Supply Cushion Methodology and Detection of Events of Interest

Directions Paper

June 4, 2012

Market Surveillance Administrator | 403.705.3181 | #500, 400 – 5th Avenue S.W., Calgary AB T2P 0L6 | www.albertamsa.ca

The Market Surveillance Administrator is an independent enforcement agency that protects and promotes the fair, efficient and openly competitive operation of Alberta's wholesale electricity markets and its retail electricity and natural gas markets. The MSA also works to ensure that market participants comply with the Alberta Reliability Standards and the Independent System Operator's rules.

Table of Contents

List of Tables and Figures

Executive Summary:

Our Motivation

One of the chief duties of the Market Surveillance Administrator is to monitor market outcomes to ensure that the conduct of participants supports the fair, efficient, and openly competitive operation of the electricity market. This involves actively monitoring the market in real time to ensure that observed outcomes are broadly consistent with market fundamentals and to identify potential anticompetitive conduct. Supply cushion analysis provides a reliable method of relating market fundamentals to market outcomes to highlight abnormal results; the MSA wishes to ensure that precisely how it goes about doing this is both transparent and robust.

Prior to 2010, the MSA relied on manual review of each hour of the market. This process was both labour intensive and reliant upon the judgment of the staff member conducting the review. As a consequence, in the Q3 report the MSA's third quarter report (Q3 2010) introduced 'Supply Cushion' analysis as a computationals method to flag hours for further review.

The supply cushion is fundamentally the difference between the energy that Alberta generation is capable of producing at a given moment in time, and what is actually demanded. If the supply cushion is small demand is close to maximum supply and we would expect a high pool price to prevail; the opposite when the supply cushion is large. Because the supply cushion is a good measure of market fundamentals, when the pool price does not reflect the supply cushion the outcome can be thought of as 'abnormal' and is subject to further scrutiny. The supply cushion is nothing more than a monitoring metric used to efficiently identify market outcomes for further review. The report identifies other fundamentals (such as, changing fuel costs and the composition and context of a given supply cushion) to be taken into consideration in this further review.

There are many different ways to calculate the relationship between pool price and the supply cushion. The method adopted in the Q3 2010 report, and applied ever since, was to separate observations of supply cushion/pool price pairs into discrete bands, and then identify abnormal hours using the calculated estimates of the mean and standard deviation within the band. The MSA selected a 250 MW band size to ensure that each band contained enough information to produce robust estimates yet included only datapoints close enough to be relevant. After a logarithmic transformation (to normalize the data), within each band over 99% of the observations are within 3 standard deviations of the mean. As a result, any that are not, are unusual enough to be flagged as events of interest for further review. The MSA chose to use a consistent dataset of the period between February 2008 and June 2010 (at the time all reliable data) to calculate the estimates, against which future hours would be considered. The MSA found that the 250 MW band method provided the best balance of simplicity, transparency, and accuracy.

In the Q3 2010 report the MSA identified a number of limitations and possible refinements. In late 2011 the MSA began further investigation as to whether these possibilities constituted a genuine improvement. At about the same time, some stakeholders raised various concerns about the methodology. As a result of this interest, the MSA formed a technical working group with representatives of some stakeholders in March of 2012 and announced that work on the supply cushion analysis would form part of the state of the market report.

What we analyzed

Internal MSA analysis and stakeholder concerns focused around two main issues: how the supply cushion estimates are calculated, and what data is used to do so. This report considers a number of alternate specifications for band size, and several methods that avoid bands altogether to produce continuous 'normal' intervals. A number of alternate datasets are compared with the one currently in use by the MSA.

Secondary issues addressed include the distribution of the data and remedial transformations, augmenting the analysis with other market fundamentals, and other questions of precise specification.

In particular, the MSA thanks the technical working group for its input on the questions of band size selection, the appropriate dataset, and alternate distributional assumptions. Attached to this report is a complete dataset which will allow interested market participants to confirm the results of the work performed.

What we concluded

The MSA initially chose the 250 MW band method because it was simple, transparent, and accurate. Having reviewed the alternatives, both the alternate band specifications and considerably more complex continuous techniques, the MSA finds that they produce materially similar results to the current method. What differences there are cannot be conclusively qualified as either an improvement or a detriment. The MSA has not been persuaded of any significant benefit from the alternatives, and as such sees no reason to replace the current method. The MSA will, from time to time, run these other techniques to ensure that the results remain broadly representative.

The February 2008 to June 2010 dataset was chosen almost by necessity when the supply cushion analysis first appeared in the quarterly reports. At the time it represented the entire universe of reliable, available data. At present it remains a desirable choice as it dates from a period where the MSA is reasonably certain there was no systematic untoward application of market power and is therefore appropriate for the current analysis. If and when the MSA is convinced the same applies to subsequent periods, they may be considered for inclusion in the benchmark dataset. Depending on the conclusions elsewhere in the state of the market report, 2011 would be an excellent candidate for instance. Like the calculation methodology itself, the MSA has not been shown any persuasive reason to deviate from the current procedure.

The MSA is releasing a comprehensive dataset along with this report, which will allow market participants who wish to confirm our results. This report and our work with the technical working group reflects the MSA's continuing desire to be open, accessible and willing to debate and discuss issues that arise as we carry out our responsibilities.

1. Introduction

Monitoring market outcomes is an important part of the MSA's work. Through monitoring the MSA tracks the conduct of market participants, the ISO and the structure and performance of the electricity market. Historically, MSA staff had monitored the market using a variety of tools but there remained a manual and subjective element to finding events of interest. Consequently, there was a desire to find a less labour intensive and more consistent method to assist the MSA in detecting such events.

In November 2010 (Q3 2010 MSA Report) the MSA presented a new metric to assist in identifying events for further scrutiny ("events of interest"). This new metric examined the relationship between market tightness or "Supply Cushion" and pool price. The metric is based on the idea that in a well-functioning energy-only market price should generally be reflective of market tightness. That is, one would expect the larger the supply cushion the lower the pool price, or vice versa. Since November 2010 the supply cushion metric has featured regularly in the MSA's quarterly reports and has proved useful in focusing the MSA's monitoring activities.

In late 2011 the MSA commenced some internal work testing the robustness of the supply cushion metric and investigating possible refinements. In Q1 2012, a number of stakeholders expressed an interest in understanding how the metric was calculated and suggested some alternatives. In response the MSA released a dataset containing supply cushion estimates and invited stakeholders to participate in a technical working group. The MSA's work and some of the output generated by the technical working group are included in this report. The MSA appreciates stakeholders' assistance in these matters.

The report is structured as follows:

> Section 2 reviews the original supply cushion analysis presented in the MSA's Q3 2010 report.
>
> Section 3 considers a number of possible refinements to the supply cushion methodology
>
> Section 4 provides the MSA's conclusions.

In addition to the report itself we are releasing a data set containing the results to allow stakeholders to compare different approaches.

2. Supply Cushion Measurement and Detecting Events of Interest

In any given hour pool price will be the result of a number of factors related to market fundamentals, the offer behavior of market participants, and actions taken by the AESO to maintain reliability. The MSA believes that in a well-functioning energy only market price should generally be reflective of market tightness. A market that repeatedly yields low prices during scarcity and high prices during surplus is unlikely to be providing the correct incentive for investment. Thus, it should be expected that there is a relationship between hourly pool price and market tightness, such that pool prices tend to be higher for periods of scarcity. The MSA's supply cushion is a direct measure of hourly market tightness.

This section is based on the MSA's original formulation presented in the Q3 2010 report.

2.1 SUPPLY CUSHION MEASUREMENT

The supply cushion measures the un-dispatched energy offers in the merit order, more formally:

$$\text{Supply Cushion} = \sum_{i=1}^{n}(\text{Available MW} - \text{Dispatched MW}) + \text{DDS dispatched} - \text{TMR dispatched}$$

where n is the number of offer blocks in the merit order.

The supply cushion measure includes an adjustment for MW providing Dispatch Down Service (DDS). Offers corresponding to units currently providing DDS are not included in the merit order. However, these MW remain available and are returned to the merit order in certain circumstances (for example, price rises to above the *reference price* specified under ISO Rule 3.10).[1]

The supply cushion measure also includes an adjustment for Transmission Must Run (TMR) to avoid counting MW that are already dispatched for TMR. More specifically, units currently dispatched for TMR also have corresponding offers in the energy merit order that indicate available MW and no energy dispatch. Should price rise such that the energy offer is in merit, the unit is dispatched off for TMR and dispatched on for energy but does not result in additional generation occurring.

[1] The merit order snapshots presented on the AESO website includes an 'offer block' labelled 'TMR' this actually corresponds to the MW of DDS that could be returned to the merit order. Including this offer block in the calculation of the supply cushion effectively adds back these available MW.

