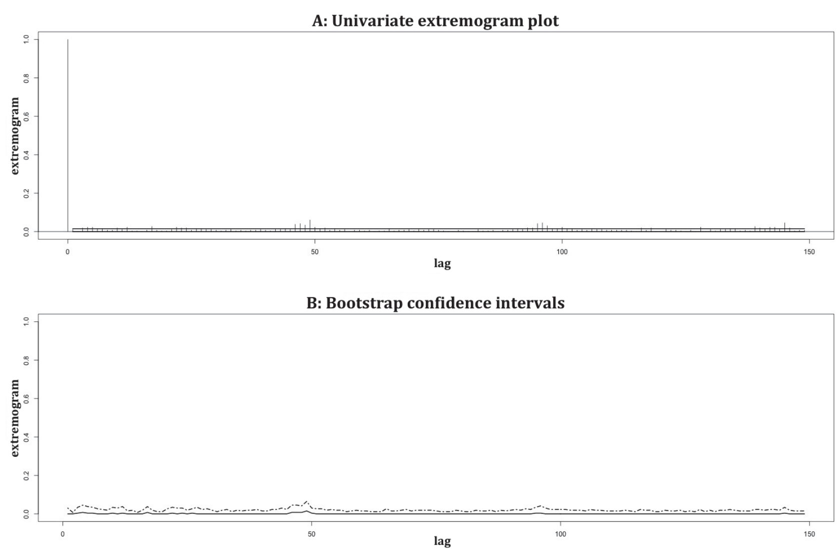

Figure A.0.32: *A: Univariate extremogram plot after fitting a GARCH(1,1) to VIC (2012 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

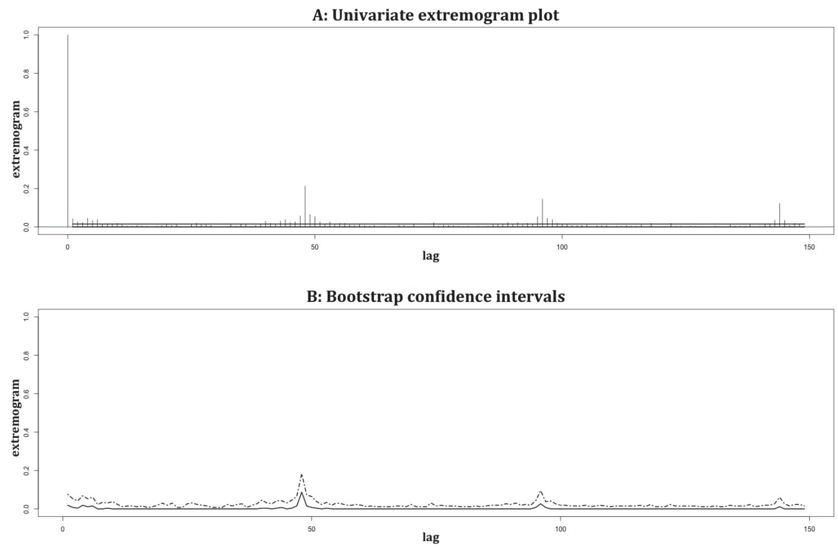

Figure A.0.33: *A: Univariate extremogram plot after fitting an ARIMA and GARCH(1,1) to NSW (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

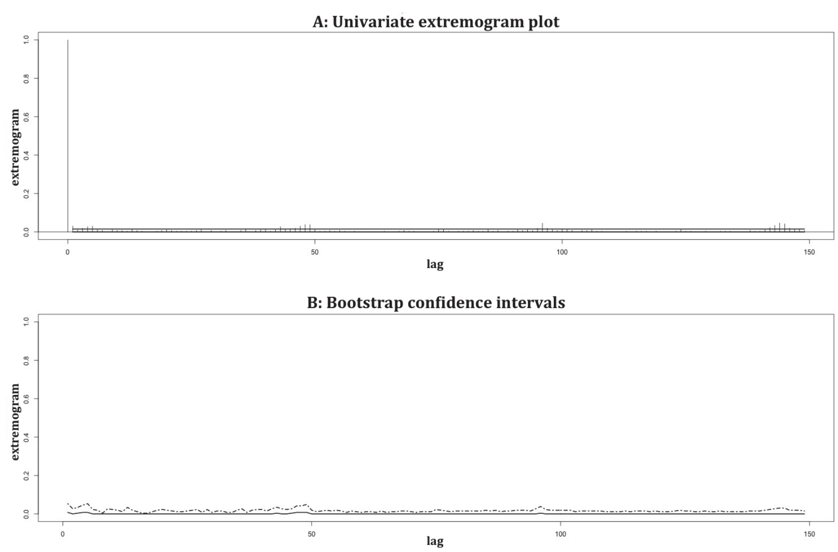

Figure A.0.34: *A: Univariate extremogram plot after fitting an ARIMA and GARCH(1,1) to QLD (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

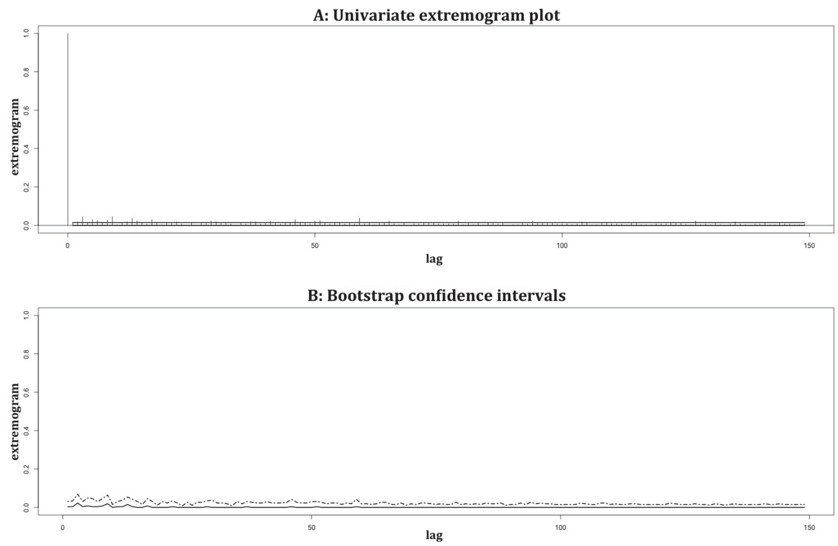

Figure A.0.35: *A: Univariate extremogram plot after fitting an ARIMA and GARCH(1,1) to SA (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

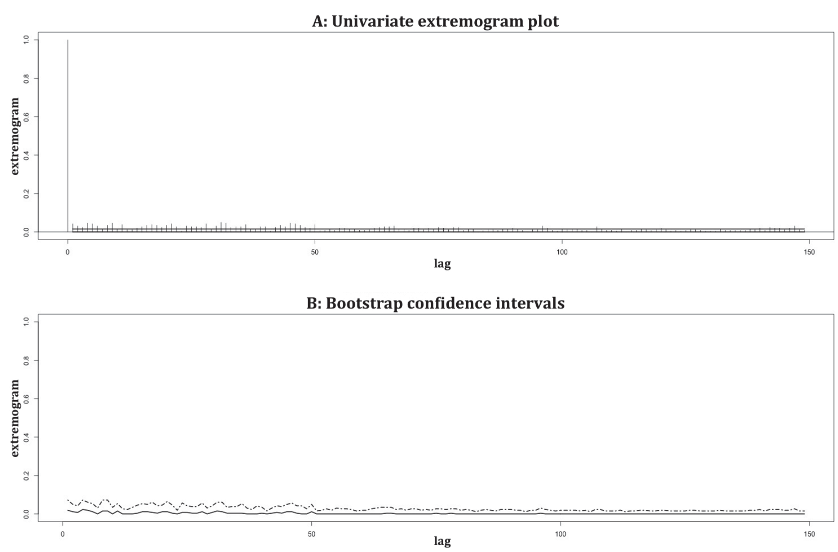

Figure A.0.36: *A: Univariate extremogram plot after fitting an ARIMA and GARCH(1,1) to TAS (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

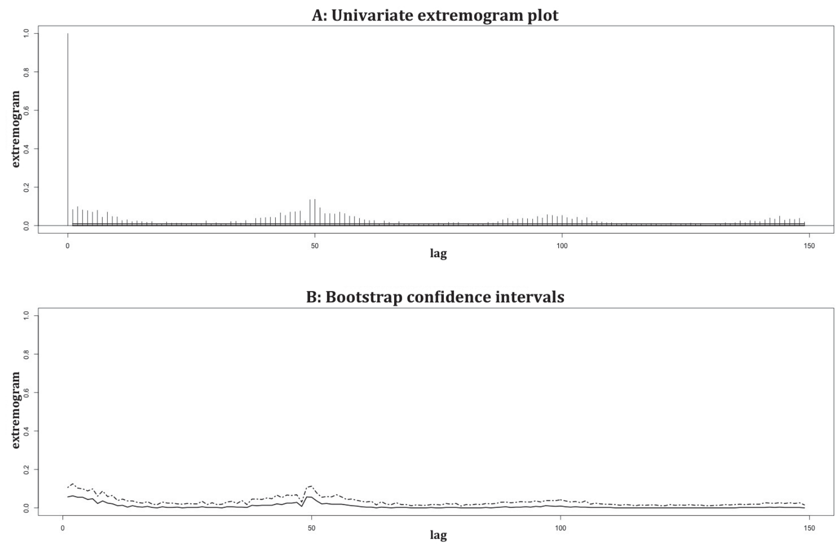

Figure A.0.37: *A: Univariate extremogram plot after fitting a GARCH(1,1) to residuals estimated for first differenced series for NSW (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

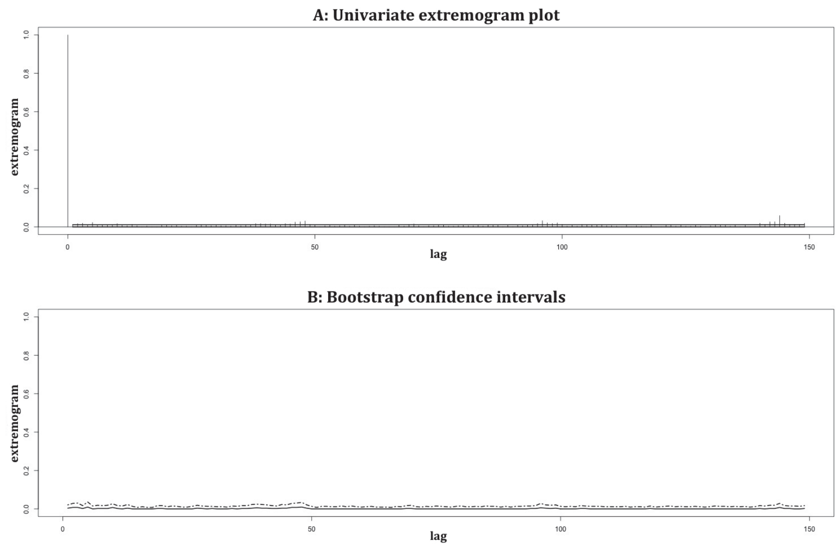

Figure A.0.38: *A: Univariate extremogram plot after fitting a GARCH(1,1) to residuals estimated for first differenced series for QLD (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

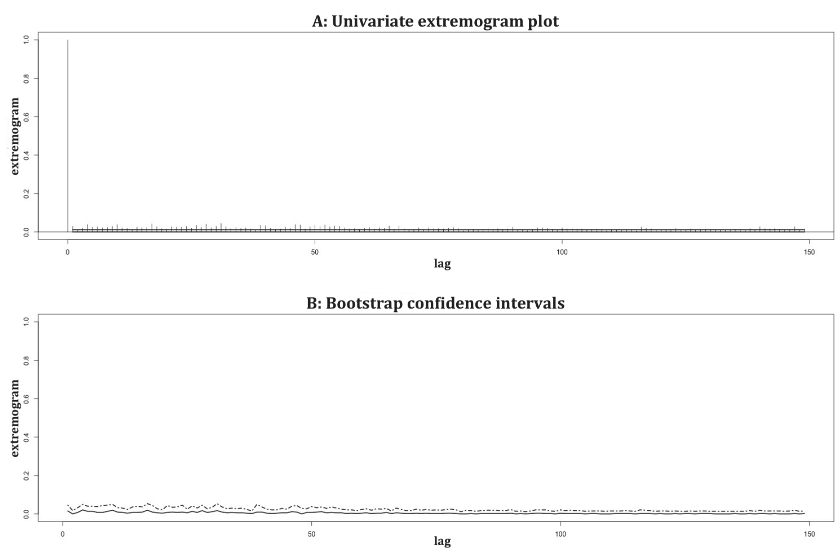

Figure A.0.39: *A: Univariate extremogram plot after fitting a GARCH(1,1) to residuals estimated for first differenced series for TAS (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

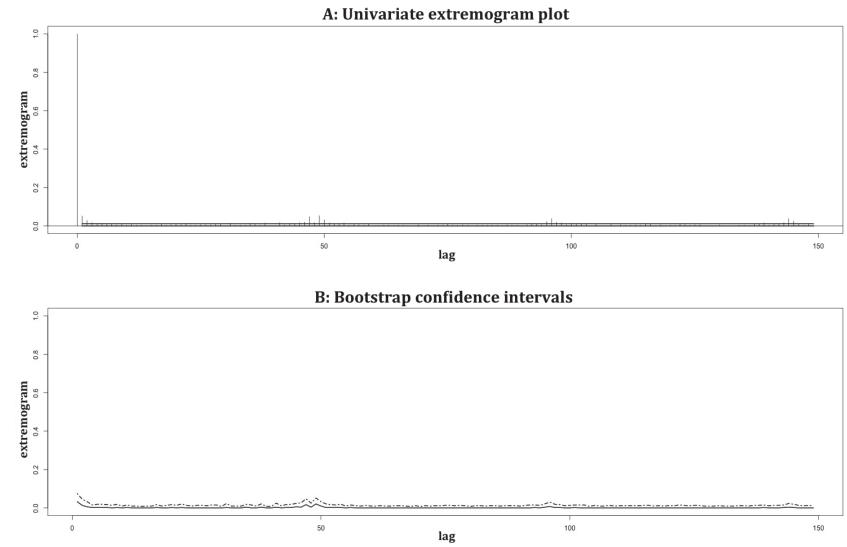

Figure A.0.40: *A: Univariate extremogram plot after fitting an GARCH(1,1) to residuals estimated for first differenced series for VIC (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