The supply cushion metric that has been used by the MSA is estimated in each hour by using a snapshot of the merit order at approximately 30 minutes into the hour. The same snapshot data (for energy, DDS and ancillary services merit orders) are made public on the AESO's website 60 days after the offers are made to the power pool. Interested stakeholders can use this publicly available data to estimate the supply cushion. As an example Table 2.1 shows the components of the supply cushion for August 1, 2010 HE1. On March 6, 2012 the MSA also posted supply cushion estimates from February 1, 2008 to December 31, 2011 to its website.[2]

Table 2.1: Supply Cushion Calculation for August 1, 2010 HE1

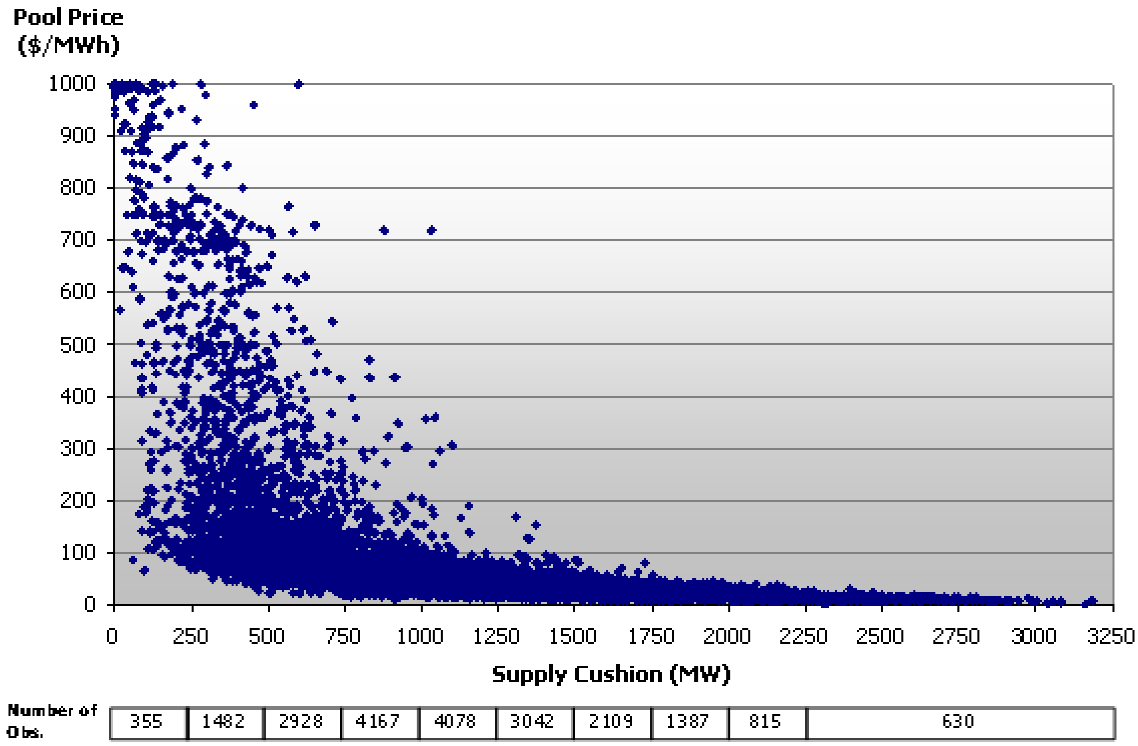
Supply Cushion	MW
$\sum_{i=1}^{n}(Available\ MW - Dispatched\ MW)$	1341
$+DDS\ dispatched$	76
$-TMR\ dispatched$	76
Total	1341

2.2 POOL PRICE AND SUPPLY CUSHION RELATIONSHIP

Figure 2.1 shows the historical relationship between supply cushion and pool price using hourly data from February 1, 2008 to June 30, 2010. Prior to the implementation of Quick Hits (December 3, 2007) equivalent data is not available to construct an estimate of the supply cushion. Further, data issues prevent construction of reliable metrics for much of the period between December 3, 2007 and February 2008. Overall, the MSA has estimates of the supply cushion in 20,993 of 21,144 hours (99.3 %) for the period of February 2008 to the end of June 2010. Of the 151 hours that were 'missed', most were due to technical issues like the dispatch tool being on maintenance and energy alert situations. Figure 2.1 also shows the number of observations within each supply cushion band, i.e. =<250 MW, >250 to =<500MW etc. Supply cushions >2250MW are relatively rare and have been grouped into a single band.

[2] See References for a link to this data.

Figure 2.1: Supply Cushion vs. Pool Price



2.3 IDENTIFICATION OF EVENTS OF INTEREST

In preparation of the Q3 2010 report the MSA considered a number of methodologies for identifying events of interest. Ultimately a very simple technique was chosen; group the observations into bands, and within each identify abnormal events based on how many standard deviations away from the mean.

From examination of the scatter plot of Figure 2.1, there is an observable non-linear relationship between supply cushion and price. Further, the data does not appear to be normally distributed (for example, there are more values above the mean than below the mean). By taking the logarithm of the price data the distribution around the mean is approximately normal, such that outcomes within two standard deviations of a mean represent something close to a 95% confidence interval.[3] Figure 2.2 shows the supply cushion plotted against price using a logarithmic scale, along with the mean and +/- 1, 2 and 3 standard deviations. Figure 2.3 also shows these same mean and standard deviation bands plotted against price – note that the construction of the standard deviation bands around the logarithm of price results in the bands being wider above

[3] The pool price data contains some instances of $0 prices. In order to obtain a logarithm of pool price, $1 is added to each price.

the mean than below. The count of observations falling within each band in shown in Table 2.2.

Figure 2.2: Supply Cushion vs. Logarithm of Pool Price, including Standard Deviation bands

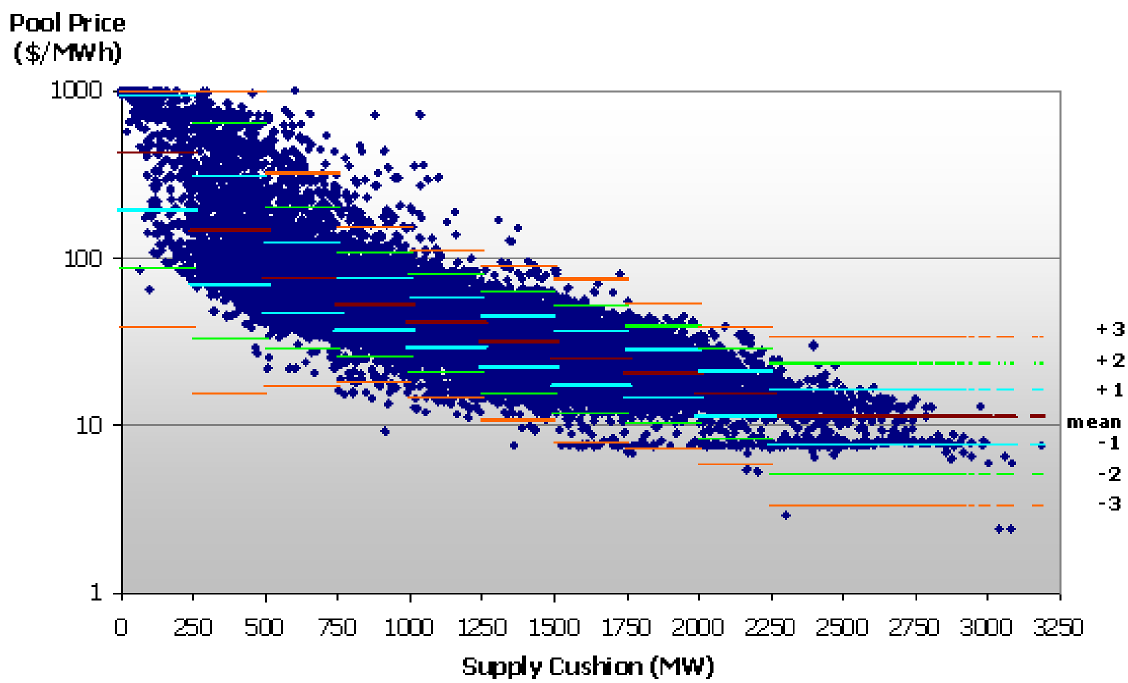


Figure 2.3: Supply Cushion vs. Pool Price, including Standard Deviation bands based on the Logarithm of Pool Price

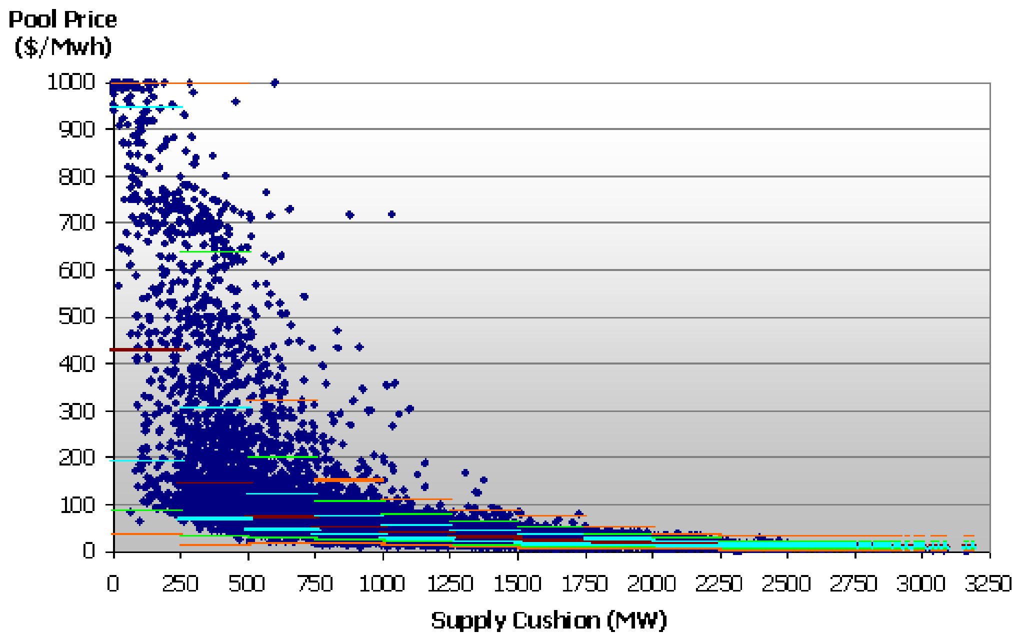


Table 2.2: Count of Hours in Each Standard Deviation Band

	=<250	>250	>500	>750	>1000	>1250	>1500	>1750	>2000	>2250	Total
>=+3			62	44	17	8	4	1	1		**137**
<+3 & >=2		86	89	60	47	29	17	25	20	3	**376**
<2 & >=1	51	196	160	453	451	383	307	222	106	36	**2365**
<1 & >=mean	168	294	950	1458	1676	1202	747	436	232	347	**7510**
<mean & >=-1	58	749	1321	1539	1340	987	661	478	377	176	**7686**
<-1 & >=-2	65	155	330	543	413	329	330	204	31	60	**2460**
<-2 & >=-3	13	2	15	55	110	100	25	20	46		**386**
<-3			1	15	24	4	18	1	2	8	**73**
Total	**355**	**1482**	**2928**	**4167**	**4078**	**3042**	**2109**	**1387**	**815**	**630**	**20993**