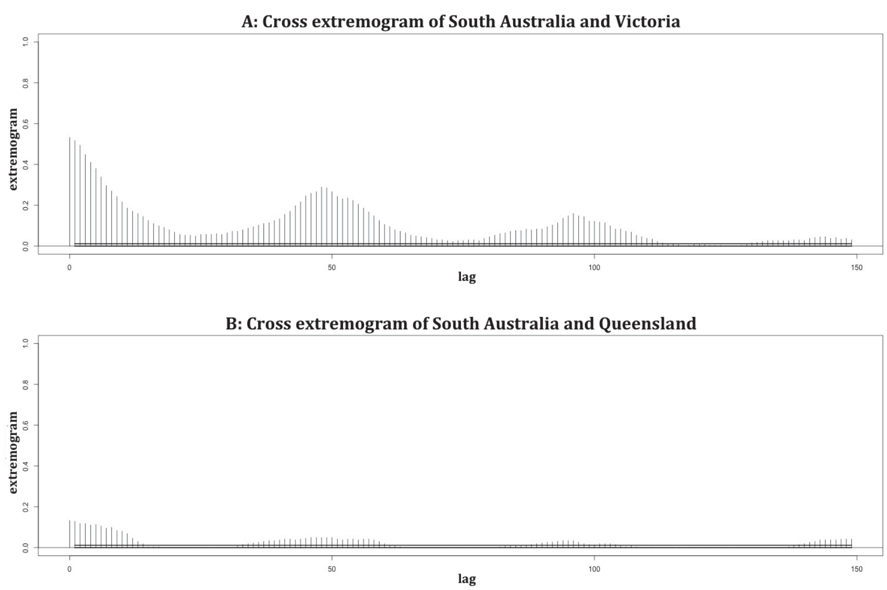

Figure A.0.41: *A: Cross extremogram for SA conditioning on VIC (2009 – 2011); B: Cross extremogram for SA conditioning on QLD (2009 – 2011);*

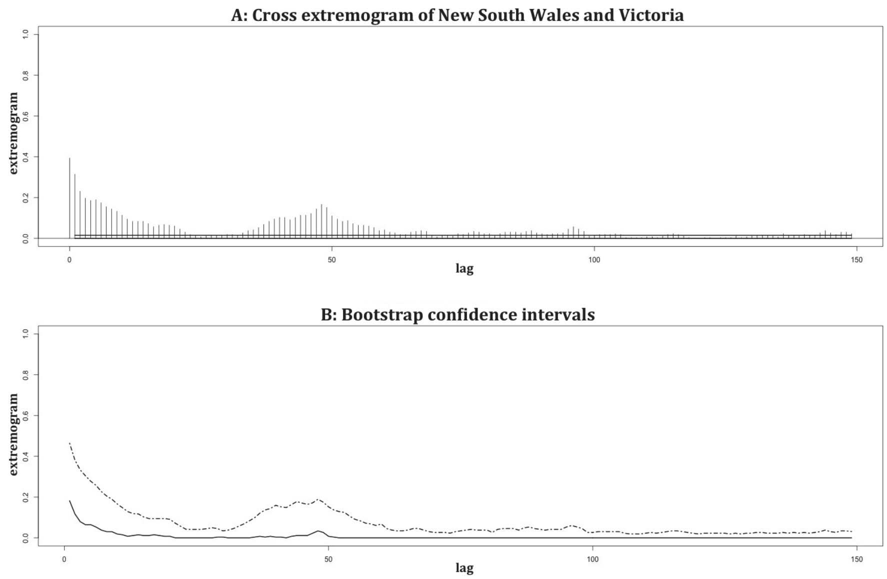

Figure A.0.42: *A: Cross extremogram for NSW conditioning on VIC (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

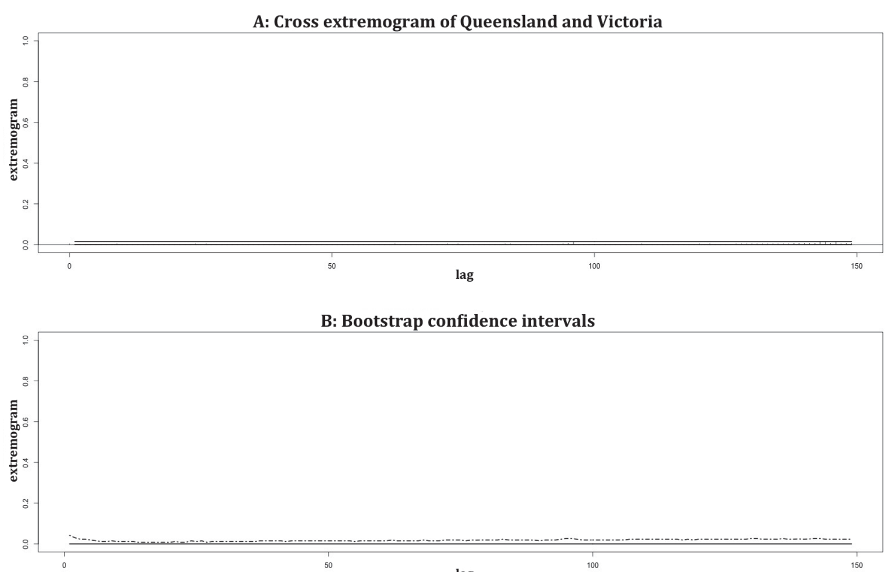

Figure A.0.43: *A: Cross extremogram for QLD conditioning on VIC (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

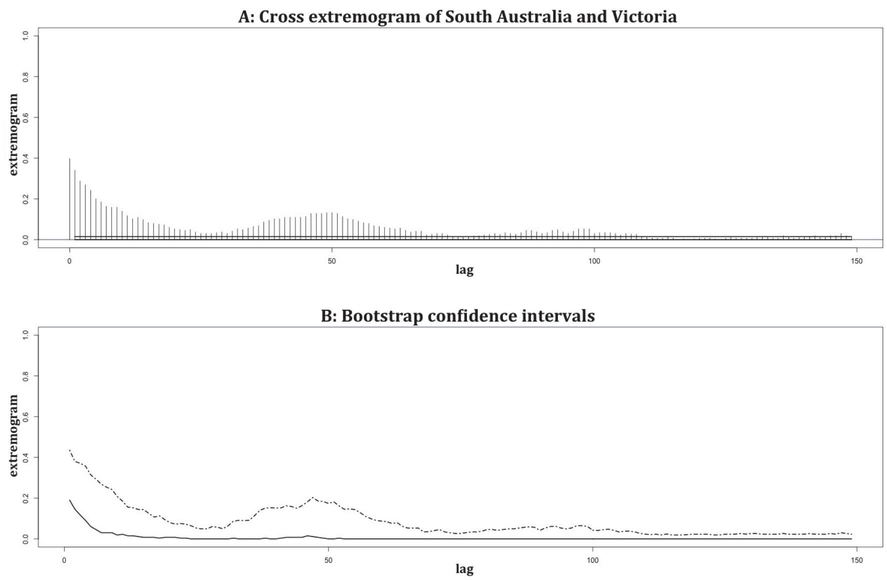

Figure A.0.44: *A: Cross extremogram for SA conditioning on VIC (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

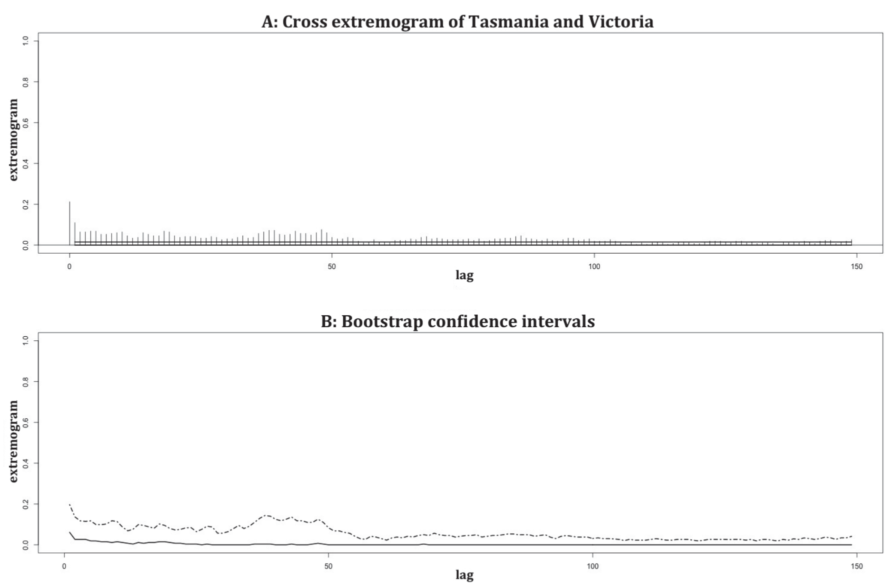

Figure A.0.45: *A: Cross extremogram for TAS conditioning on VIC (2009 – 2011); B: Stationary bootstrap confidence intervals with mean block size = 100;*

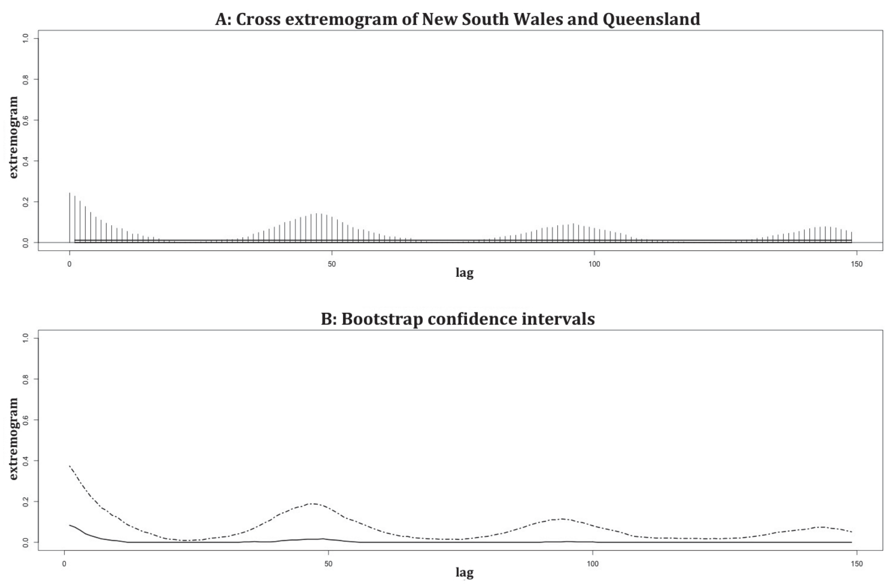

Figure A.0.46: *A: Cross extremogram for NSW conditioning on QLD (2009 – 2014); B: Stationary bootstrap confidence intervals with mean block size = 100;*

A.0.2 Appendix 2

Package ‘extremogram’

March 28, 2016

Type Package

Title Estimation of Extreme Value Dependence for Time Series Data

Version 1.0.1

Date 2015-09-24

Author Nadezda Frolova, Ivor Cribben

Maintainer Nadezda Frolova <nfrolova@ualberta.ca>

Description Estimation of the sample univariate, cross and return time extremograms. The package can also add empirical confidence bands to each of the extremogram plots via a permutation procedure under the assumption that the data are independent. Finally, the stationary bootstrap allows us to construct credible confidence bands for the extremograms.

License GPL-3

Imports boot(>= 1.3-11), MASS(>= 7.3-31), parallel(>= 3.1.1)

Depends R (>= 3.1.0)

NeedsCompilation no

Repository CRAN

Date/Publication 2016-03-28 23:17:50

R topics documented:

extremogram-package *extremogram*

Description

The package estimates the sample univariate, cross and return time extremograms. It can also add empirical confidence bands to each of the extremogram plots via a permutation procedure under the assumption that the data are independent. Finally, the stationary bootstrap allows us to construct credible confidence bands for the extremograms.

Functions:

1. `extremogram1`
2. `extremogram2`
3. `extremogramr`
4. `bootconf1`
5. `bootconf2`
6. `bootconfr`
7. `permfn1`
8. `permfn2`
9. `permfnr`

Author(s)

Nadezda Frolova <nfrolova@ualberta.ca>, Ivor Cribben <cribben@ualberta.ca>

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

bootconf1 *Confidence bands for the sample univariate extremogram*

Description

The function estimates confidence bands for the sample univariate extremogram using the stationary bootstrap.

Usage

```
bootconf1(x, R, l, maxlag, quant, type, par, start = 1, cutoff = 1,
  alpha = 0.05, ...)
```

Arguments

`x`	Univariate time series (a vector).
`R`	Number of bootstrap replications (an integer).
`l`	Mean block size for stationary bootstrap or mean of the geometric distribution used to generate resampling blocks (an integer that is not longer than the length of the time series).
`maxlag`	Number of lags to include in the extremogram (an integer).
`quant`	Quantile of the time series to indicate an extreme event (a number between 0 and 1).
`type`	Extremogram type (see function `extremogram1`).
`par`	If par = 1, the bootstrap replication procedure will be parallelized. If par = 0, no parallelization will be used.
`start`	The lag that the extremogram plots starts at (an integer not greater than `maxlag`, default is 1).
`cutoff`	The cutoff of the y-axis on the plot (a number between 0 and 1, default is 1).
`alpha`	Significance level for the confidence bands (a number between 0 and 1, default is 0.05).
`...`	further arguments: plot and axis names.

Value

Returns a plot of the confidence bands for the sample univariate extremogram.

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega  = 1
alpha  = 0.1
beta   = 0.6
n      = 1000
quant  = 0.95
type   = 1
maxlag = 70
df     = 3
R      = 10
l      = 30
par    = 0
G = extremogram:::garchsim(omega,alpha,beta,n,df)

extremogram1(G, quant, maxlag, type, 1, 1, 0)
bootconf1(G, R, l, maxlag, quant, type, par, 1, 1, 0.05)
```

`bootconf2` *Confidence bands for the sample cross extremogram*

Description

The function estimates confidence bands for the sample cross extremogram using the stationary bootstrap.

Usage

```
bootconf2(x, R, l, maxlag, quant1, quant2, type, par, start = 1, cutoff = 1,
  alpha = 0.05, ...)
```

Arguments

`x`	Bivariate time series (n by 2 matrix).
`R`	Number of bootstrap replications (an integer).
`l`	Mean block size for stationary bootstrap or mean of the geometric distribution used to generate resampling blocks (an integer that is not longer than the length of the time series).
`maxlag`	Number of lags to include in the extremogram (an integer).
`quant1`	Quantile of the first time series to indicate an extreme event (a number between 0 and 1).
`quant2`	Quantile of the second series to indicate an extreme event (a number between 0 and 1).
`type`	Extremogram type (see function `extremogram2`).
`par`	If par = 1, the bootstrap replication procedure will be parallelized. If par = 0, no parallelization will be used.
`start`	The lag that the extremogram plots starts at (an integer not greater than `maxlag`, default is 1).
`cutoff`	The cutoff of the y-axis on the plot (a number between 0 and 1, default is 1).
`alpha`	Significance level for the confidence bands (a number between 0 and 1, default is 0.05).
`...`	further arguments: plot and axis names.

Value

Returns a plot of the confidence bands for the sample cross extremogram.

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega  = 1
alpha1 = 0.1
beta1  = 0.6
alpha2 = 0.11
beta2  = 0.78
n      = 1000
quant  = 0.95
type   = 1
maxlag = 70
df     = 3
R      = 10
l      = 30
par    = 0
G1     = extremogram:::garchsim(omega,alpha1,beta1,n,df)
G2     = extremogram:::garchsim(omega,alpha2,beta2,n,df)
data   = cbind(G1, G2)

extremogram2(data, quant, quant, maxlag, type, 1, 1, 0)
bootconf2(data, R, l, maxlag, quant, quant, type, par, 1, 1, 0.05)
```

`bootconfr` *Confidence bands for the sample return time extremogram*

Description

The function estimates confidence bands for the sample return time extremogram using the stationary bootstrap.

Usage

```
bootconfr(x, R, l, maxlag, uplevel = 1, lowlevel = 0, type, par,
  start = 1, cutoff = 1, alpha = 0.05, ...)
```

Arguments

`x`	Univariate time series (a vector).
`R`	Number of bootstrap replications (an integer).
`l`	Mean block size for stationary bootstrap or mean of the geometric distribution used to generate resampling blocks (an integer that is not longer than the length of the time series).
`maxlag`	Number of lags to include in the extremogram (an integer)
`uplevel`	Quantile of the time series to indicate a upper tail extreme event (a number between 0 and 1, default is 1).
`lowlevel`	Quantile of the time series to indicate a lower tail extreme event (a number between 0 and 1, default is 0).

type	Extremogram type (see function extremogramr).
par	If par = 1, the bootstrap replication procedure will be parallelized. If par = 0, no parallelization will be used.
start	The lag that the extremogram plots starts at (an integer not greater than maxlag, default is 1).
cutoff	The cutoff of the y-axis on the plot (a number between 0 and 1, default is 1).
alpha	Significance level for the confidence bands (a number between 0 and 1, default is 0.05).
...	further arguments: plot and axis names.

Value

Returns a plot of the confidence bands for the sample return time extremogram.

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega    = 1
alpha    = 0.1
beta     = 0.6
n        = 1000
uplevel  = 0.95
lowlevel = 0.05
type     = 3
maxlag   = 70
df       = 3
R        = 10
l        = 30
par      = 0
G = extremogram:::garchsim(omega,alpha,beta,n,df)

extremogramr(G, type, maxlag, uplevel, lowlevel, 1, 1)
bootconfr(G, R, l, maxlag, uplevel, lowlevel, type, par, 1, 1, 0.05)
```

---

extremogram1 *Sample univariate extremogram*

Description

The function estimates the sample univariate extremogram and creates an extremogram plot.

Usage

```
extremogram1(x, quant, maxlag, type, ploting = 1, cutoff = 1, start = 0,
  ...)
```

Arguments

`x`	Univariate time series (a vector).
`quant`	Quantile of the time series to indicate an extreme event (a number between 0 and 1).
`maxlag`	Number of lags to include in the extremogram (an integer).
`type`	Extremogram type. If type = 1, the upper tail extremogram is estimated. If type = 2, the lower tail extremogram is estimated.
`ploting`	An extremogram plot. If ploting = 1, a plot is created (default). If ploting = 0, no plot is created.
`cutoff`	The cutoff of the y-axis on the plot (a number between 0 and 1, default is 1).
`start`	The lag that the extremogram plots starts at (an integer not greater than `maxlag`, default is 0).
`...`	further arguments: plot and axis names.

Value

Extremogram values and a plot (if requested).

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega  = 1
alpha  = 0.1
beta   = 0.6
n      = 1000
quant  = 0.95
type   = 1
maxlag = 70
df     = 3
G = extremogram:::garchsim(omega,alpha,beta,n,df)

extremogram1(G, quant, maxlag, type, 1, 1, 0)
```

`extremogram2` *Sample cross extremogram*

Description

The function estimates the sample cross extremogram and creates an extremogram plot.

Usage

```
extremogram2(a, quant1, quant2, maxlag, type, ploting = 1, cutoff = 1,
  start = 0, ...)
```

Arguments

`a`	Bivariate time series (n by 2 matrix).
`quant1`	Quantile of the first time series to indicate an extreme event (a number between 0 and 1).
`quant2`	Quantile of the second time series to indicate an extreme event (a number between 0 and 1).
`maxlag`	Number of lags to include in the extremogram (an integer).
`type`	If type=1, the upper tail extremogram is estimated - P(Y>y,X>x). If type=2, the lower tail extremogram is estimated - P(Y<y,X<x). If type=3, the extremogram is estimated for a lower tail extreme value in the first time series and an upper tail extreme value in the second time series - P(Y>y,X<x). If type=4, the extremogram is estimated for a lower tail extreme value in the second time series and an upper tail extreme value in the first time series - P(Y<y,X>x).
`ploting`	An extremogram plot. If ploting = 1, a plot is created (default). If ploting = 0, no plot is created.
`cutoff`	The cutoff of the y-axis on the plot (a number between 0 and 1, default is 1).
`start`	The lag that the extremogram plots starts at (an integer not greater than `maxlag`, default is 0).
`...`	further arguments: plot and axis names.

Value

Cross extremogram values and a plot (if requested).

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega  = 1
alpha1 = 0.1
beta1  = 0.6
alpha2 = 0.11
beta2  = 0.78
n      = 1000
quant  = 0.95
type   = 1
maxlag = 70
df     = 3
G1     = extremogram:::garchsim(omega,alpha1,beta1,n,df)
G2     = extremogram:::garchsim(omega,alpha2,beta2,n,df)
data   = cbind(G1, G2)

extremogram2(data, quant, quant, maxlag, type, 1, 1, 0)
```

`extremogramr` *Sample return time extremogram*

Description

The function estimates the sample return time extremogram and creates an extremogram plot.

Usage

```
extremogramr(x, type, maxlag, uplevel = 1, lowlevel = 0, histogram = 1,
  cutoff = 1, ...)
```

Arguments

`x`	Univariate time series (a vector).
`type`	Extremogram type. If type = 1, the upper tail extremogram is estimated. If type = 2, the lower tail extremogram is estimated. If type = 3, both upper and lower tail extremogram is estimated.
`maxlag`	Number of lags to include in the extremogram (an integer).
`uplevel`	Quantile of the time series to indicate a upper tail extreme event (a number between 0 and 1, default is 1).
`lowlevel`	Quantile of the time series to indicate a lower tail extreme event (a number between 0 and 1, default is 0).
`histogram`	An extremogram plot. If histogram = 1, a plot is created (default). If histogram = 0, no plot is created.
`cutoff`	The cutoff of the y-axis on the plot (a number between 0 and 1, default is 1).
`...`	further arguments: plot and axis names.

Value

Extremogram values, return time for extreme events, mean return time and a plot (if requested).

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega    = 1
alpha    = 0.1
beta     = 0.6
n        = 1000
uplevel  = 0.95
lowlevel = 0.05
type     = 3
maxlag   = 70
df       = 3
G = extremogram:::garchsim(omega,alpha,beta,n,df)

extremogramr(G, type, maxlag, uplevel, lowlevel, 1, 1)
```

---

permfn1 — *Confidence bands for the sample univariate extremogram*

Description

The function estimates empirical confidence bands for the sample univariate extremogram via a permutation procedure under the assumption that the data are independent.

Usage

```
permfn1(x, p, m, type, exttype, maxlag, start = 1, alpha = 0.05)
```

Arguments

x	Univariate time series (a vector).
p	Quantile of the time series to indicate an extreme event (a number between 0 and 1).
m	Number of permutations (an integer).
type	Type of confidence bands. If type=1, it adds all permutations to the sample extremogram plot. If type=2, it adds the alpha/2 and (1-alpha)/2 empirical confidence bands for each lag. If type=3, it calculates the lag 1 alpha/2 and (1-alpha)/2 empirical confidence bands lag and uses them for all of the lags.

exttype	Extremogram type (see extremogram1).
maxlag	Number of lags to include in the extremogram (an integer).
start	The lag that the extremogram plots starts at (an integer not greater than maxlag, default is 1).
alpha	Significance level for the confidence bands (a number between 0 and 1, default is 0.05).

Value

The empirical confidence bands are added to the sample univariate extremogram plot.

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega   = 1
alpha   = 0.1
beta    = 0.6
n       = 1000
quant   = 0.95
exttype = 1
maxlag  = 70
df      = 3
type    = 3
m       = 10
G = extremogram:::garchsim(omega,alpha,beta,n,df)

extremogram1(G, quant, maxlag, exttype, 1, 1, 0)
permfn1(G, quant, m, type, exttype, maxlag, 1, 0.05)
```

---

permfn2 *Confidence bands for the sample cross extremogram*

Description

The function estimates empirical confidence bands for the sample cross extremogram via a permutation procedure under the assumption that the data are independent.

Usage

```
permfn2(x, p1, p2, m, type, exttype, maxlag, start = 1, alpha = 0.05)
```

Arguments

x	Bivariate time series (n by 2 matrix).
p1	Quantile of the first time series to indicate an extreme event (a number between 0 and 1).
p2	Quantile of the second time series to indicate an extreme event (a number between 0 and 1).
m	Number of permutations (an integer).
type	Type of confidence bands. If type=1, it adds all permutations to the sample extremogram plot. If type=2, it adds the `alpha`/2 and (1-`alpha`)/2 empirical confidence bands for each lag. If type=3, it calculates the lag 1 `alpha`/2 and (1-`alpha`)/2 empirical confidence bands lag and uses them for all of the lags.
exttype	Extremogram type (see `extremogram2`).
maxlag	Number of lags to include in the extremogram (an integer).
start	The lag that the extremogram plots starts at (an integer not greater than `maxlag`, default is 1).
alpha	Significance level for the confidence bands (a number between 0 and 1, default is 0.05).

Value

The empirical confidence bands are added to the sample cross extremogram plot.

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega   = 1
alpha1  = 0.1
beta1   = 0.6
alpha2  = 0.11
beta2   = 0.78
n       = 1000
quant   = 0.95
exttype = 1
maxlag  = 70
df      = 3
type    = 3
m       = 10
G1      = extremogram:::garchsim(omega,alpha1,beta1,n,df)
G2      = extremogram:::garchsim(omega,alpha2,beta2,n,df)
data    = cbind(G1, G2)
```

```
extremogram2(data, quant, quant, maxlag, type, 1, 1, 0)
permfn2(data, quant, quant, m, type, exttype, maxlag, 1, 0.05)
```

---

permfnr *Confidence bands for the sample return time extremogram*

Description

The function estimates empirical confidence bands for the sample returt time extremogram via a permutation procedure under the assumption that the data are independent.

Usage

```
permfnr(x, m, type, exttype, maxlag, uplevel = 1, lowlevel = 0, start = 1,
  alpha = 0.05)
```

Arguments

`x`	Univariate time series (a vector).
`m`	Number of permutations (an integer).
`type`	Type of confidence bands. If type=1, it adds all permutations to the sample extremogram plot. If type=2, it adds the `alpha`/2 and (1-`alpha`)/2 empirical confidence bands for each lag. If type=3, it calculates the lag 1 `alpha`/2 and (1-`alpha`)/2 empirical confidence bands lag and uses them for all of the lags.
`exttype`	Extremogram type (see `extremogramr`).
`maxlag`	Number of lags to include in the extremogram (an integer).
`uplevel`	Quantile of the time series to indicate a upper tail extreme event (a number between 0 and 1, default is 1).
`lowlevel`	Quantile of the time series to indicate a lower tail extreme event (a number between 0 and 1, default is 0).
`start`	The lag that the extremogram plots starts at (an integer not greater than `maxlag`, default is 1).
`alpha`	Significance level for the confidence bands (a number between 0 and 1, default is 0.05).

References

1. Davis, R. A., Mikosch, T., & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. Journal of Econometrics,170(1), 142-152.
2. Davis, R. A., Mikosch, T., & Cribben, I. (2011). Estimating extremal dependence in univariate and multivariate time series via the extremogram.arXiv preprint arXiv:1107.5592.

Examples

```
# generate a GARCH(1,1) process
omega    = 1
alpha    = 0.1
beta     = 0.6
n        = 1000
uplevel  = 0.95
lowlevel = 0.05
exttype  = 3
maxlag   = 70
type     = 3
m        = 10
df       = 3
G = extremogram:::garchsim(omega,alpha,beta,n,df)

extremogramr(G, type, maxlag, uplevel, lowlevel, 1, 1)
permfnr(G, m, type, exttype, maxlag, uplevel, lowlevel, 1, 0.05)
```

Index

A.0.3 Appendix 3

Intercept	QLD (lag 1)
14.22441	2.17666
(0.78674)	*(0.12236)*
Intercept	**QLD (lag 48)**
44.69375	1.69766
(0.43872)	*(0.21072)*
Intercept	**QLD (lag 96)**
102.70195	0.48040
(0.00000)	*(0.13817)*

Intercept	SA (lag 1)
107.56082	0.33981
(0.00000)	*(0.00035)*
Intercept	**SA (lag 48)**
94.62020	0.58163
(0.01145)	*(0.51101)*
Intercept	**SA (lag 96)**
99.57194	0.71198
(0.00000)	*(0.00000)*

Intercept	TAS (lag 1)
127.73722	0.00248
(0.00015)	*(0.99760)*
Intercept	**TAS (lag 96)**
126.20501	0.04896
(0.00000)	*(0.49312)*

Intercept	VIC (lag 1)
59.59740	1.12128
(0.58001)	*(0.69178)*
Intercept	**VIC (lag 48)**
28.23487	2.11012
(0.00000)	*(0.00000)*
Intercept	**VIC (lag 96)**
46.54683	1.93507
(0.00004)	*(0.00000)*

Table A.1: *Estimated QR models for NSW using one predictor at lag h (specified separately for each model), $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**NSW (lag 1)**
-33.23774	-0.00326	3.86442
(0.01539)	*(0.84831)*	*(0.00000)*
Intercept	**QLD (lag 48)**	**NSW (lag 48)**
-18.39471	-0.00347	3.39593
(0.77475)	*(0.99623)*	*(0.06892)*
Intercept	**QLD (lag 96)**	**NSW (lag 96)**
11.99812	0.06405	2.58221
(0.95016)	*(0.67025)*	*(0.63888)*