3. Possible Refinements and Alternative Approaches

The MSA's quarterly reports have noted a number of limitations with the supply cushion estimation and the methodology for selecting events of interest. The Supply Cushion Technical Working group also identified a number of alternate approaches. The MSA's approach has been to consider refinements and alternatives and to present the results of each so that both the MSA and stakeholders may come to an informed opinion. The refinements and alternatives fall into a number of categories, each of which is examined in a subsection:

- Section 3.1 - Band size selection
- Section 3.2 - Alternative distribution assumptions
- Section 3.3 - Benchmark data set selection
- Section 3.4 - Additional fundamentals
- Section 3.5 – More discrete data

3.1 BAND SIZE SELECTION

In the analysis described in section 2 the MSA selected a band size of 250 MW (along with a single large band to capture observations above 2250 MW). The band size was selected to satisfy two competing goals. First, to ensure enough observations are within each band to get robust estimates of the mean and standard deviation. Second, to have enough bands so as to represent the negative relationship between pool price and supply cushion. Obviously a number of other ways of selecting band size would also satisfy these requirements.

In section 3.1.1 we consider simple alternatives to a 250 MW band size selection. We note that all fixed band size methodologies suffer from the same disadvantage – that they are unlikely to be good predictors of events of interest for supply cushion observations close to the edge of a band. We then provide a sensitivity analysis to show how outliers change for an alternative band size.

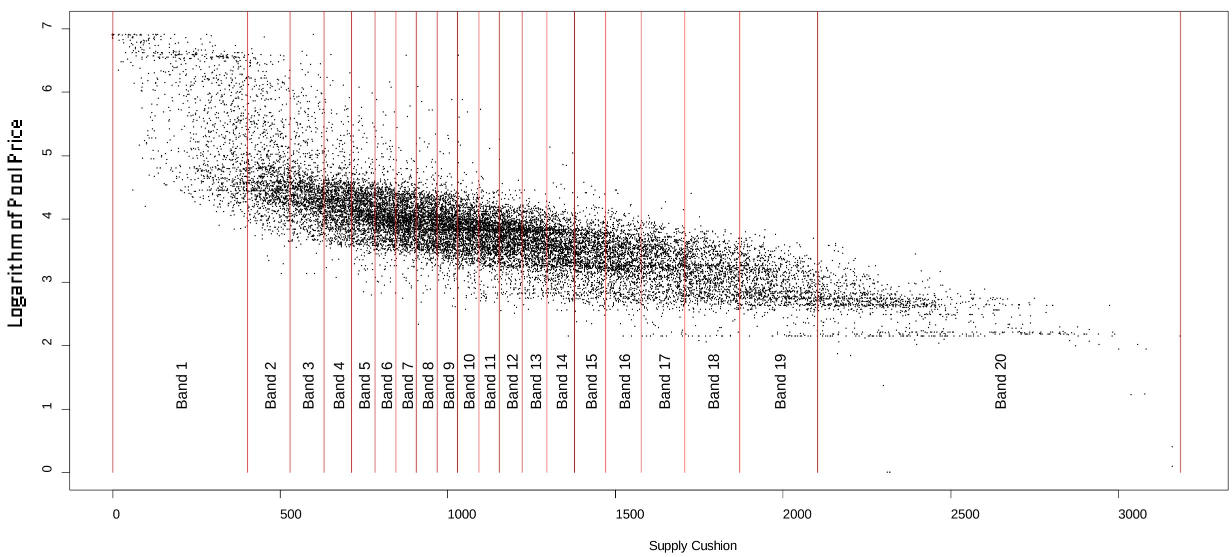
In section 3.1.2 we consider two methodologies that avoid the specification of a band size. Firstly, a parametric approach based on an ordinary least squares regression and secondly a non-parametric approach. The non-parametric approach applies a weighting function to other observations in order to construct an estimate of the relationship between pool price and supply cushion. We provide a comparison of results between the parametric, non-parametric and 250 MW band size methodologies. Finally, we investigate whether estimation methods might be subject to bias as a result of the price cap and floor.

3.1.1. Alternative band size selection

3.1.1.1. Band Size Based on Number of Observations

Instead of defining a band by the number of MW in the supply cushion it spans, one can define it so it includes a fixed number of observations. In Figure 3.1 we show the bands that would result assuming 1050 observations in each band. [5]

Figure 3.1: Bands based on 1050 observations per band (Dataset: Feb 2008 to Jun 2010 inclusive).



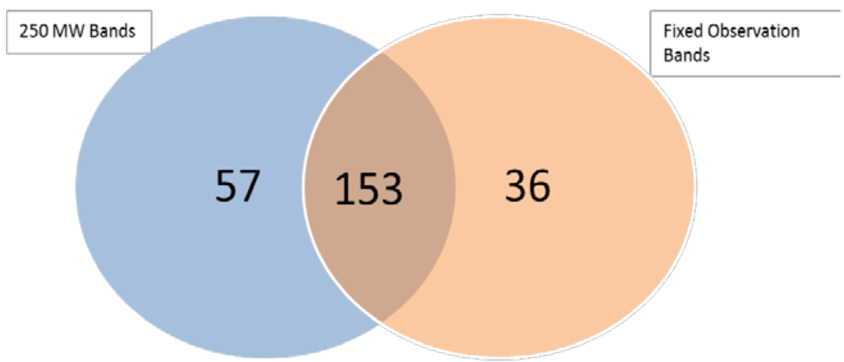
The band size based on observations can be designed to satisfy our first requirement (ensuring enough observations within each band to get robust estimates of the mean and standard deviation) but results in quite large bands where there are relatively few observations. Since all values within a band contribute equally to the determination of a mean, combining points from widely differing supply cushions may fail to provide a good approximation of relationship between pool price and supply cushion. For example, in Figure 3.1 above the first band ranges from a supply cushion level of 0 MW to 403 MW. This method ignores that we would expect a very different price between those two market conditions. Further, a wide band may inflate the estimate of the standard deviation which in turn causes fewer hours to be flagged for review when the market is tight.

[5] Selecting 1050 observations in each band creates 19 equal sized bands with the twentieth band containing 1043 observations (for a total of 20,993 observations).

Basing band sizes on observations also has a disadvantage in the case where the benchmark data is updated (see discussion in Section 3.3). In this case the starting and end points for bands may change (for example, one benchmark data set may have 1050 observations between 1000 MW and 1200 MW supply cushion whereas another may have 1050 observations between 1000 MW and 1100 MW). This would make it more difficult to compare results to determine if there was a trend in a particular band.

In terms of detecting events of interest the 1050 observation approach detects many of the same events as the 250 MW band approach described in Section 2. The figure below shows this result in the form of a Venn diagram – the overlapping area shows the number of outliers common to both approaches.

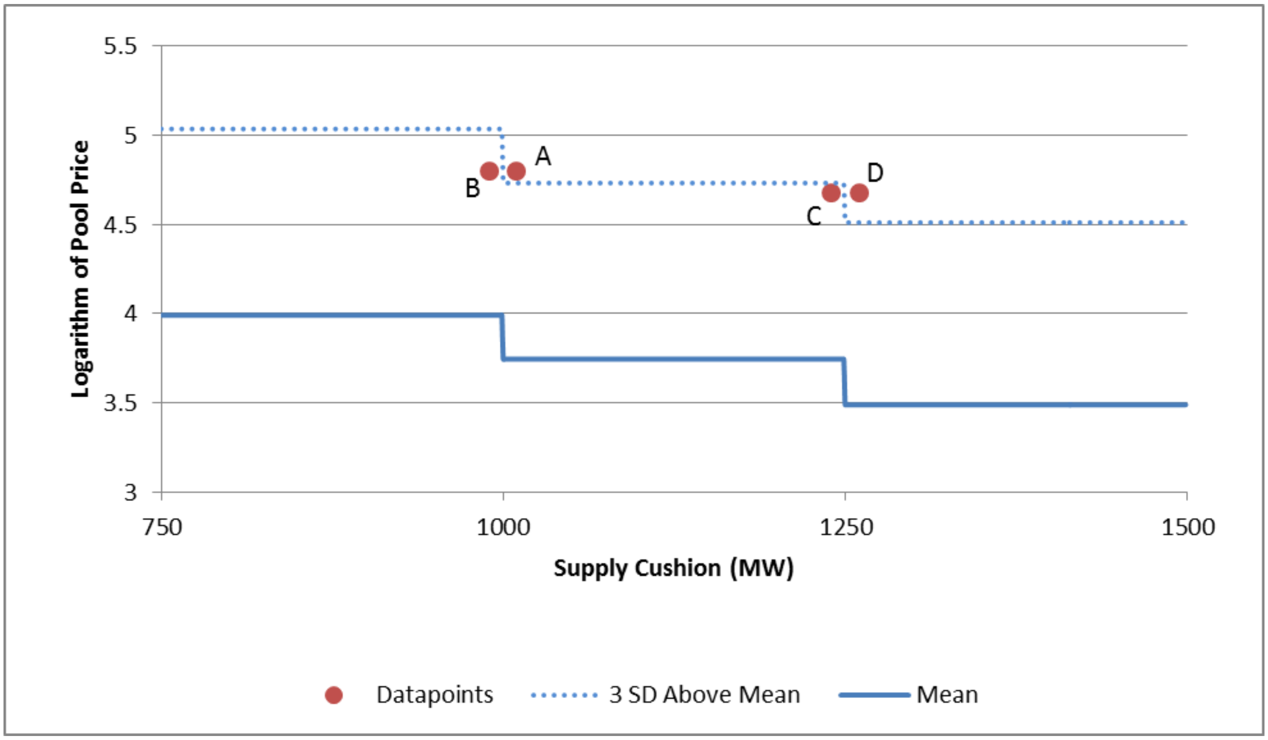
Figure 3.2: Venn diagram showing intersection between +/-3 standard deviation outliers using a 250 MW band size and a fixed 1050 observations per band (Dataset: Feb 2008 to Jun 2010 inclusive)