Intercept	**SA (lag 1)**	**NSW (lag 1)**
-44.39694	0.03511	5.04100
(0.57333)	*(0.76377)*	*(0.03947)*
Intercept	**SA (lag 48)**	**NSW (lag 48)**
-11.70523	0.23251	2.86402
(0.92936)	*(0.10778)*	*(0.44793)*
Intercept	**SA (lag 96)**	**NSW (lag 96)**
31.31101	0.41566	1.78317
(0.70521)	*(0.08280)*	*(0.43506)*

Intercept	**TAS (lag 1)**	**NSW (lag 1)**
-33.19013	-0.00285	3.86395
(0.57861)	*(0.98688)*	*(0.04397)*
Intercept	**TAS (lag 48)**	**NSW (lag 48)**
-18.23732	-0.00341	3.38603
(0.92416)	*(0.95987)*	*(0.53757)*
Intercept	**TAS (lag 96)**	**NSW (lag 96)**
10.72805	0.04622	2.64998
(0.95015)	*(0.89584)*	*(0.61950)*

Intercept	**VIC (lag 1)**	**NSW (lag 1)**
-34.17078	0.08780	3.85938
(0.27170)	*(0.01343)*	*(0.00184)*
Intercept	**VIC (lag 48)**	**NSW (lag 48)**
-11.83119	0.72732	2.31899
(0.88422)	*(0.33559)*	*(0.53866)*
Intercept	**VIC (lag 96)**	**NSW (lag 96)**
22.41415	0.97991	1.41082
(0.53116)	*(0.43300)*	*(0.00023)*

Table A.2: *Estimated QR models for NSW using one predictor and NSW at lag h (specified separately for each model), $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
-35.79831	-0.00299	-0.00357	-0.00343	3.35085	0.56128	-0.00269
(0.20990)	*(0.84111)*	*(0.91470)*	*(0.83429)*	*(0.00631)*	*(0.00000)*	*(0.92802)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
-36.98930	-0.00350	0.03819	0.07406	3.39797	0.52931	-0.00409
(0.66730)	*(0.98690)*	*(0.86505)*	*(0.45955)*	*(0.24866)*	*(0.30748)*	*(0.95108)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
-36.34528	0.07014	0.09022	0.06541	3.27430	0.46315	-0.00413
(0.98049)	*(0.68139)*	*(0.86629)*	*(0.93797)*	*(0.00000)*	*(0.44774)*	*(0.98846)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
-35.63340	-0.00320	-0.00282	-0.00017	3.34825	0.55772	-0.00404
(0.03370)	*(0.97782)*	*(0.72505)*	*(0.98660)*	*(0.00006)*	*(0.18005)*	*(0.95523)*

Table A.3: *Estimated QR models for NSW using each of the other states and NSW at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	QLD (lag 1)
45.03242	2.17806
(0.98878)	*(0.85190)*
Intercept	**QLD (lag 48)**
66.06208	1.69767
(0.00000)	*(0.00000)*
Intercept	**QLD (lag 96)**
83.44513	0.40552
(0.00001)	*(0.44050)*

Intercept	SA (lag 1)
85.41773	0.33991
(0.00000)	*(0.00023)*
Intercept	**SA (lag 48)**
80.59169	0.57269
(0.00928)	*(0.00000)*
Intercept	**SA (lag 96)**
91.82294	0.71203
(0.00000)	*(0.00000)*

Intercept	TAS (lag 1)
92.44967	0.00018
(0.00000)	*(0.99938)*
Intercept	**TAS (lag 48)**
92.52577	-0.00203
(0.00000)	*(0.95839)*
Intercept	**TAS (lag 96)**
92.10384	0.03088
(0.00000)	*(0.02790)*

Intercept	VIC (lag 1)
65.69960	0.91225
(0.00002)	*(0.47043)*
Intercept	**VIC (lag 48)**
66.69631	1.88722
(0.20784)	*(0.40737)*
Intercept	**VIC (lag 96)**
79.16123	1.93508
(0.00185)	*(0.00114)*

Table A.4: *Estimated QR models for NSW using one predictor at lag h (specified separately for each model) for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**NSW (lag 1)**
60.50722	-0.00317	3.86793
(0.00002)	*(0.96427)*	*(0.33863)*
Intercept	**QLD (lag 48)**	**NSW (lag 48)**
60.14399	-0.00331	3.54469
(0.03233)	*(0.99760)*	*(0.67302)*
Intercept	**QLD (lag 96)**	**NSW (lag 96)**
67.43810	0.05054	2.11864
(0.00004)	*(0.39554)*	*(0.64734)*
Intercept	**SA (lag 1)**	**NSW (lag 1)**
61.32061	0.04089	3.86594
(0.00000)	*(0.38712)*	*(0.00002)*
Intercept	**SA (lag 48)**	**NSW (lag 48)**
58.72528	0.23170	3.14760
(0.04170)	*(0.32865)*	*(0.00020)*
textbfIntercept	**SA (lag 96)**	**NSW (lag 96)**
74.38864	0.44144	1.44112
(0.00000)	*(0.13482)*	*(0.28424)*
Intercept	**TAS (lag 1)**	**NSW (lag 1)**
60.50752	-0.00271	3.86756
(0.00000)	*(0.99168)*	*(0.08518)*
Intercept	**TAS (lag 48)**	**NSW (lag 48)**
59.92127	-0.00328	3.53389
(0.00229)	*(0.98854)*	*(0.33753)*
Intercept	**TAS (lag 96)**	**NSW (lag 96)**
67.98428	0.05544	2.17468
(0.00000)	*(0.94992)*	*(0.41463)*
Intercept	**VIC (lag 1)**	**NSW (lag 1)**
62.22476	0.08840	3.86358
(0.00000)	*(0.28269)*	*(0.08188)*
Intercept	**VIC (lag 48)**	**NSW (lag 48)**
55.76794	0.68074	2.52439
(0.00000)	*(0.66941)*	*(0.55492)*
Intercept	**VIC (lag 96)**	**NSW (lag 96)**
69.69614	0.79283	1.46478
(0.00167)	*(0.59634)*	*(0.00021)*

Table A.5: *Estimated QR models for NSW using one predictor and NSW at lag h (specified separately for each model) for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
58.93041	-0.00297	-0.00346	-0.00333	3.35220	0.56222	-0.00271
(0.00000)	*(0.96039)*	*(0.96219)*	*(0.77979)*	*(0.29927)*	*(0.00000)*	*(0.98090)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
61.60552	0.02068	0.05049	0.07238	3.35057	0.54155	-0.00499
(0.00000)	*(0.89178)*	*(0.78860)*	*(0.18135)*	*(0.11196)*	*(0.05161)*	*(-0.00499)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
59.97877	0.07082	0.08644	0.06567	3.27509	0.46897	-0.00519
(0.00000)	*(0.66042)*	*(0.40500)*	*(0.45545)*	*(0.00000)*	*(0.00000)*	*(0.22447)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
58.80381	-0.00298	-0.00158	-0.00148	3.34968	0.55904	-0.00502
(0.00000)	*(0.95184)*	*(0.89263)*	*(0.93680)*	*(0.00003)*	*(0.05076)*	*(0.89739)*

Table A.6: *Estimated QR models for NSW using each of the other states and NSW at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**NSW (lag 1)**
14.22441	2.17666
(0.78674)	*(0.12236)*
Intercept	**NSW (lag 48)**
44.69375	1.69766
(0.43872)	*(0.21072)*
Intercept	**NSW (lag 96)**
102.70195	0.48040
(0.00000)	*(0.13817)*
Intercept	**SA (lag 1)**
107.56082	0.33981
(0.00000)	*(0.00035)*
Intercept	**SA (lag 48)**
94.62020	0.58163
(0.01145)	*(0.51101)*
Intercept	**SA (lag 96)**
99.57194	0.71198
(0.00000)	*(0.00000)*
Intercept	**VIC (lag 1)**
59.59740	1.12128
(0.58001)	*(0.69178)*
Intercept	**VIC (lag 48)**
28.23487	2.11012
(0.00000)	*(0.00000)*
Intercept	**VIC (lag 96)**
46.54683	1.93507
(0.00004)	*(0.00000)*
Intercept	**TAS (lag 1)**
127.73722	0.00248
(0.00015)	*(0.99760)*
Intercept	**TAS (lag 96)**
126.20501	0.04896
(0.00000)	*(0.49312)*

Table A.7: *Estimated QR models for QLD using one predictor at lag h (specified separately for each model), $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**NSW (lag 1)**	**QLD (lag 1)**
-33.23774	-0.00326	3.86442
(0.01539)	*(0.84831)*	*(0.00000)*
Intercept	**NSW (lag 48)**	**QLD (lag 48)**
-18.39471	-0.00347	3.39593
(0.77475)	*(0.99623)*	*(0.06892)*
Intercept	**NSW (lag 96)**	**QLD (lag 96)**
11.99812	0.06405	2.58221
(0.95016)	*(0.67025)*	*(0.63888)*
Intercept	**SA (lag 1)**	**QLD (lag 1)**
-33.34964	0.00966	3.86355
(0.42755)	*(0.31272)*	*(0.00755)*
Intercept	**SA (lag 48)**	**QLD (lag 48)**
-11.70523	0.23251	2.86402
(0.92936)	*(0.10778)*	*(0.44793)*
textbfIntercept	**SA (lag 96)**	**QLD (lag 96)**
31.31101	0.41566	1.78317
(0.70521)	*(0.08280)*	*(0.43506)*
Intercept	**TAS (lag 1)**	**QLD (lag 1)**
-33.19013	-0.00285	3.86395
(0.57861)	*(0.98688)*	*(0.04397)*
Intercept	**TAS (lag 48)**	**QLD (lag 48)**
-18.23732	-0.00341	3.38603
(0.92416)	*(0.95987)*	*(0.53757)*
Intercept	**TAS (lag 96)**	**QLD (lag 96)**
10.72805	0.04622	2.64998
(0.95015)	*(0.89584)*	*(0.61950)*
Intercept	**VIC (lag 1)**	**QLD (lag 1)**
-34.17078	0.08780	3.85938
(0.72360)	*(0.01343)*	*(0.00184)*
Intercept	**VIC (lag 48)**	**QLD (lag 48)**
-11.83119	0.72732	2.31899
(0.88422)	*(0.33559)*	*(0.53866)*
Intercept	**VIC (lag 96)**	**QLD (lag 96)**
22.41415	0.97991	1.41082
(0.53116)	*(0.43300)*	*(0.00023)*

Table A.8: *Estimated QR models for QLD using one predictor and QLD at lag h (specified separately for each model), $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
-35.79831	-0.00299	-0.00357	-0.00343	3.35085	0.56128	-0.00269
(0.20990)	*(0.84111)*	*(0.91470)*	*(0.83429)*	*(0.00631)*	*(0.00000)*	*(0.92802)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
-36.98930	-0.00350	0.03819	0.07406	3.39797	0.52931	-0.00409
(0.66730)	*(0.98690)*	*(0.86505)*	*(0.45955)*	*(0.24866)*	*(0.30748)*	*(0.95108)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
-36.34528	0.07014	0.09022	0.06541	3.27430	0.46315	0.44774
(0.98049)	*(0.68139)*	*(0.86629)*	*(0.93797)*	*(0.00000)*	*(0.44774)*	*(0.98846)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
-35.63340	-0.00320	-0.00282	-0.00017	3.34825	0.55772	0.18005
(0.03370)	*(0.97782)*	*(0.72505)*	*(0.98660)*	*(0.00006)*	*(0.18005)*	*(0.95523)*

Table A.9: *Estimated QR models for QLD using each of the other states and QLD at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	NSW (lag 1)
45.03242	2.17806
(0.98878)	*(0.85190)*

Intercept	NSW (lag 48)
66.06208	1.69767
(0.00000)	*(0.00000)*

Intercept	NSW (lag 96)
83.44513	0.40552
(0.00001)	*(0.44050)*

Intercept	SA (lag 1)
85.41773	0.33991
(0.00000)	*(0.00023)*

Intercept	SA (lag 48)
80.59169	0.57269
(0.00928)	*(0.00000)*

Intercept	SA (lag 96)
91.82294	0.71203
(0.00000)	*(0.00000)*

Intercept	TAS (lag 1)
92.44967	0.00018
(0.00000)	*(0.99938)*

Intercept	TAS (lag 48)
92.52577	-0.00203
(0.00000)	*(0.95839)*

Intercept	TAS (lag 96)
92.10384	0.03088
(0.00000)	*(0.02790)*

Intercept	VIC (lag 1)
65.69960	0.91225
(0.00002)	*(0.47043)*

Intercept	VIC (lag 48)
66.69631	1.88722
(0.20784)	*(0.40737)*

Intercept	VIC (lag 96)
79.16123	1.93508
(0.00185)	*(0.00114)*

Table A.10: *Estimated QR models for QLD using one predictor at lag h (specified separately for each model) for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**NSW (lag 1)**	**QLD (lag 1)**
60.50722	-0.00317	3.86793
(0.00002)	*(0.96427)*	*(0.33863)*
Intercept	**NSW (lag 48)**	**QLD (lag 48)**
60.14399	-0.00331	3.54469
(0.03233)	*(0.99760)*	*(0.67302)*
Intercept	**NSW (lag 96)**	**QLD (lag 96)**
67.43810	0.05054	2.11864
(0.00004)	*(0.39554)*	*(0.64734)*
Intercept	**SA (lag 1)**	**QLD (lag 1)**
61.32061	0.04089	3.86594
(0.00000)	*(0.38712)*	*(0.00002)*
Intercept	**SA (lag 48)**	**QLD (lag 48)**
58.72528	0.23170	3.14760
(0.04170)	*(0.32865)*	*(0.00020)*
textbfIntercept	**SA (lag 96)**	**QLD (lag 96)**
74.38864	0.44144	1.44112
(0.00000)	*(0.13482)*	*(0.28424)*
Intercept	**TAS (lag 1)**	**QLD (lag 1)**
60.50752	-0.00271	3.86756
(0.00000)	*(0.99168)*	*(0.08518)*
Intercept	**TAS (lag 48)**	**QLD (lag 48)**
59.92127	-0.00328	3.53389
(0.00229)	*(0.98854)*	*(0.33753)*
Intercept	**TAS (lag 96)**	**QLD (lag 96)**
67.98428	0.05544	2.17468
(0.00000)	*(0.94992)*	*(0.41463)*
Intercept	**VIC (lag 1)**	**QLD (lag 1)**
62.22476	0.08840	3.86358
(0.00000)	*(0.28269)*	*(0.08188)*
Intercept	**VIC (lag 48)**	**QLD (lag 48)**
55.76794	0.68074	2.52439
(0.00000)	*(0.66941)*	*(0.55492)*
Intercept	**VIC (lag 96)**	**QLD (lag 96)**
69.69614	0.79283	1.46478
(0.00167)	*(0.59634)*	*(0.00021)*

Table A.11: *Estimated QR models for QLD using one predictor and QLD at lag h (specified separately for each model) for the deseasonalized data,* $\tau = 0.995$*. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
58.88936	3.34964	0.55901	-0.00503
(0.00000)	*(0.22688)*	*(0.20571)*	*(0.95275)*

Table A.12: *Estimated QR models for QLD using QLD at lags 1, 48 and 96 as predictors for deseasonalized data set, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
58.93041	-0.00297	-0.00346	-0.00333	3.35220	0.56222	-0.00271
(0.00000)	*(0.96039)*	*(0.96219)*	*(0.77979)*	*(0.29927)*	*(0.00000)*	*(0.98090)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
61.60552	0.02068	0.05049	0.07238	3.35057	0.54155	-0.00499
(0.00000)	*(0.89178)*	*(0.78860)*	*(0.18135)*	*(0.11196)*	*(0.05161)*	*(0.96174)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
59.97877	0.07082	0.08644	0.06567	3.27509	0.46897	-0.00519
(0.00000)	*(0.66042)*	*(0.40500)*	*(0.45545)*	*(0.00000)*	*(0.00000)*	*(0.78960)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
58.80381	-0.00298	-0.00158	-0.00148	3.34968	0.55904	-0.00502
(0.00000)	*(0.95184)*	*(0.89263)*	*(0.93680)*	*(0.00003)*	*(0.05076)*	*(0.89739)*

Table A.13: *Estimated QR models for QLD using each of the other states and QLD at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**SA (lag 1)**
29.32469	0.03699	1.85125
(0.49258)	*(0.91709)*	*(0.04710)*
Intercept	**QLD (lag 48)**	**SA (lag 48)**
136.66261	0.13683	1.16941
(0.00000)	*(0.78947)*	*(0.00812)*
Intercept	**QLD (lag 96)**	**SA (lag 96)**
179.42646	0.03106	0.98191
(0.00120)	*(0.97026)*	*(0.05096)*
Intercept	**NSW (lag 1)**	**SA (lag 1)**
10.43728	0.58450	1.83216
(0.90929)	*(0.80491)*	*(0.00055)*
Intercept	**NSW (lag 48)**	**SA (lag 48)**
121.69576	0.40151	1.17024
(0.65646)	*(0.95294)*	*(0.60475)*
textbfIntercept	**NSW (lag 96)**	**SA (lag 96)**
178.93671	0.02980	0.98194
(0.00243)	*(0.46814)*	*(0.00000)*
Intercept	**TAS (lag 1)**	**SA (lag 1)**
31.72468	-0.00183	1.85103
(0.39933)	*(0.99209)*	*(0.00000)*
Intercept	**TAS (lag 48)**	**SA (lag 48)**
143.11380	-0.01202	1.16917
(0.00000)	*(0.97951)*	*(0.00000)*
Intercept	**TAS (lag 96)**	**SA (lag 96)**
181.89585	-0.01557	0.98193
(0.03143)	*(0.99306)*	*(0.02494)*
Intercept	**VIC (lag 1)**	**SA (lag 1)**
7.35713	0.96106	1.48560
(0.92373)	*(0.69076)*	*(0.15185)*
Intercept	**VIC (lag 48)**	**SA (lag 48)**
66.70433	2.01422	0.98234
(0.20885)	*(0.10599)*	*(0.41063)*
Intercept	**VIC (lag 96)**	**SA (lag 96)**
172.64115	0.18137	0.97910
(0.55116)	*(0.97910)*	*(0.70386)*

Table A.14: *Estimated QR models for SA using one predictor and SA at lag h (specified separately for each model), $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**
112.50573	2.11538	0.34425	0.18118
(0.49371)	*(0.58853)*	*(0.83837)*	*(0.95287)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
-69.20180	7.87351	-0.01121	-0.01054
(0.92183)	*(0.68164)*	*(0.99465)*	*(0.99084)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
-29.81378	5.46102	0.25245	0.16662
(0.66386)	*(0.00015)*	*(0.88306)*	*(0.89890)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
232.28807	0.02508	0.01046	0.02093
(0.10169)	*(0.99128)*	*(0.97148)*	*(0.99265)*

Table A.15: *Estimated QR models for NSW using each of the other states at lags 1, 48 and 96 as predictors, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	SA (lag 1)	SA (lag 48)	SA (lag 96)
27.50610	1.71328	0.27106	0.05614
(0.70914)	*(0.29586)*	*(0.44503)*	*(0.46410)*

Table A.16: *Estimated QR models for SA using SA at lags 1, 48 and 96 as predictors, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
27.56941	0.00619	-0.00191	-0.00408	1.71324	0.27148	0.05624
(0.42243)	*(0.97411)*	*(0.97966)*	*(0.91448)*	*(0.00162)*	*(0.59891)*	*(0.83532)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
14.92845	0.40894	-0.00059	-0.00549	1.71046	0.24546	0.05236
(0.72824)	*(0.01708)*	*(0.98019)*	*(0.96397)*	*(0.01213)*	*(0.11673)*	*(0.93670)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
4.49097	0.97030	-0.00401	0.08795	1.43945	0.25853	0.02679
(0.96188)	*(0.72317)*	*(0.99223)*	*(0.87974)*	*(0.00002)*	*(0.70730)*	*(0.88473)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
28.87975	-0.00508	0.00155	-0.00342	1.71242	0.27620	0.05767
(0.61103)	*(0.97772)*	*(0.99549)*	*(0.99121)*	*(0.09214)*	*(0.57295)*	*(0.83480)*

Table A.17: *Estimated QR models for SA using each of the other states and SA at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	QLD (lag 1)
175.74926	1.44204
(0.00000)	*(0.50489)*
Intercept	**QLD (lag 48)**
188.17660	0.36998
(0.39531)	*(0.82365)*
Intercept	**QLD (lag 96)**
194.18325	0.28699
(0.00187)	*(0.49270)*

Intercept	NSW (lag 1)
135.73749	6.88502
(0.13132)	*(0.81242)*
Intercept	**NSW (lag 48)**
155.28897	1.89346
(0.93915)	*(0.87502)*
Intercept	**NSW (lag 96)**
179.07792	0.90499
(0.00040)	*(0.80542)*

Intercept	TAS (lag 1)
198.06969	0.02611
(0.00000)	*(0.96291)*
Intercept	**TAS (lag 96)**
202.24650	0.01775
(0.00000)	*(0.95443)*

Intercept	VIC (lag 1)
124.65933	5.71064
(0.00000)	*(0.00081)*
Intercept	**VIC (lag 48)**
126.33515	4.41997
(0.03910)	*(0.01020)*
Intercept	**VIC (lag 96)**
177.34340	0.93796
(0.26108)	*(0.93895)*

Table A.18: *Estimated QR models for SA using one predictor at lag h (specified separately for each model) for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	QLD (lag 1)	SA (lag 1)
57.97522	0.03846	1.87103
(0.00032)	*(0.94311)*	*(0.04654)*
Intercept	**QLD (lag 48)**	**SA (lag 48)**
146.75021	0.06825	1.16944
(0.00872)	*(0.94661)*	*(0.35269)*
Intercept	**QLD (lag 96)**	**SA (lag 96)**
181.05104	-0.01445	0.98183
(0.000003)	*(0.89401)*	*(0.00000)*

Intercept	NSW (lag 1)	SA (lag 1)
55.44352	0.60557	1.84629
(0.31774)	*(0.87193)*	*(0.03202)*
Intercept	**NSW (lag 48)**	**SA (lag 48)**
146.07407	0.06536	1.16952
(0.00000)	*(0.62382)*	*(0.12661)*
Intercept	**NSW (lag 96)**	**SA (lag 96)**
179.98325	0.01008	0.98191
(0.02000)	*(0.73829)*	*(0.12892)*

Intercept	TAS (lag 1)	SA (lag 1)
58.43932	-0.00123	1.87495
(0.55097)	*(0.99769)*	*(0.20029)*
Intercept	**TAS (lag 48)**	**SA (lag 48)**
148.61395	-0.01189	1.16919
(0.06419)	*(0.98483)*	*(0.37475)*
Intercept	**TAS (lag 96)**	**SA (lag 96)**
180.62578	-0.01556	0.98191
(0.01276)	*(0.99429)*	*(0.42337)*

Intercept	VIC (lag 1)	SA (lag 1)
54.16271	0.97766	1.52345
(0.30880)	*(0.70788)*	*(0.06804)*
Intercept	**VIC (lag 48)**	**SA (lag 48)**
136.79213	1.51295	0.98315
(0.05174)	*(0.69972)*	*(0.57832)*
Intercept	**VIC (lag 96)**	**SA (lag 96)**
181.14250	-0.00502	0.98188
(0.00000)	*(0.99702)*	*(0.00000)*

Table A.19: *Estimated QR models for SA using one predictor and SA at lag h (specified separately for each model) for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
61.48716	0.006711	-0.00146	-0.00379	1.71257	0.28227	0.06829
(0.03653)	*(0.99038)*	*(0.98768)*	*(0.93739)*	*(0.00146)*	*(0.03025)*	*(0.82111)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
61.11964	0.45702	0.00213	-0.00487	1.70805	0.27970	0.06596
(0.40073)	*(0.77675)*	*(0.99580)*	*(0.98833)*	*(0.13559)*	*(0.76236)*	*(0.89603)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
58.72938	0.98028	-0.00343	0.07136	1.44057	0.25834	0.04730
(0.00000)	*(0.01439)*	*(0.98928)*	*(0.95756)*	*(0.00000)*	*(0.62404)*	*(0.69271)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
61.67090	-0.00476	0.00199	-0.00377	1.71240	0.28204	0.06860
(0.03700)	*(0.97823)*	*(0.98780)*	*(0.98965)*	*(0.00000)*	*(0.55488)*	*(0.78297)*

Table A.20: *Estimated QR models for SA using each of the other states and SA at lags 1, 48 and 96 as predictors for the deseasonalized data,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**VIC (lag 1)**
-9.11693	0.06061	2.39732
(0.87049)	*(0.81077)*	*(0.22462)*
Intercept	**QLD (lag 48)**	**VIC (lag 48)**
0.48485	0.02799	2.92131
(0.99615)	*(0.98577)*	*(0.22161)*
Intercept	**QLD (lag 96)**	**VIC (lag 96)**
80.50883	0.11592	0.65246
(0.07521)	*(0.84455)*	*(0.67126)*

Intercept	**SA (lag 1)**	**VIC (lag 1)**
-5.68868	0.17085	2.10899
(0.89006)	*(0.18250)*	*(0.11026)*
Intercept	**SA (lag 48)**	**VIC (lag 48)**
7.36157	0.26328	2.35919
(0.95648)	*(0.80701)*	*(0.53247)*
textbfIntercept	**SA (lag 96)**	**VIC (lag 96)**
84.26517	0.06748	0.60986
(0.00000)	*(0.76681)*	*(0.04511)*

Intercept	**TAS (lag 1)**	**VIC (lag 1)**
-8.67323	-0.00142	2.47251
(0.86849)	*(0.88665)*	*(0.16176)*
Intercept	**TAS (lag 48)**	**VIC (lag 48)**
0.47018	-0.00485	2.97173
(0.99445)	*(0.88826)*	*(0.11131)*
Intercept	**TAS (lag 96)**	**VIC (lag 96)**
84.82352	-0.00525	0.67031
(0.00728)	*(0.99061)*	*(0.34275)*

Intercept	**NSW (lag 1)**	**VIC (lag 1)**
-14.57636	0.34328	2.36533
(0.68528)	*(0.64580)*	*(0.09230)*
Intercept	**NSW (lag 48)**	**VIC (lag 48)**
-3.77388	0.21111	2.88536
(0.97742)	*(0.95971)*	*(0.21339)*
Intercept	**NSW (lag 96)**	**VIC (lag 96)**
62.67706	0.56928	0.65512
(0.23138)	*(0.66284)*	*(0.00000)*

Table A.21: *Estimated QR models for VIC using one predictor and VIC at lag h (specified separately for each model),* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	QLD (lag 1)	QLD (lag 48)	QLD (lag 96)
70.82095	0.70415	0.17278	0.07372
(0.00672)	*(0.14959)*	*(0.65005)*	*(0.37540)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
32.07926	1.20378	0.20867	0.01150
(0.42936)	*(0.00000)*	*(0.71982)*	*(0.93465)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
-17.76033	2.98851	-0.00170	-0.00311
(0.86488)	*(0.31246)*	*(0.99438)*	*(0.93353)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
110.98827	0.08416	0.02276	0.01876
(0.00000)	*(0.35430)*	*(0.65321)*	*(0.85342)*

Table A.22: *Estimated QR models for VIC using each of the other states at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	VIC (lag 1)	VIC (lag 48)	VIC (lag 96)
-12.00219	2.38342	0.20091	0.10641
(0.85091)	*(0.26092)*	*(0.60683)*	*(0.94039)*

Table A.23: *Estimated QR models for VIC using VIC at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
-12.12397	0.04647	-0.00293	-0.00205	2.38212	0.19783	0.01034
(0.67176)	*(0.92379)*	*(0.94436)*	*(0.94063)*	*(0.00947)*	*(0.45300)*	*(0.98396)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
-7.77061	0.16933	0.02717	0.00561	1.92975	0.17687	0.03233
(0.74523)	*(0.30369)*	*(0.02055)*	*(0.85379)*	*(0.00613)*	*(0.64498)*	*(0.95725)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
-15.82413	0.25981	-0.00242	-0.00312	2.36808	0.16646	0.00488
(0.58728)	*(0.01010)*	*(0.95893)*	*(0.82302)*	*(0.04005)*	*(0.65438)*	*(0.97941)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
-12.03441	-0.00279	0.00011	-0.00064	2.38341	0.20076	0.04531
(0.71778)	*(0.82896)*	*(0.99299)*	*(0.96185)*	*(0.05602)*	*(0.74672)*	*(0.94193)*