3.1.1.2. Problems with Band Size Selection

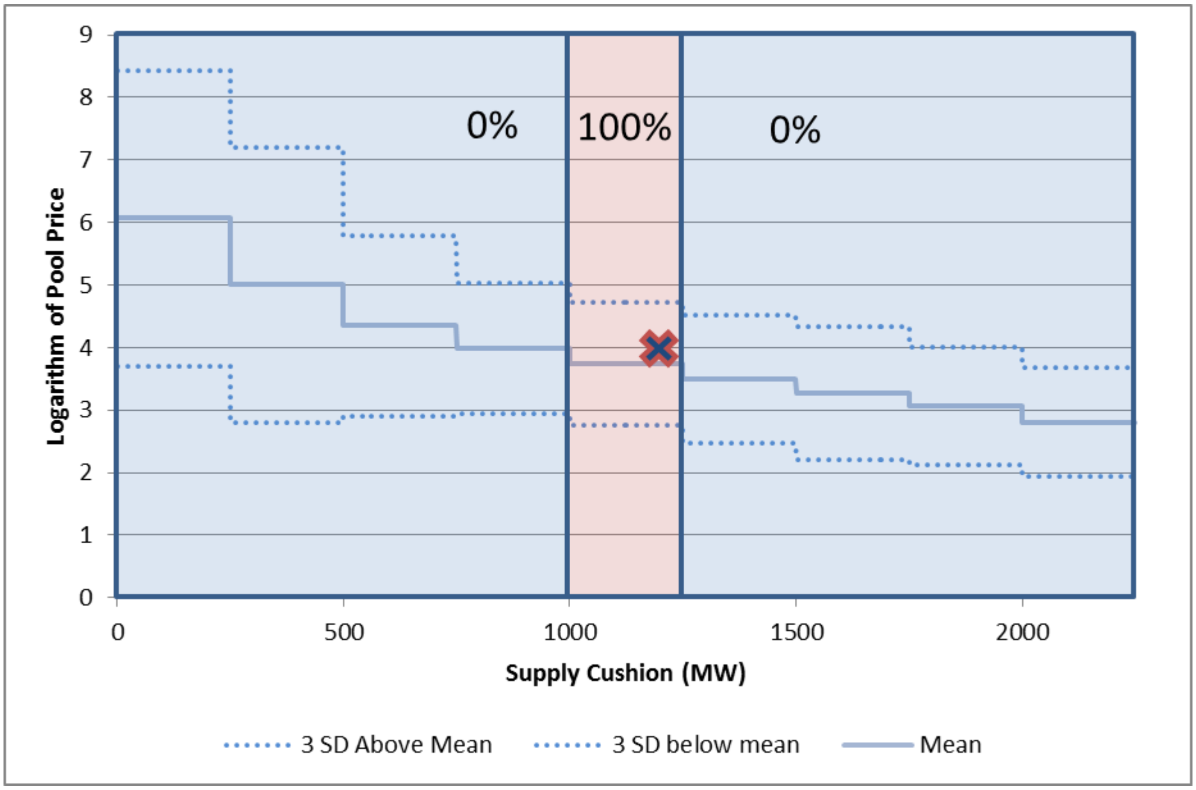
All fixed band size methodologies suffer from the same disadvantage, namely they are unlikely to be good predictors of events of interest for supply cushion observations close to the edge of a band. Figure 3.3 shows this problem graphically. Point A with a supply cushion of 1001 MW is in the band from 1000 MW to 1250 and is flagged as an event of interest. Point B with a supply cushion of 999 MW and the same pool price is in the band 750 MW to 1000 MW, but in this band is not flagged as an outlier. An analogous issue can be seen at the lower end of the band (Points C and D). We consider alternatives that avoid selection of a specific band in section 3.1.2.

Figure 3.3: Illustration of Problem with Band Size Selection



By grouping observations into different bands before calculating the price distribution, a band-based method ignores the information embedded in the observations that are outside of the band, regardless of how salient it may be. Conversely, it weights all observations within a band equally, regardless of appropriateness. This is illustrated in Figure 3.4.

Figure 3.4: Band Selection and Weight



To illustrate the problem with band size selection further we compare the results shown in Section 2 (based on a 250 MW band size) with the results based on a 300 MW band size.[6] The number of observations outside three standard deviations from the mean is similar between both approaches (see the Venn diagram shown in Figure 3.5) but the actual observations flagged are different. This difference is shown more clearly in Figure 3.6.

[6] The first band contains supply cushion levels from 0 up to 300 MW, the second 300 to 600MW and so on until 2100 MW where all greater supply cushions are grouped into one band

Figure 3.5: Venn Diagram Showing Intersection Between +-3 Standard Deviation Outliers Using a 250 MW Band Size and a 300 MW Band Size (Dataset: Feb 2008 to Jun 2010 Inclusive)

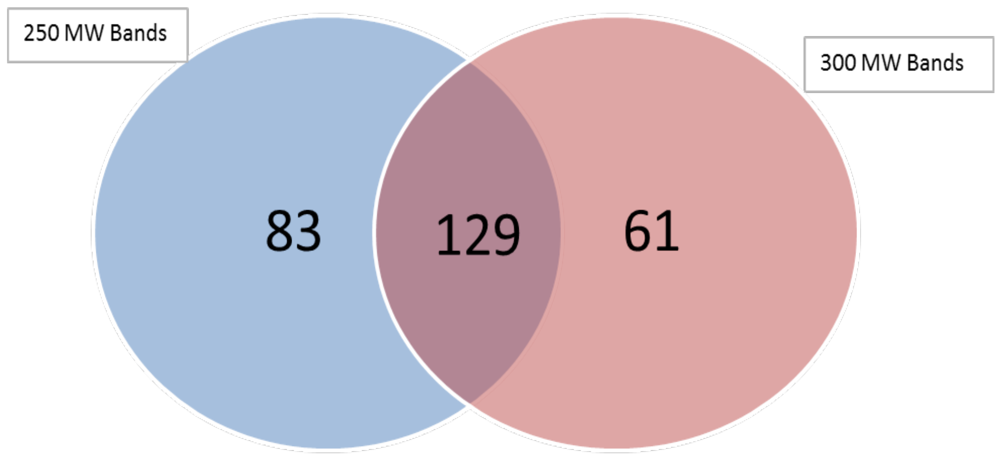
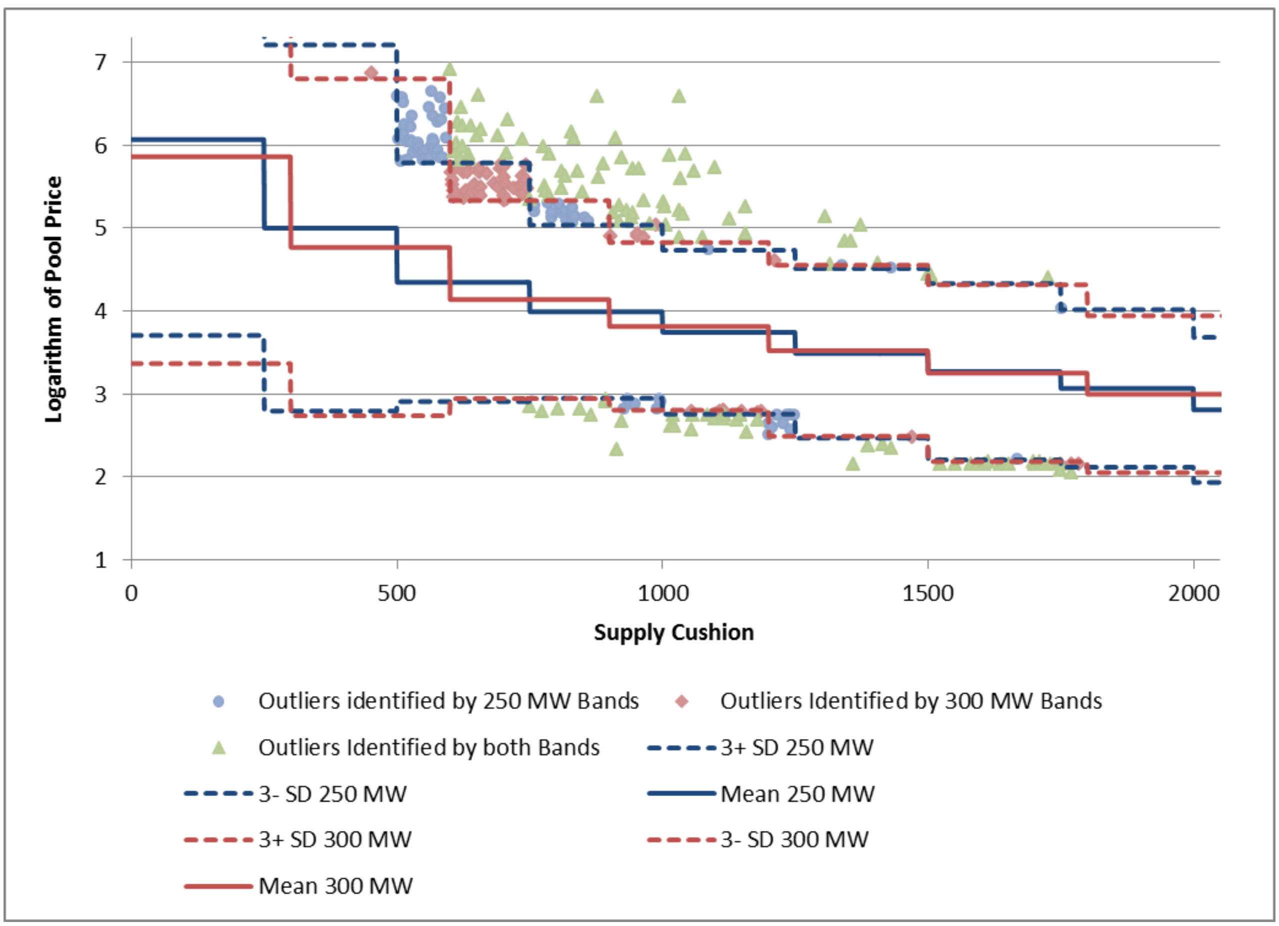


Figure 3.6: Comparison of +-3 Standard Deviation Outliers Between 250 MW Band Size and 300 MW Band Size (Dataset: Feb 2008 to Jun 2010 Inclusive)



3.1.2. Methods that Avoid the Specification of a Band Size

In this section we consider two methodologies that avoid the specification of a band size. Firstly, we present a parametric approach based on an ordinary least squares regression and secondly a non-parametric approach.