Table A.24: *Estimated QR models for VIC using each of the other states and VIC at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	**QLD (lag 1)**
68.35838	0.70508
(0.00017)	*(0.00000)*
Intercept	**QLD (lag 48)**
79.86536	0.30022
(0.00002)	*(0.02264)*
Intercept	**QLD (lag 96)**
83.13649	0.14720
(0.00166)	*(0.91728)*
Intercept	**SA (lag 1)**
45.86676	1.20525
(0.00002)	*(0.00000)*
Intercept	**SA (lag 48)**
65.90306	0.83139
(0.05298)	*(0.08063)*
Intercept	**SA (lag 96)**
82.69939	0.12174
(0.00000)	*(0.76106)*
Intercept	**TAS (lag 1)**
80.78948	0.08454
(0.26050)	*(0.96436)*
Intercept	**TAS (lag 48)**
84.53524	0.04705
(0.00000)	*(0.77194)*
Intercept	**TAS (lag 96)**
84.81606	0.03484
(0.61937)	*(0.98907)*
Intercept	**NSW (lag 1)**
45.39299	2.79510
(0.00580)	*(0.53905)*
Intercept	**NSW (lag 48)**
57.18625	1.96785
(0.00194)	*(0.29802)*
Intercept	**NSW (lag 96)**
71.69860	0.86859
(0.00063)	*(0.66893)*

Table A.25: *Estimated QR models for VIC using one predictor at lag h (specified separately for each model) for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**VIC (lag 1)**
36.99030	0.05721	2.38522
(0.00000)	*(0.88002)*	*(0.17782)*
Intercept	**QLD (lag 48)**	**VIC (lag 48)**
63.63323	0.02854	2.93858
(0.34640)	*(0.97954)*	*(0.55149)*
Intercept	**QLD (lag 96)**	**VIC (lag 96)**
72.44630	0.07453	0.65144
(0.00000)	*(0.81130)*	*(0.16487)*
Intercept	**SA (lag 1)**	**VIC (lag 1)**
35.16823	0.19632	2.08556
(0.00000)	*(0.10345)*	*(0.17330)*
Intercept	**SA (lag 48)**	**VIC (lag 48)**
60.85325	0.30247	2.35583
(0.00000)	*(0.50652)*	*(0.03299)*
Intercept	**SA (lag 96)**	**VIC (lag 96)**
73.75071	0.06955	0.56980
(0.00000)	*(0.77099)*	*(0.02931)*
Intercept	**TAS (lag 1)**	**VIC (lag 1)**
37.23265	-0.00027	2.42078
(0.00000)	*(0.96530)*	*(0.00000)*
Intercept	**TAS (lag 48)**	**VIC (lag 48)**
63.86682	-0.00448	2.96021
(0.02974)	*(0.97207)*	*(0.25272)*
Intercept	**TAS (lag 96)**	**VIC (lag 96)**
73.51360	-0.00534	0.65113
(0.00077)	*(0.98814)*	*(0.00000)*
Intercept	**NSW (lag 1)**	**VIC (lag 1)**
39.36958	0.25804	2.371078
(0.00000)	*(0.14678)*	*(0.30639)*
Intercept	**NSW (lag 48)**	**VIC (lag 48)**
64.06369	0.20994	2.91413
(0.12303)	*(0.52222)*	*(0.09900)*
Intercept	**NSW (lag 96)**	**VIC (lag 96)**
70.49986	0.34733	0.65378
(0.00000)	*(0.47412)*	*(0.00172)*

Table A.26: *Estimated QR models for VIC using one predictor and VIC at lag h (specified separately for each model) for the deseasonalized data,* $\tau = 0.995$*. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
39.84489	0.05936	-0.00252	-0.00183	2.38234	0.197893	-0.00061
(0.00000)	*(0.86138)*	*(0.94391)*	*(0.92702)*	*(0.00118)*	*(0.45965)*	*(0.99921)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
34.72722	0.16976	0.02886	0.00730	1.93171	0.17381	0.02160
(0.00000)	*(0.41755)*	*(0.06146)*	*(0.89942)*	*(0.30896)*	*(0.84930)*	*(0.97007)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
41.90123	0.25261	-0.00218	-0.00287	2.36898	0.16784	0.00472
(0.00000)	*(0.67372)*	*(0.98630)*	*(0.82284)*	*(0.09877)*	*(0.16931)*	*(0.99443)*
Intercept	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
39.92525	-0.00254	0.00104	-0.00099	2.38397	0.20234	0.03974
(0.00006)	*(0.86991)*	*(0.95176)*	*(0.94349)*	*(0.28830)*	*(0.03318)*	*(0.94388)*

Table A.27: *Estimated QR models for VIC using each of the other states and VIC at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	QLD (lag 1)
141.79012	0.12295
(0.00000)	*(0.75994)*
Intercept	**QLD (lag 48)**
145.60111	-0.00760
(0.00000)	*(0.91696)*
Intercept	**QLD (lag 96)**
145.64924	-0.00700
(0.00000)	*(0.85261)*

Intercept	SA (lag 1)
140.16916	0.15882
(0.00000)	*(0.56107)*
Intercept	**SA (lag 48)**
143.50788	0.17511
(0.00000)	*(0.62206)*
Intercept	**SA (lag 96)**
144.45895	0.00235
(0.00000)	*(0.98420)*

Intercept	NSW (lag 1)
123.61402	0.52304
(0.17762)	*(0.82982)*
Intercept	**NSW (lag 48)**
145.75637	-0.00843
(0.00000)	*(0.98256)*
Intercept	**NSW (lag 96)**
143.55338	0.04108
(0.03376)	*(0.94440)*

Intercept	VIC (lag 1)
115.02462	0.66216
(0.00000)	*(0.17143)*
Intercept	**VIC (lag 48)**
100.73934	1.35897
(0.43759)	*(0.38993)*

Table A.28: *Estimated QR models for TAS using one predictor at lag h (specified separately for each model),* $\tau = 0.995$*. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**TAS (lag 1)**
42.72702	0.06167	1.04424
(0.08806)	*(0.89762)*	*(0.05206)*
Intercept	**QLD (lag 48)**	**TAS (lag 48)**
117.59814	0.03361	0.99030
(0.00000)	*(0.90391)*	*(0.00000)*
Intercept	**QLD (lag 96)**	**TAS (lag 96)**
141.26272	0.05624	0.10139
(0.00000)	*(0.45944)*	*(0.78468)*

Intercept	**SA (lag 1)**	**TAS (lag 1)**
42.12136	0.10327	1.06793
(0.24322)	*(0.77850)*	*(0.20956)*
Intercept	**SA (lag 48)**	**TAS (lag 48)**
110.32343	0.28351	0.98895
(0.00016)	*(0.42675)*	*(0.06832)*
textbfIntercept	**SA (lag 96)**	**TAS (lag 96)**
137.93269	0.08755	0.13548
(0.00106)	*(0.87569)*	*(0.86129)*

Intercept	**NSW (lag 1)**	**TAS (lag 1)**
28.25307	0.44707	1.07480
(0.91907)	*(0.83998)*	*(0.01428)*
Intercept	**NSW (lag 48)**	**TAS (lag 48)**
86.65837	0.75085	0.98750
(0.23955)	*(0.71608)*	*(0.05365)*
Intercept	**NSW (lag 96)**	**TAS (lag 96)**
136.28200	0.17224	0.11079
(0.08337)	*(0.89386)*	*(0.20208)*

Intercept	**VIC (lag 1)**	**TAS (lag 1)**
20.65300	0.64199	1.13999
(0.12860)	*(0.00000)*	*(0.00000)*
Intercept	**VIC (lag 48)**	**TAS (lag 48)**
66.50537	1.07930	1.00766
(0.02701)	*(0.20702)*	*(0.00000)*
Intercept	**VIC (lag 96)**	**TAS (lag 96)**
136.65752	0.14175	0.13284
(0.00000)	*(0.78075)*	*(0.75649)*

Table A.29: *Estimated QR models for TAS using one predictor and TAS at lag h (specified separately for each model), $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
38.74335	0.09222	-0.00045	-0.00310	0.91780	0.31532	0.05293
(0.16437)	*(0.87061)*	*(0.99641)*	*(0.99164)*	*(0.00000)*	*(0.19368)*	*(0.69878)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
32.94793	0.06465	0.13701	0.04769	0.94123	0.46270	0.05452
(0.14163)	*(0.68715)*	*(0.49392)*	*(0.63184)*	*(0.00001)*	*(0.33934)*	*(0.90613)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
2.61942	0.56358	0.27933	0.05843	1.03250	0.61571	0.04591
(0.95285)	*(0.00000)*	*(0.63200)*	*(0.75911)*	*(0.13193)*	*(0.52887)*	*(0.89933)*
Intercept	**NSW (lag 1)**	**NSW(lag 48)**	**NSW (lag 96)**	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
8.86270	0.63632	0.21964	0.00028	0.92110	0.45195	0.05728
(0.86458)	*(0.29251)*	*(0.87024)*	*(0.99661)*	*(0.00002)*	*(0.17775)*	*(0.80482)*

Table A.30: *Estimated QR models for TAS using each of the other states and TAS at lags 1, 48 and 96 as predictors,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	QLD (lag 1)
111.96437	0.12400
(0.00001)	*(0.90863)*
Intercept	**QLD (lag 48)**
111.77356	-0.00766
(0.00000)	*(0.48098)*
Intercept	**QLD (lag 96)**
111.67372	-0.00704
(0.98662)	*(0.99695)*

Intercept	SA (lag 1)
112.21183	0.15871
(0.00048)	*(0.83016)*
Intercept	**SA (lag 48)**
114.90650	0.17527
(0.99882)	*(0.99879)*
Intercept	**SA (lag 96)**
110.64865	0.00242
(0.00584)	*(0.99500)*

Intercept	NSW (lag 1)
105.94638	0.53326
(0.09037)	*(0.81572)*
Intercept	**NSW (lag 48)**
111.88531	-0.00850
(0.00704)	*(0.98311)*
Intercept	**NSW (lag 96)**
110.81834	0.04127
(0.00000)	*(0.83766)*

Intercept	VIC (lag 1)
100.10672	0.66294
(0.00000)	*(0.00022)*
Intercept	**VIC (lag 48)**
109.15782	1.35982
(0.00000)	*(0.40662)*
Intercept	**VIC (lag 96)**
110.75859	0.00701
(0.00000)	*(0.93876)*

Table A.31: *Estimated QR models for TAS using one predictor at lag h (specified separately for each model) for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**TAS (lag 1)**
45.46341	0.05877	1.04350
(0.96291)	*(0.95640)*	*(0.11400)*
Intercept	**QLD (lag 48)**	**TAS (lag 48)**
118.05009	0.03379	0.99034
(0.01299)	*(0.96714)*	*(0.26421)*
Intercept	**QLD (lag 96)**	**TAS (lag 96)**
111.97793	0.05665	0.09848
(0.00000)	*(0.01433)*	*(0.65790)*

Intercept	**SA (lag 1)**	**TAS (lag 1)**
47.35543	0.10863	1.06836
(0.00000)	*(0.02728)*	*(0.00000)*
Intercept	**SA (lag 48)**	**TAS (lag 48)**
120.30070	0.28340	0.98913
(0.00000)	*(0.41124)*	*(0.00703)*
textbfIntercept	**SA (lag 96)**	**TAS (lag 96)**
111.57473	0.07404	0.13327
(0.00000)	*(0.77831)*	*(0.72308)*

Intercept	**NSW (lag 1)**	**TAS (lag 1)**
43.50121	0.44927	1.07714
(0.23006)	*(0.90066)*	*(0.05271)*
Intercept	**NSW (lag 48)**	**TAS (lag 48)**
110.49102	0.70391	0.98844
(0.00000)	*(0.67012)*	*(0.03181)*
Intercept	**NSW (lag 96)**	**TAS (lag 96)**
111.23786	0.06974	0.10899
(0.00000)	*(0.00523)*	*(0.71716)*

Intercept	**VIC (lag 1)**	**TAS (lag 1)**
44.70080	0.64248	1.14101
(0.00000)	*(0.00000)*	*(0.00000)*
Intercept	**VIC (lag 48)**	**TAS (lag 48)**
101.34039	1.0861	1.00891
(0.00002)	*(0.41183)*	*(0.21664)*
Intercept	**VIC (lag 96)**	**TAS (lag 96)**
111.67924	0.11721	0.13339
(0.00000)	*(0.64750)*	*(0.41993)*

Table A.32: *Estimated QR models for TAS using one predictor and TAS at lag h (specified separately for each model) for the deseasonalized data,* $\tau = 0.995$. *p-values given in parenthesis.*

Intercept	QLD (lag 1)	QLD (lag 48)	QLD (lag 96)
112.14326	0.12384	-0.00821	-0.00682
(0.93179)	*(0.93937)*	*(0.99434)*	*(0.99854)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**
114.59806	0.13511	0.16826	-0.00109
(0.00000)	*(0.13684)*	*(0.00000)*	*(0.52991)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**
105.93968	0.62773	0.81540	-0.01017
(0.00001)	*(0.64460)*	*(0.12082)*	*(0.83922)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**
105.92345	0.62773	-0.01246	0.04365
(0.00255)	*(0.78375)*	*(0.93750)*	*(0.44075)*

Table A.33: *Estimated QR models for TAS using each of the other states at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	TAS (lag 1)	TAS (lag 48)	TAS (lag 96)
51.64350	0.91201	0.32547	0.05097
(0.98208)	*(0.00000)*	*(0.87153)*	*(0.98126)*

Table A.34: *Estimated QR models for TAS using TAS at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

Intercept	**QLD (lag 1)**	**QLD (lag 48)**	**QLD (lag 96)**	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
50.71039	0.09157	-0.00017	-0.00309	0.91751	0.31589	0.05330
(0.00346)	*(0.88932)*	*(0.99948)*	*(0.99536)*	*(0.00000)*	*(0.52522)*	*(0.89750)*
Intercept	**SA (lag 1)**	**SA (lag 48)**	**SA (lag 96)**	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
56.96354	0.06607	0.13595	0.04764	0.94513	0.46029	0.05703
(0.00000)	*(0.56106)*	*(0.03716)*	*(0.71012)*	*(0.01751)*	*(0.01251)*	*(0.82944)*
Intercept	**VIC (lag 1)**	**VIC (lag 48)**	**VIC (lag 96)**	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
53.20616	0.56270	0.28207	0.05896	1.02785	0.60948	0.05565
(0.00000)	*(0.16670)*	*(0.03047)*	*(0.41739)*	*(0.37006)*	*(0.17584)*	*(0.77383)*
Intercept	**NSW (lag 1)**	**NSW (lag 48)**	**NSW (lag 96)**	**TAS (lag 1)**	**TAS (lag 48)**	**TAS (lag 96)**
50.24934	0.80102	0.20772	0.00064	0.91735	0.44361	0.06035
(0.11944)	*(0.51042)*	*(0.92233)*	*(0.99800)*	*(0.00003)*	*(0.00412)*	*(0.95003)*

Table A.35: *Estimated QR models for TAS using each of the other states and TAS at lags 1, 48 and 96 as predictors for the deseasonalized data, $\tau = 0.995$. p-values given in parenthesis.*

A.0.4 Appendix 4

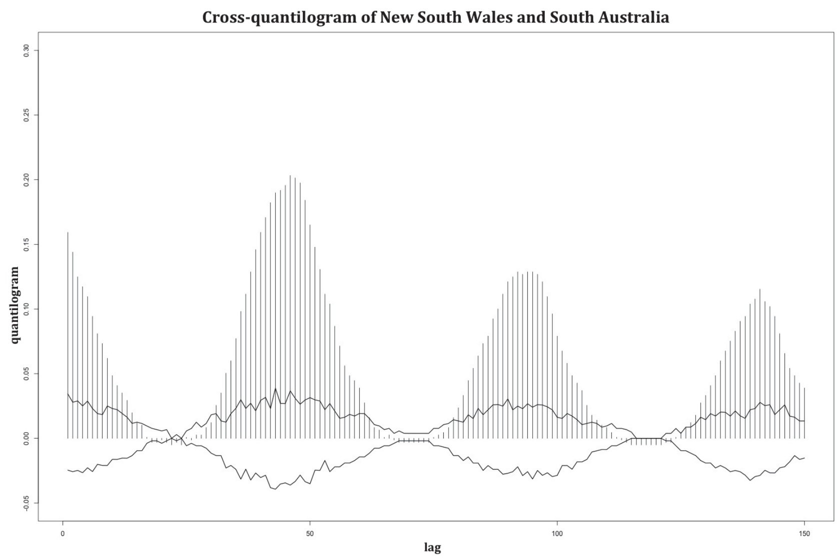

Figure A.0.47: *A: Cross-quantilograms for NSW conditioning on SA (2009 – 2014).*

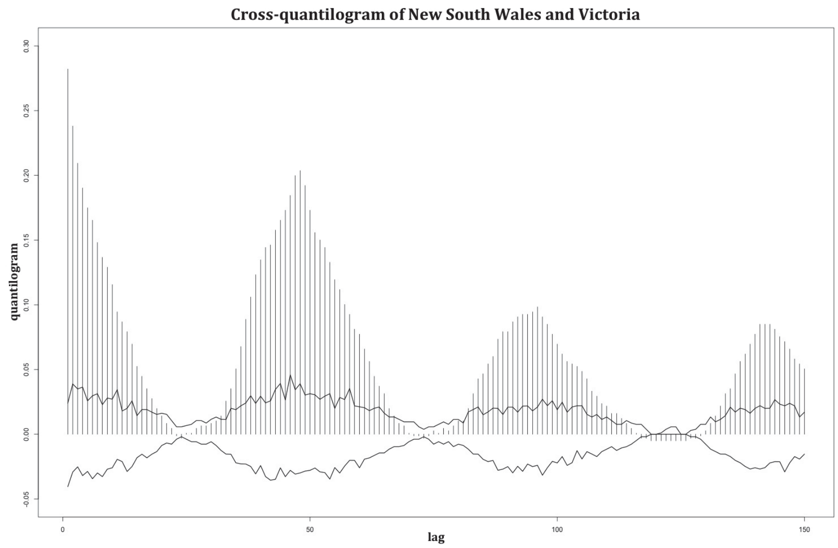

Figure A.0.48: *A: Cross-quantilograms for NSW conditioning on VIC (2009 – 2014).*

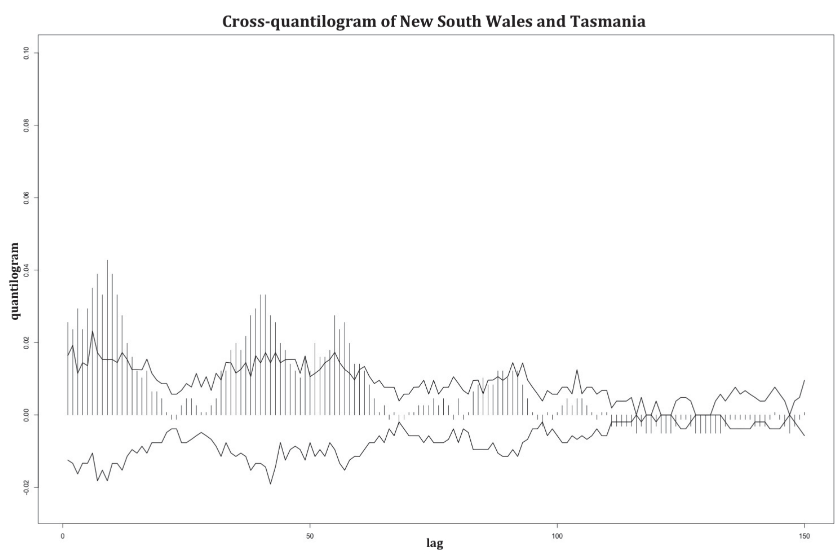

Figure A.0.49: *A: Cross-quantilograms for NSW conditioning on TAS (2009 – 2014).*

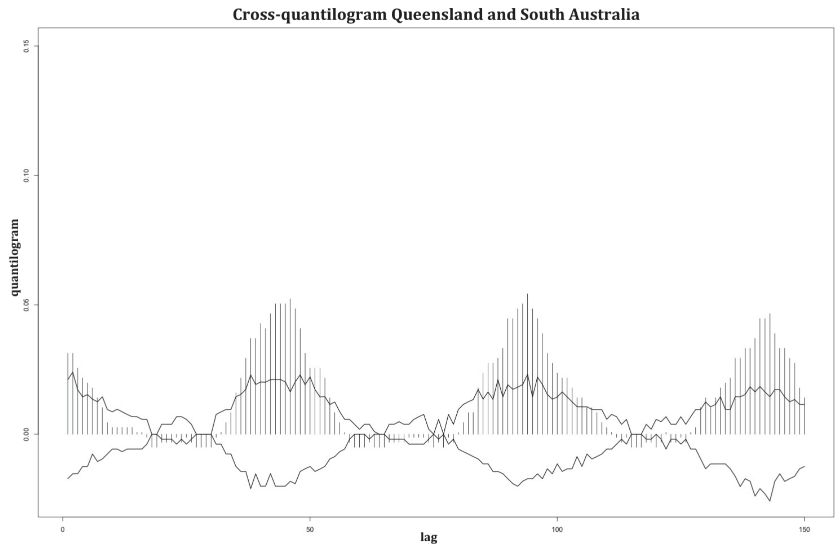

Figure A.0.50: *A: Cross-quantilograms for QLD conditioning on SA (2009 – 2014).*

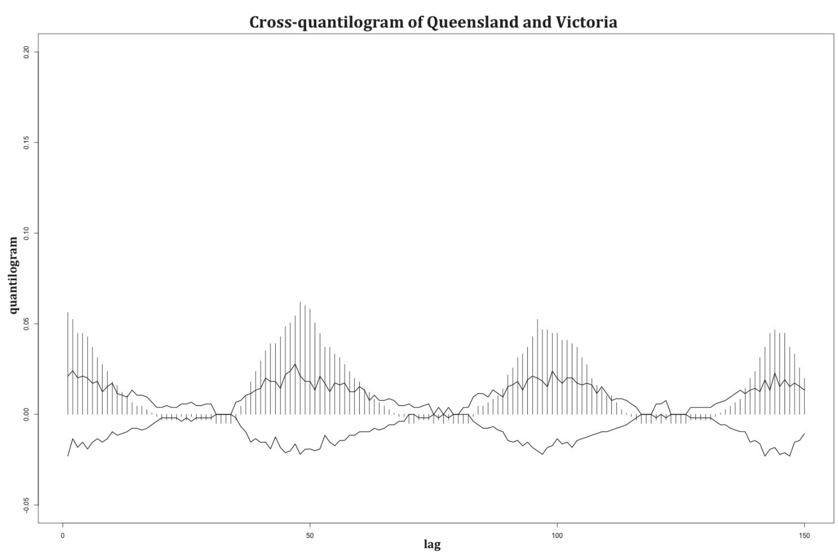

Figure A.0.51: *A: Cross-quantilograms for QLD conditioning on VIC (2009 – 2014).*

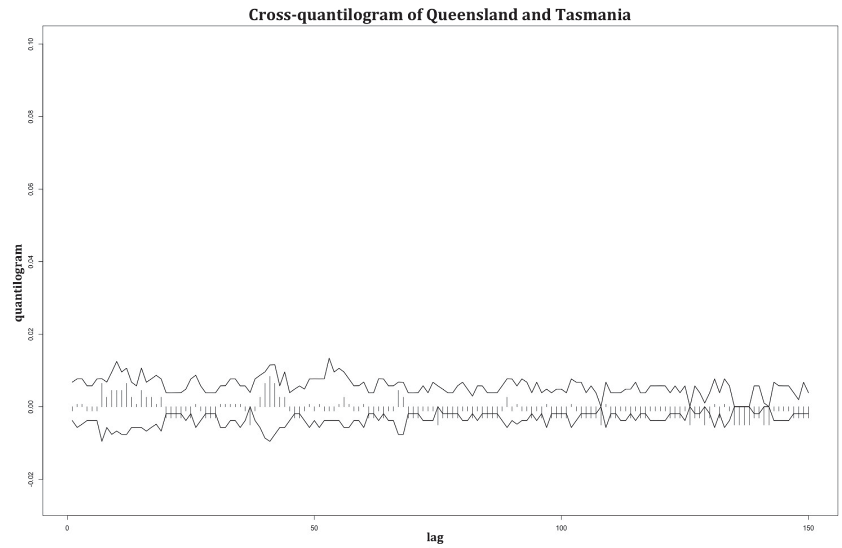

Figure A.0.52: *A: Cross-quantilograms for QLD conditioning on TAS (2009 – 2014).*

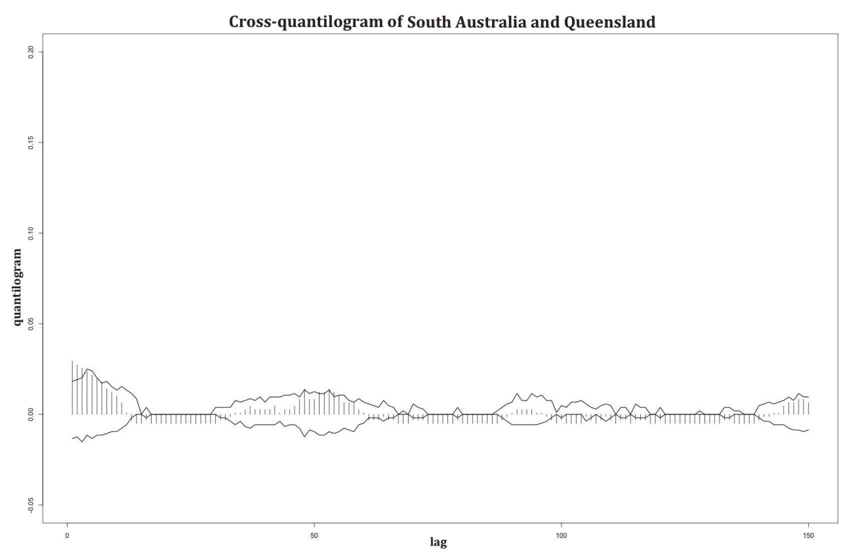

Figure A.0.53: *A: Cross-quantilograms for SA conditioning on QLD (2009 – 2014).*

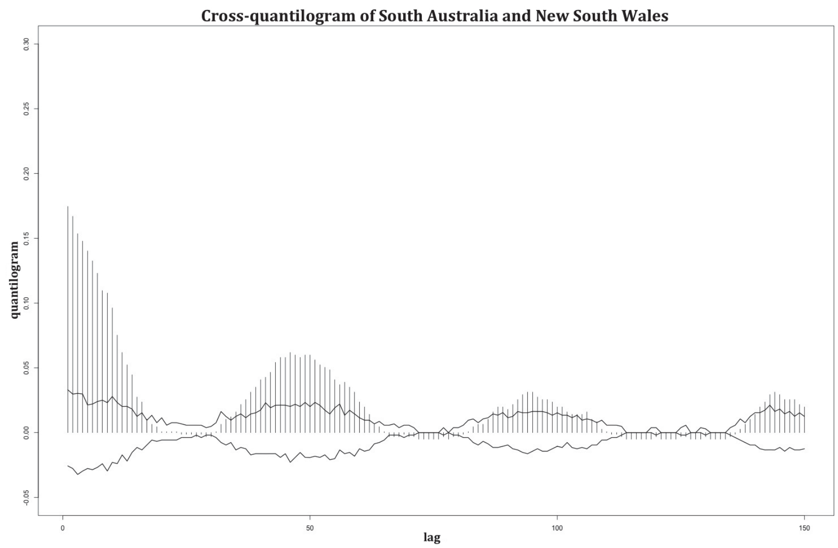

Figure A.0.54: *A: Cross-quantilograms for SA conditioning on NSW (2009 – 2014).*

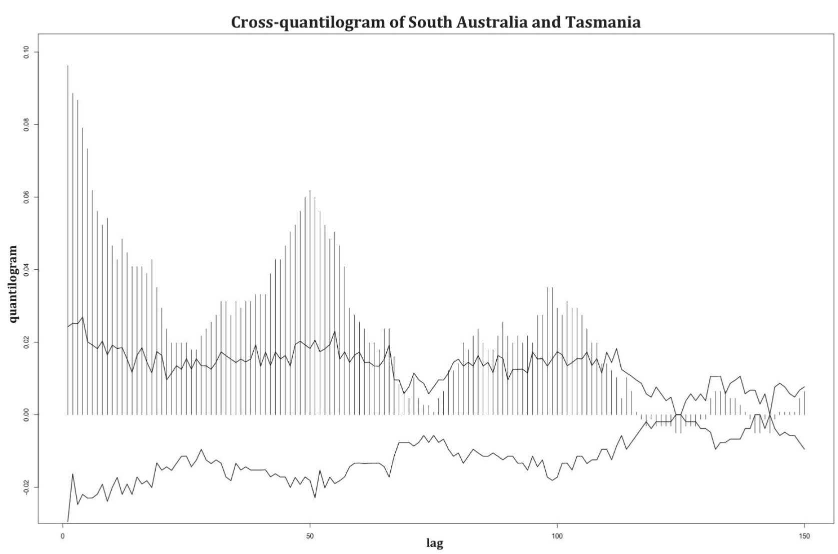

Figure A.0.55: *A: Cross-quantilograms for SA conditioning on TAS (2009 – 2014).*