3.1.2.1. Ordinary Least Squares (OLS) Projection Method

The OLS method projects the expected value of the pool price conditional on the supply cushion, $E(y\,|\,x)$, by estimating a linear functional form. That is

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varsigma$$

where y is the natural logarithm of pool price, x is the supply cushion, and ς is a random term and is orthogonal to x (uncorrelated with x). The quadratic term, x^2, is included to capture the non-linear curvature of the relationship between pool price and supply cushion. Additional terms could be included to capture higher order non-linearity (e.g. x^3).

To determine a range of 'normal' pool prices for a given supply cushion we need to first estimate the variance of the pool price conditional on the supply cushion, $Var(y\,|\,x)$, which is not the same as the variance of the expected pool price conditional on supply cushion, $Var(E(y\,|\,x))$. The latter is used to derive a confidence interval that will give us a range that we expect will contain the mean of the data at that point. (This is what the OLS regression predicts.) In contrast, what we want is a *prediction* interval which gives us the range where we would expect to find another observation for that supply cushion level. Recall that the mean is only the center of the data, and as such we would expect observations to be distributed equally around it, both above and below. If the true mean is very close to the edge of the confidence interval (as we cannot rule out), it would not be surprising to find that a significant portion of the distribution around that mean falls outside. Hence, when trying to predict where another observation would lie for that supply cushion we must give a larger range than the confidence interval (where we expect the mean to lie). Having acknowledged that the mean may fall anywhere in the confidence interval, we must then acknowledge the distribution of the data around all of the possible means. In other words, the prediction interval takes into account uncertainty about both where the data falls around the mean and uncertainty regarding the location of the mean itself. More formally,

$$Var(y\,|\,x) = Var(y - E(y\,|\,x) + E(y\,|\,x)\,|\,x)$$

$$= Var(\varsigma\,|\,x) + Var(E(y\,|\,x))$$

Where the first term represents variance around a mean, and the second the variance of the mean itself. We can then construct a range of +/- 3 standard deviations around the prediction for every supply cushion level.

Confidence Intervals and Prediction Intervals

Readers who have studied some statistics are likely familiar with the concept of confidence intervals but less so with prediction intervals. The following is a brief description to distinguish between the two:

Confidence intervals give an indication of how well you have determined the mean. A 95% confidence interval of the mean tells you that if you randomly took many new samples of data from the population and calculated the confidence interval for each sample, you would expect 95% of those intervals to include the true value of the population mean.

Prediction intervals give an indication of where you can expect to see the next data point sampled. A 95% prediction interval tells you that if you randomly took many new samples of data from the population you would expect each value to lie within the prediction interval 95% of the time.

In summary: The confidence interval tells you about the likely location of the true population mean. The prediction interval tells you about the distribution of values, not just the uncertainty in determining the population mean.

The main disadvantage with an OLS method is that while the observation of the data suggests a nonlinear relationship between pool price and supply cushion it does not suggest a specific functional form. While we may believe that the quadratic form is a good *approximation* of the relationship between the supply cushion and the pool price, we cannot know if it is the *right* one. The non-parametric method discussed in the next section avoids the need to specify the functional form of the relationship between pool price and supply cushion.

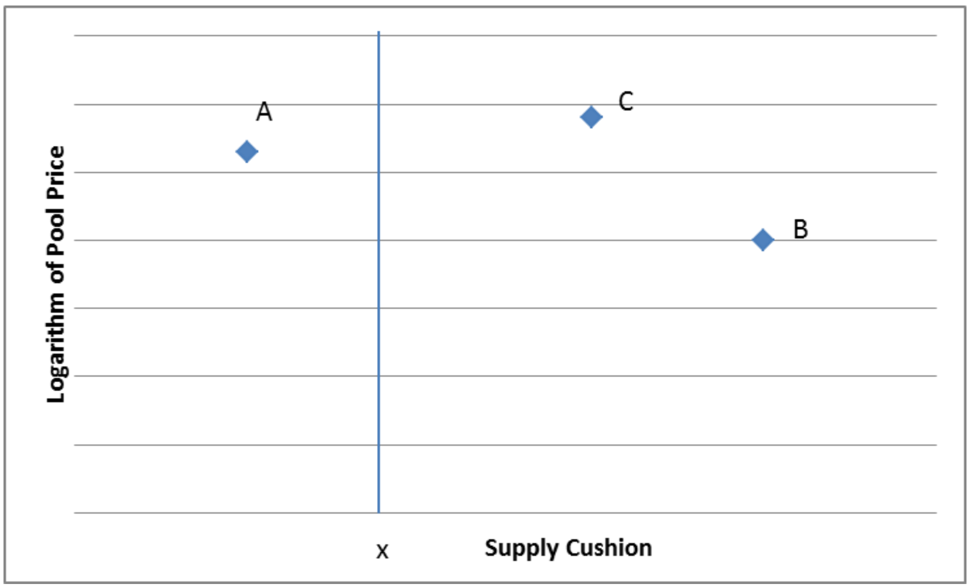
3.1.2.2. Non parametric Method[7]

In Figure 3.4 we noted that any method based on band size implies an equal weight to all observations within a band and no weight to observations outside of it. This issue can be addressed directly by weighting observations according to the distance away from a given supply cushion level, i.e. the further away from a given level the less

[7] For an introductory to nonparametric econometrics, see J. Racine (2008). Q. Li and J. Racine (2007) offers a comprehensive discussion.

weight is assigned to the observation. We show an illustration of this in Figure 3.7. The vertical line at X represents a supply cushion level for which we are interested in assessing the likely pool price. Consider the three observations of the pool price and supply cushion relationship labeled A, B and C. Since A and C are closer to the vertical line than B they should include more useful information about the price for a supply cushion of X and thus should be given more weight than B. Further, as A and C are the same distance from the vertical line it is reasonable to assume that they should be accorded the same weight in determining the price at X.

Figure 3.7: Illustration of Weights Applied to Observations in Determination of Pool Price at X.



More specifically, let's assume

$$y = m(x) + \varepsilon$$

where y is the pool price in logarithm, x is the supply cushion, $m(x)$ is some unknown function, and ε is random term and is orthogonal to x. Note the function form of $m(x)$ is not specified.

Further assume the density function of x is $f(x)$, and the joint density function of (y, x) is $g(y, x)$. We are interested in the expected value of y conditional on x, which is $E(y \mid x) = E(m(x) + \varepsilon \mid x) = m(x)$. It can be estimated by

$$m(x) = \int \frac{yg(y,x)dy}{f(x)} = \frac{\sum_{i=1}^{n} y_i K(\frac{x_i - x}{h_x})}{\sum_{i=1}^{n} K(\frac{x_i - x}{h_x})}$$

where $K(\bullet)$ is the Kernel function[8] and h_x is the bandwidth, a smoothing parameter. For our estimate we use the second-order Gauss Kernel[9], namely $K(x, x_i, h_x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\frac{x_i - x}{h_x})^2\right)$. The least-squares cross-validation is used to select the optimal bandwidth.[10]

The variance of $m(x)$ is given by

$$\frac{\sigma^2(x)}{f(x)nh} \int K^2(t)dt$$

where $\sigma^2(x)$ is the conditional variance of y.

The last step is to derive the variance used to construct the prediction interval.

$$Var(y \mid x - E(y \mid x)) = Var(y \mid x) + Var(E(y \mid x))$$

$$= \sigma^2(x) + Var(m(x))$$

$$= \sigma^2(x)\left(1 + \frac{1}{f(x)nh} \int K^2(t)dt\right)$$

As with the other methods we then determine outliers based on +- 3 standard deviations from the mean.

[8] A Kernel function is just a weighting function.

[8] Other Kernel functions could be used, e.g., Epanechnikov or Quartic. The Gauss Kernel is selected so as to include more information on the relationship between pool and supply cushion.

[10] See Li and Racine *(2007)* and Hardle, *(1990)* for a thorough discussion.

3.1.2.3. Comparison of Results[11]

Figure 3.8 shows a comparison of results between the OLS projection and non parametric estimation methodologies. Both cases show an interval based on 3 standard deviations from the mean. Note that:

- The non parametric method identifies more downside outliers, particularly at low supply cushion levels.
- The simple functional form specified for the OLS projection method results in an estimation of a smooth relationship between pool price and supply cushion. This flags more upside outliers at very low levels of supply cushion but less between 400 and 1000 MW.

A further comparison in Figure 3.9 shows the results for the non parametric method and the 250 MW band size method. Note that:

- The non parametric method identifies more downside outliers at low supply cushion levels and less downside outliers at high supply cushion levels.
- The non parametric method identifies more upside outliers below supply cushion levels of 750 MW. This is in part due to the band method not being able to identify any outliers below 500 MW (i.e. the 3 standard deviation level is higher than the price cap). Above a 750 MW supply cushion the non parametric method identifies a few less outliers.

[11] The non parametric method relies upon constructing a value based on neighboring observations. Close to the edge of the dataset this neighborhood becomes asymmetric as data points beyond the edge that would have been given non-trivial weight had they existed are not observed, potentially distorting the estimates. As a result, whenever non parametric results are reported, observations with supply cushions less than 100 MW or greater than 2300 MW are omitted (<u>including where results from other methods are compared to the non parametric estimates</u>).

Figure 3.8: Comparison of OLS Projection and Non Parametric Method ((Dataset: Feb 2008 to Jun 2010 Inclusive)

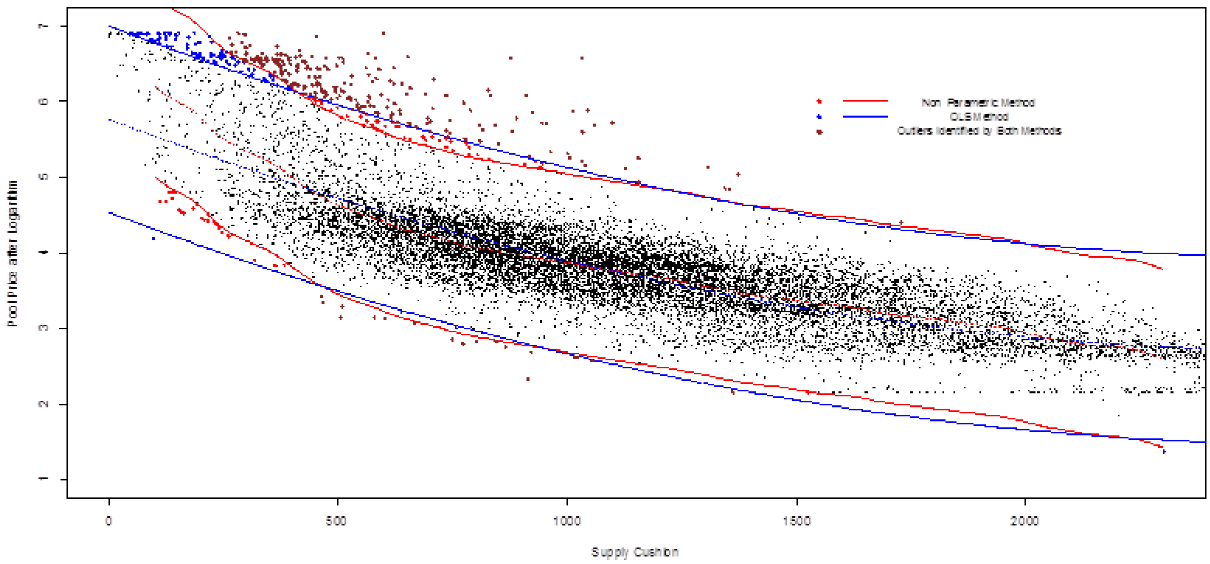


Figure 3.9: Comparison of Non Parametric and 250 MW Band Size Method (Dataset: Feb 2008 to Jun 2010 Inclusive)

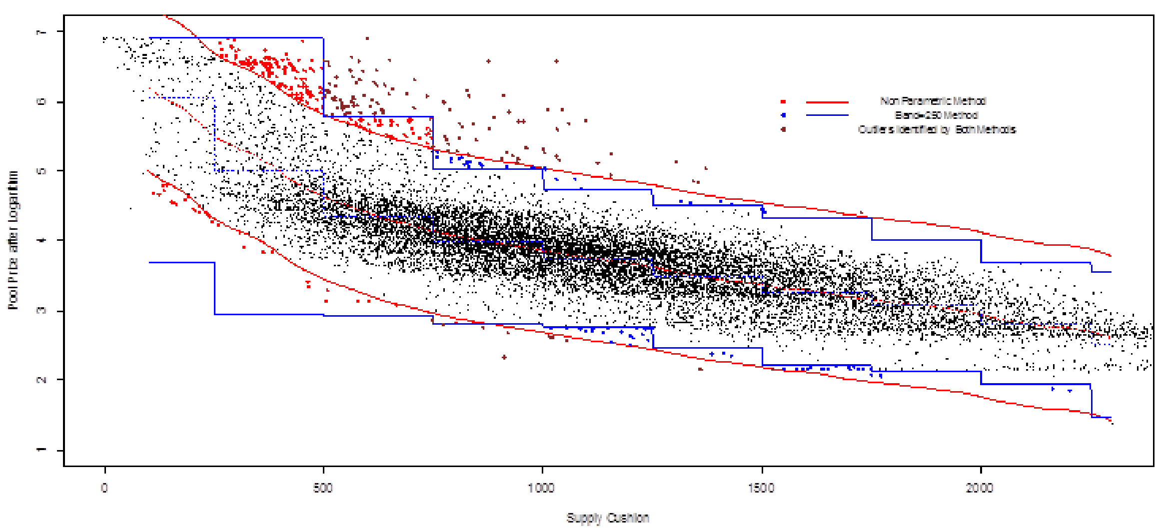
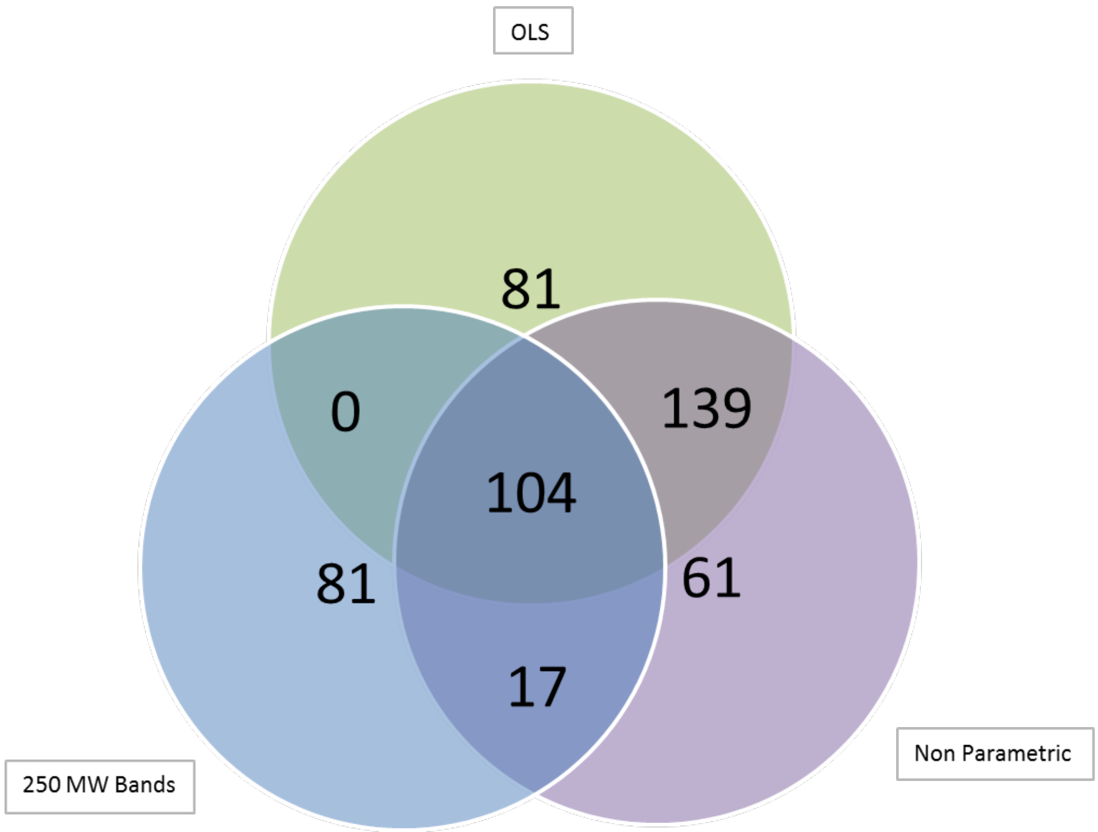


Figure 3.10: Venn Diagram Showing Intersection OLS Projection, Non Parametric and 250 MW Band Size Methods (Dataset: Feb 2008 to Jun 2010 Inclusive)