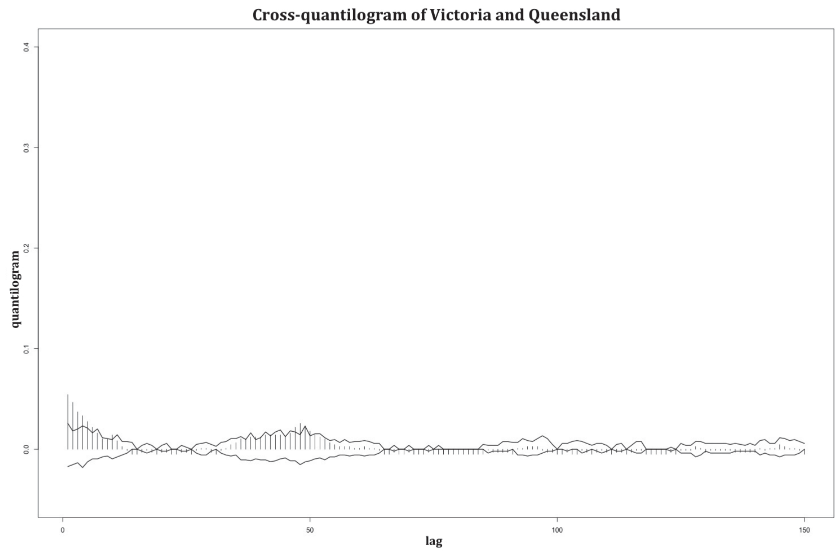

Figure A.0.56: *A: Cross-quantilograms for VIC conditioning on QLD (2009 – 2014).*

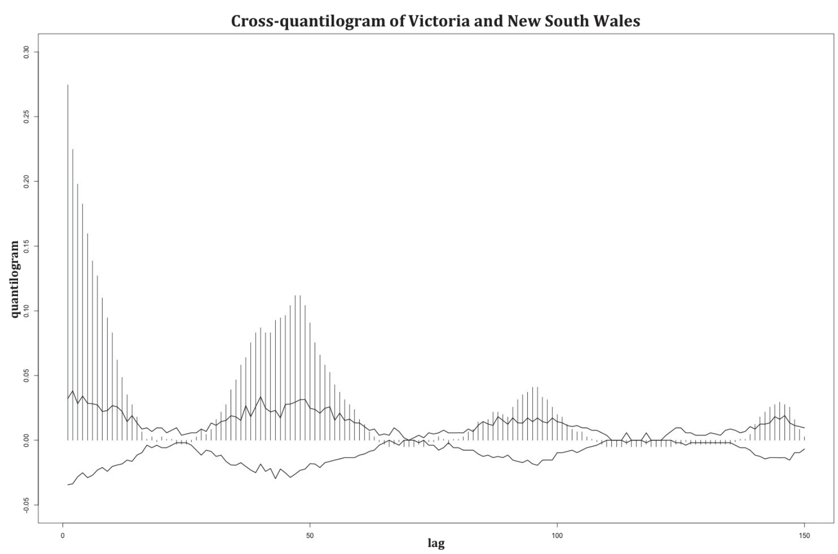

Figure A.0.57: *A: Cross-quantilograms for VIC conditioning on NSW (2009 – 2014).*

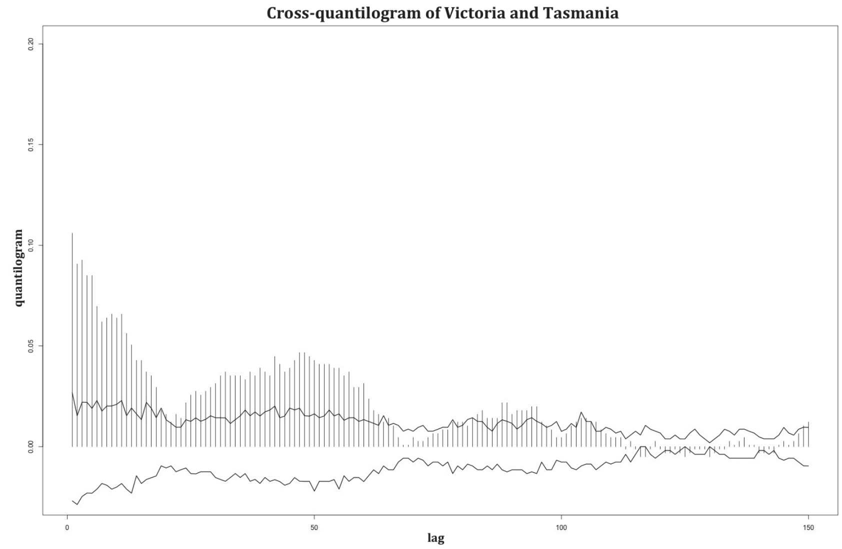

Figure A.0.58: *A: Cross-quantilograms for VIC conditioning on TAS (2009 – 2014).*

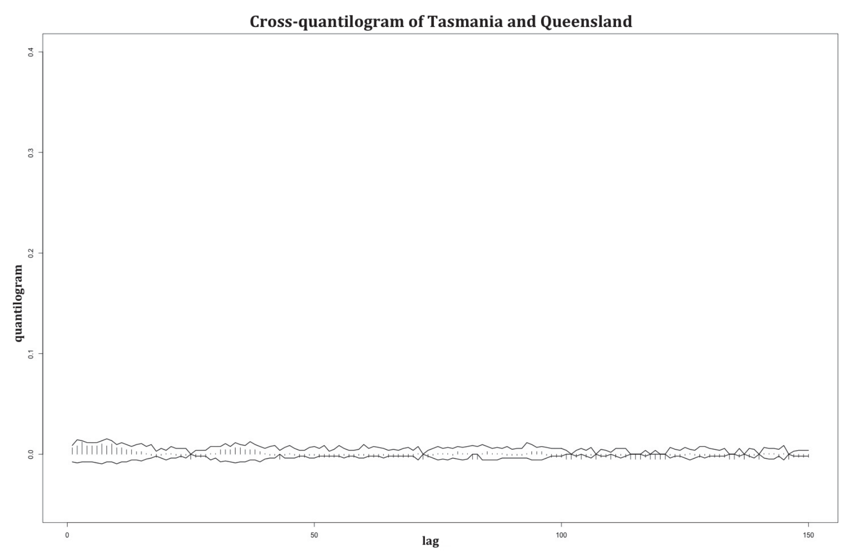

Figure A.0.59: *A: Cross-quantilograms for TAS conditioning on QLD (2009 – 2014).*

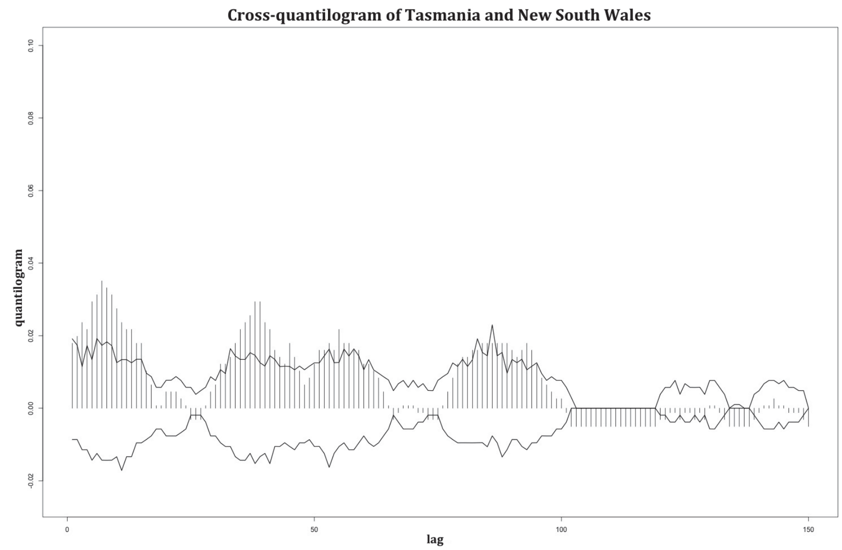

Figure A.0.60: *A: Cross-quantilograms for TAS conditioning on NSW (period: 2009 – 2014).*

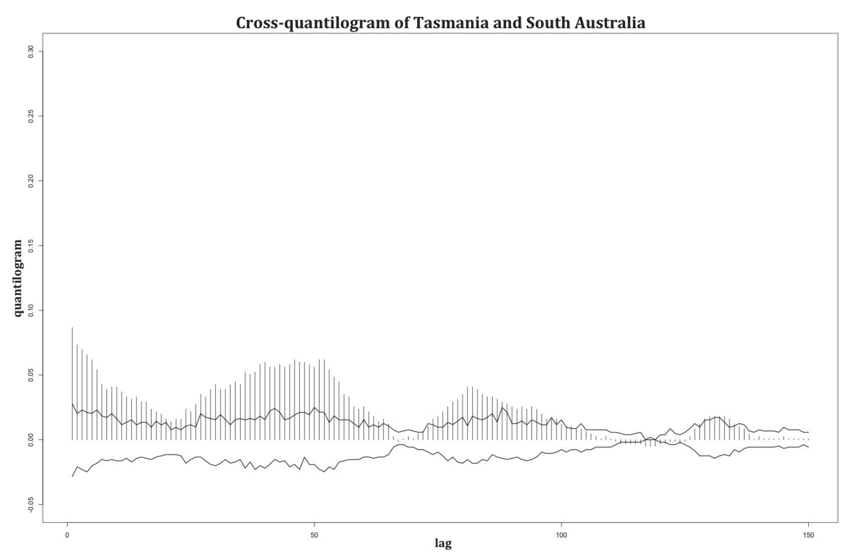

Figure A.0.61: *A: Cross-quantilograms for TAS conditioning on SA (2009 – 2014).*

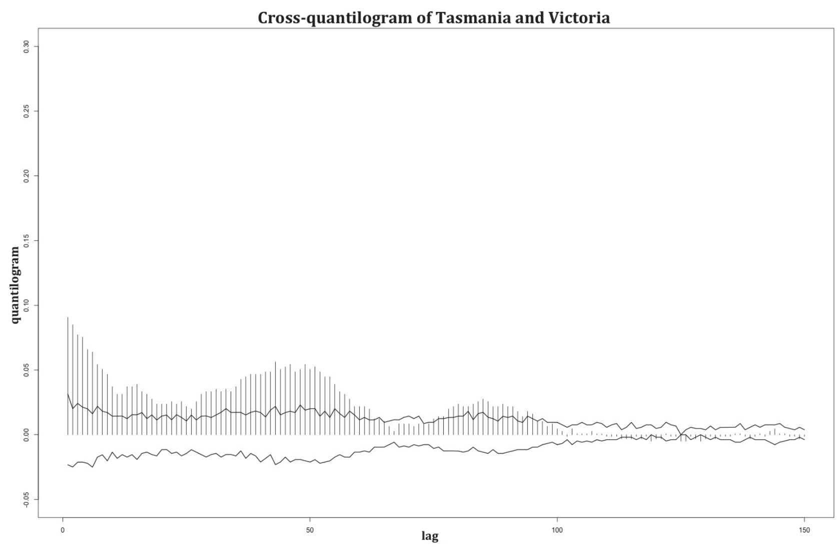

Figure A.0.62: *A: Cross-quantilograms for TAS conditioning on VIC (2009 – 2014).*

Bibliography

[1] Davis R. A and T. Mikosch. Limit theory for the sample acf of stationary process with heavy tails with applications to arch. *Annals of Statistics*, 26:2049–2080, 1998.

[2] Davis R. A., T. Mikosch, and Y. Zhao. Measures of serial extremal dependence and their estimation. *Stochastic Processes and their Applications*, 123(7):2575–2602, 2013.

[3] Davis R. A., Mikosch T., and I. Cribben. Towards estimating extremal serial dependence via the bootstrapped extremogram. *Journal of Econometrics*, 170(1):142–152, 2012.

[4] Davis R. A., Mikosch T., et al. The extremogram: A correlogram for extreme events. *Bernoulli*, 15(4):977–1009, 2009.

[5] Davis R. A., Mikosch T., and Cribben I. Estimating extremal dependence in univariate and multivariate time series via the extremogram. *arXiv preprint arXiv:1107.5592*, 2011.

[6] Ledford A.W. and Tawn J. A. Diagnostics for dependence within time series extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):521–543, 2003.

[7] Hill J. B. et al. Gaussian tests of" extremal white noise" for dependent, heterogeneous, heavy tailed time series with an application. *Manuscript, Florida International University*, 2005.

[8] Nelsen R. B. *An Introduction to Copulas (Springer Series in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[9] Davison A. C. and Smith R. L. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 393–442, 1990.

[10] Box G. EP., Jenkins G. M., Reinsel G. C., and Ljung G. M. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[11] Engle R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[12] Engle R. F. and Manganelli S. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381, 2004.

[13] De Luca G. and Rivieccio G. Multivariate tail dependence coefficients. *Statistical methods for the analysis of large data-sets, Book of short papers*, 2009.

[14] Kendall M. G. A new measure of rank correlation. *Biometrika*, 30:81–93, 1938.

[15] Han H., Linton O., Tatsushi O., , and Whang Y. The cross-quantilogram: Measuring quantile dependence and testing directional predictability between time series. *Journal of Econometrics*, 193(1):251 – 270, 2016.

[16] Fan J. and Li R. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96:1348–1360, 2001.

[17] Galambos J. *The Asymptotic Theory of Extreme Order Statistics*. Krieger R. E., Malabar, FL, 1987. Second Edition.

[18] Smith R. L. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, pages 367–377, 1989.

[19] Sklar M. *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8, 1959.

[20] Frolova N. and Cribben I. *extremogram: Estimation of Extreme Value Dependence for Time Series Data*, 2015. R package version 1.0.1.

[21] Politis D. N. and Romano J. P. The stationary bootstrap. *Journal of the American Statistical association*, 89(428):1303–1313, 1994.

[22] Linton O. and Whang Y. The quantilogram: With an application to evaluating directional predictability. *Journal of Econometrics*, 141(1):250 – 282, 2007.

[23] Hall P., Peng L., and Q. Yao. Moving-maximum models for extrema of time series. *Journal of Statistical Planning and Inference*, 103(1):51–63, 2002.

[24] Koenker R. Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1):74 – 89, 2004.

[25] Koenker R. and Bassett Jr G. Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50, 1978.

[26] Koenker R. and Z. Xiao. Quantile autoregression. *Journal of the American Statistical Association*, 101(475):980–990, 2006.

[27] Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–288, 1994.

[28] Coles S., Bawa J., Trenner L., and Dorazio P. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.

[29] Li T. Quantile periodograms. *Journal of the American Statistical Association*, 107(498):765–776, 2012.

[30] Li T. Quantile periodogram and time-dependent variance. *Journal of Time Series Analysis*, 35(4):322–340, 2014.

[31] Fasen V., Klüppelberg C., and M. Schlather. High-level dependence in time series models. *Extremes*, 13(1):1–33, 2010.

[32] Wu Y. and Liu Y. Variable selection in quantile regression. *Statistica Sinica*, 19(2):801–817, 2009.