3.1.2.4. Impact of the Price Cap and Floor on Estimation

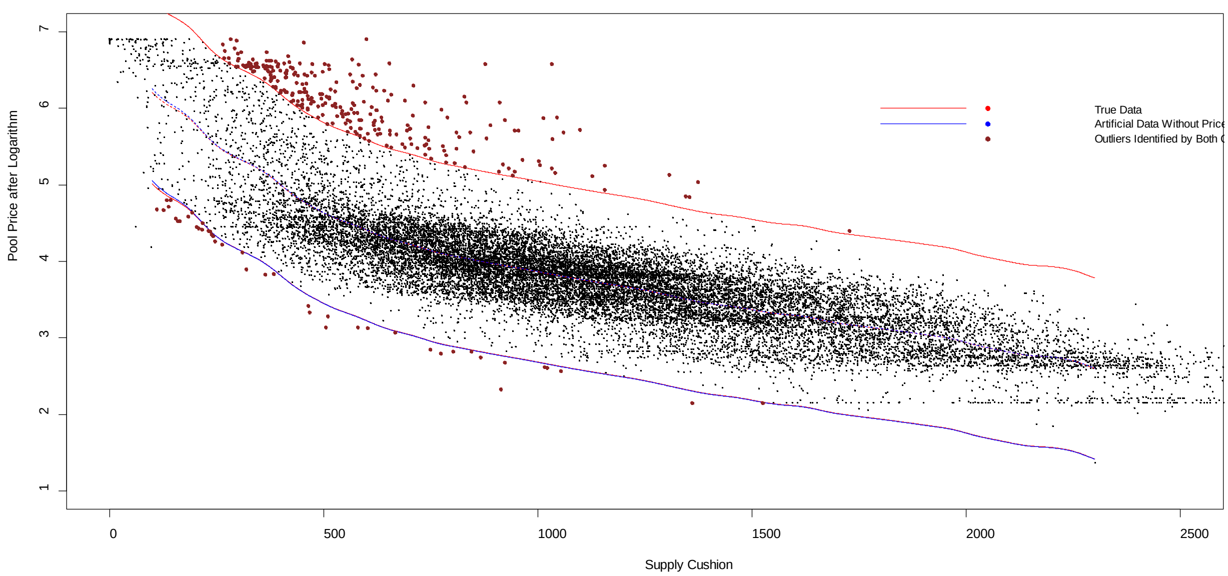
In this section we investigate whether estimation methods might be subject to bias as a result of the price cap and floor. Simply put, estimates of the expected value and variance of pool price at a given supply cushion level may be biased if price is in some way constrained. The price cap ($999.99) and floor ($0) are two obvious constraints. Without the price cap the realized pool price may have been greater, truncating our observations. A similar effect might be observed at the price floor.

It may well be the case that absent the constraints the greater variation in pool price would have produced a larger normal range. As a result, some hours end up flagged as events of interest that would not have been without the price cap and Floor. To assess the effect (if any) we consider a method for correcting for this possible source of bias. It must be stressed that this correction is purely illustrative, and has not been applied elsewhere in our analysis.

First, we assume that observations are only affected by the cap when the pool price comes very close to the maximum, specifically when pool price is greater than \$990. In these cases the observation is assigned a new price of \$999.99 +x, where x is drawn randomly from a normal distribution with mean of 1000 and standard deviation of 500 (i.e. on average it would assign a new value of \$1999.99).

In Figure 3.11 we compare the results from the non parametric method with and without an adjustment for the price cap. Since observations of pool price above \$990 occur at lower supply cushion levels, the impact is limited to the estimation of the mean and prediction intervals when the market is very tight. Overall, there is little impact on the number of hours flagged.

Figure 3.11: Comparison of Non Parametric Method With and Without Price Cap Adjustment (Dataset: Feb 2008 to Jun 2010 Inclusive)



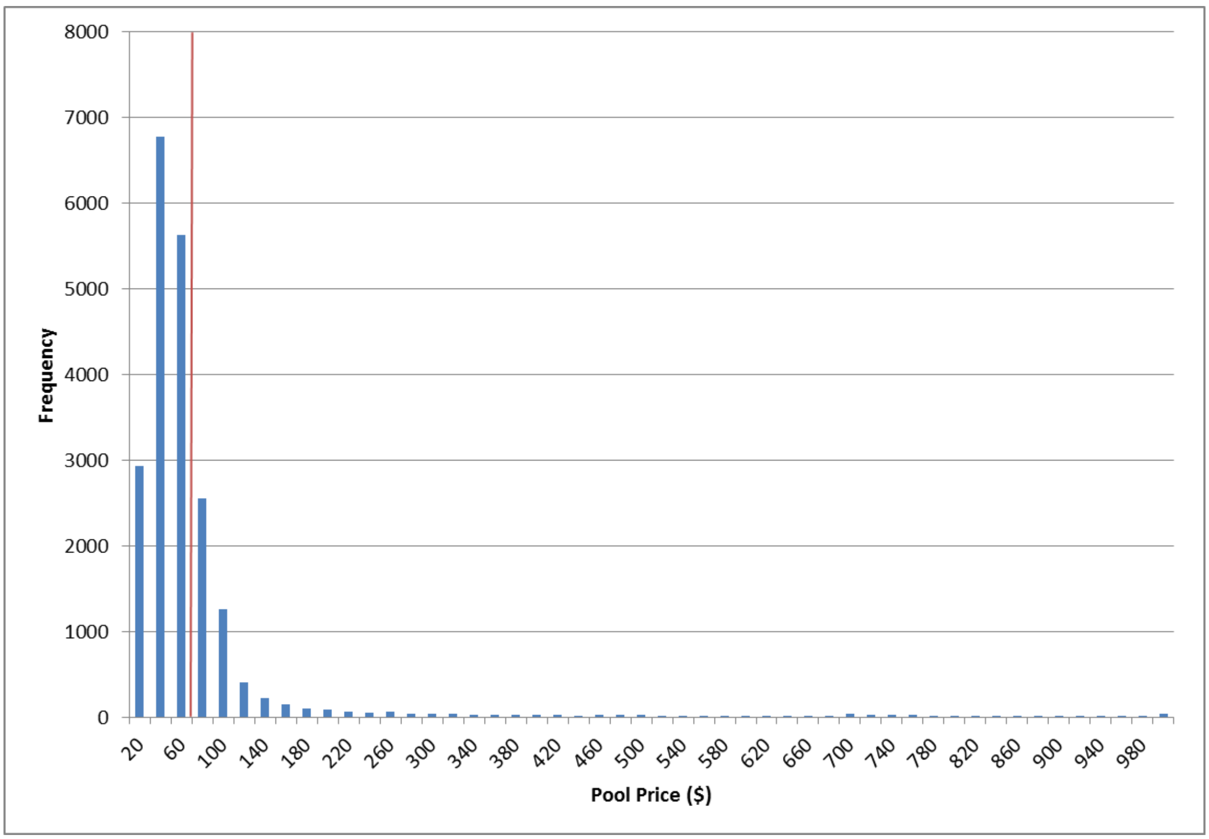
3.2 ALTERNATIVE DISTRIBUTION ASSUMPTIONS

In this report we have set out a number of methods for examining the relationship between pool price and supply cushion. All the methods considered have first transformed the pool price data by adding one to the pool price and then taking the natural logarithm (i.e. ln(pool price +1)).

Using a log transformation is a simple method for dealing with a number of common problems in data sets. A log transformation tends to squeeze together any values larger than 1, with the larger the value the more the data value is squeezed. Similarly, values less than 1 are stretched out.[12] The squeezing of the data can address a number of issues including:

> **Skewed data** – Skewness measures the symmetry of a distribution around the mean. A positive (or right) skew represents the situation where the bulk of the values are below the mean and the right tail of the distribution is long. Negative (or left skew) indicates the opposite. Pool price data has a positive skew. Figure 3.12 shows a histogram of pool price from our benchmark data set (Feb 2008 – Jun 2010) with a definite positive skew. Figure 3.13 shows the effect on the distribution of the benchmark dataset of the logarithmic transformation.[13]

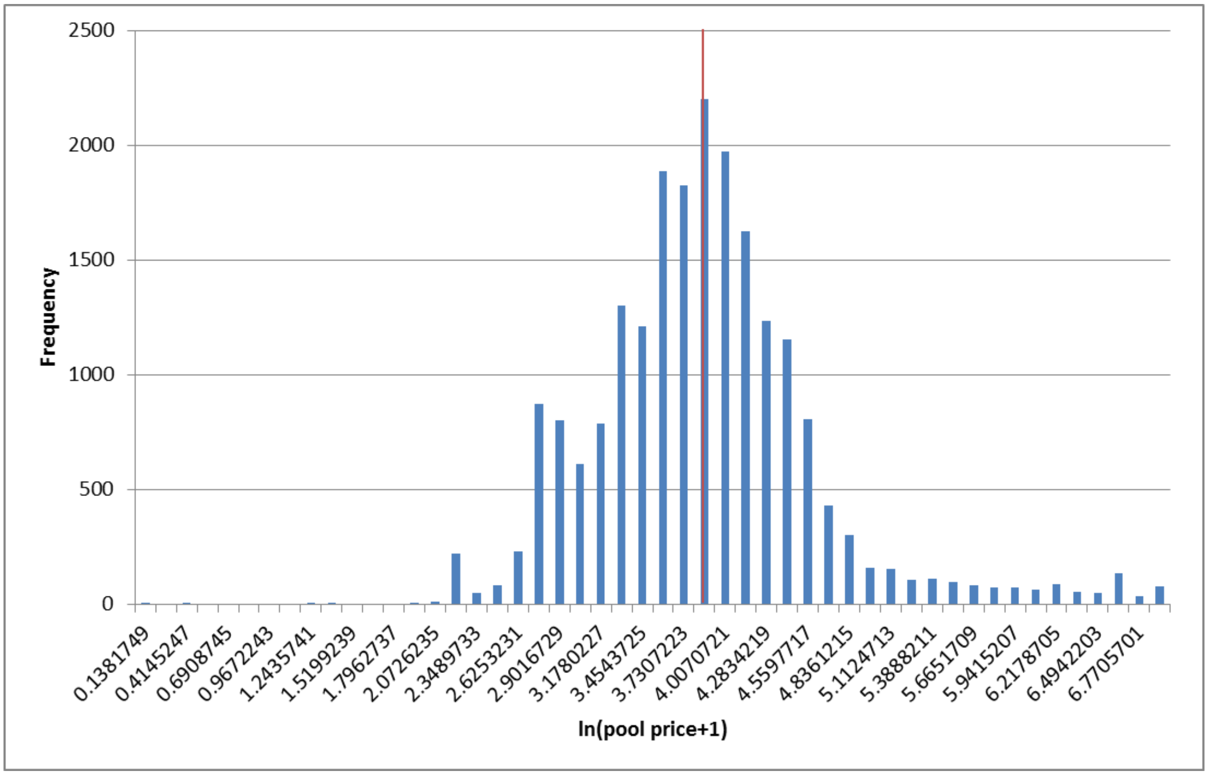
Figure 3.12: Histogram of Pool Price Frequency (Dataset: February 2008 to June 2010 Inclusive, Mean=$65.56)



[12] Note in the transformation we have applied we have taken ln(pool price +1) in order to include events with a $0 price (no natural logarithm exists for zero) but as a consequence no values exist less than 1 so no values are stretched out.

[13] Note that the when the data is examined in supply cushion bands the skew is more pronounced for low levels of supply cushion.

Figure 3.13: Histogram of ln(Pool Price +1) frequency (Dataset: February 2008 to June 2010 Inclusive, Mean=3.8)



Extreme values – a log transformation is sometimes preferred since the squeezing of data can reduce the influence of extreme values. In the case of pool price data we don't expect this to have much impact since any truly extreme values are already truncated by the price cap or floor.

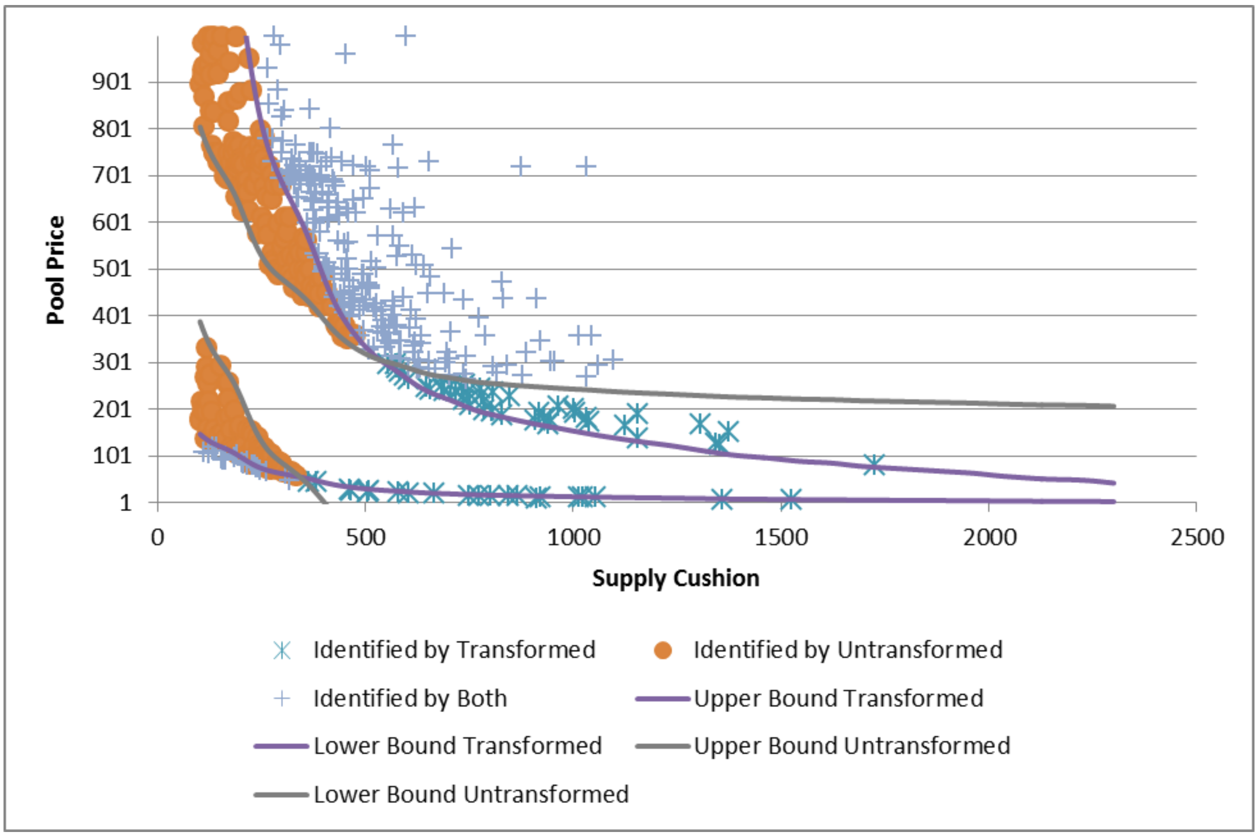
Unequal variation – A common pattern observed in data is that groups (or in this case, supply cushion bands) with large means have large standard deviations. This property is expected in observations of pool price, i.e. we might expect price observations to have a low standard deviation at low levels where the merit order is well populated and a higher standard deviation at high price levels were the merit order tends to be thinner. Again the squeezing transformation helps to equalize the variation in different groups and makes subsequent statistical interpretation easier**.**

It is important to recognize that a logarithmic transformation may not resolve these issues completely but it is a relatively simple and in many cases an effective step to mitigate them. In the next section we show the results with and without a log transformation.

3.2.1. Comparison of Results With and Without Log Transformation

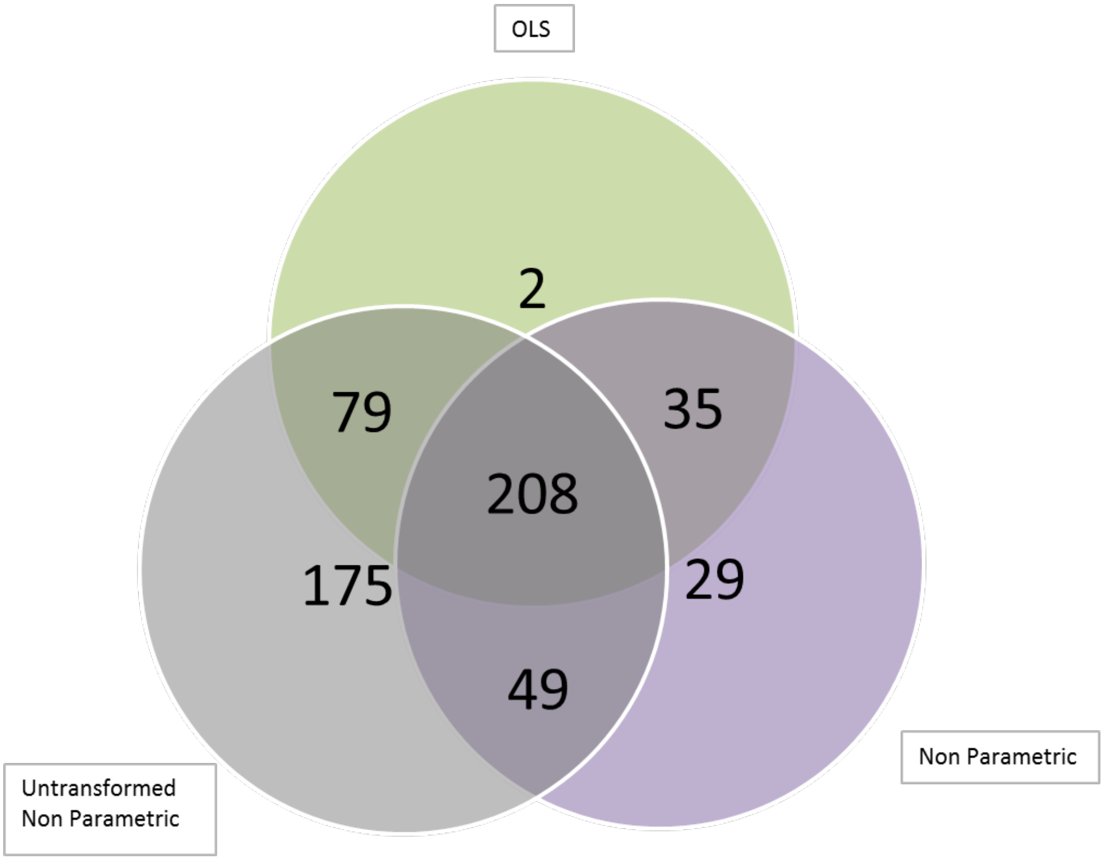
As noted above, the logarithmic transformation goes a long way to making the data more tractable. It reduces skewness and makes the distribution much closer to normal. If the non parametric method is applied to untransformed data, it identifies 511 outliers, compared to the 320 when used on the transformed dataset. Figure 3.14 Shows the hours and ranges identified by both options.

Figure 3.14: Hours Identified by the Non Parametric Method applied to Transformed & Untransformed Pool Prices (Dataset: February 2008 to June 2010 Inclusive)



When the data is transformed, it becomes much easier to identify events of interest for higher values of the supply cushion, while the untransformed identifes many more hours when the supply cushion is small. Figure 3.15 shows how all 3 continuous options (OLS, NP, and untransformed NP) compare:

Figure 3.15: Venn Diagram of Events of Interest Identified by Different Continuous Methods (Dataset: February 2008 to June 2010 Inclusive)



Without the logarithmic transformation, the non parametric estimates identify considerably more hours, particularly when the market is tight. It also has the downside that it cannot flag unusually low prices, except at small supply cushion levels. The MSA believes that the logarithmic transformation benefits rather than harms the analysis, and there is no comparable benefit to working with untransformed data for the above methodologies.

3.2.2. Fitting distributions to the data

An alternative to using a log transformation would be to consider other options that would address skewness and the other undesirable features of the data. The supply cushion technical working group produced a method called BRAUMS, or Bounds by Regression Analysis Upon Modeled Statistics described in Appendix A (the results are included in the data file released with this report). The MSA believes the approach has merit but relies on software not currently available to the MSA such that we have been unable to confirm the results.

Figure 3.16 compares the outliers identified by the BRAUMS, Non Parametric, and 250 MW Band methods:

Figure 3.16: Venn Diagram of Events of Interest Identified by BRAUMS, 250 MW Band, and Non-Parametric Methodologies (Dataset: February 2008 to June 2010 inclusive)

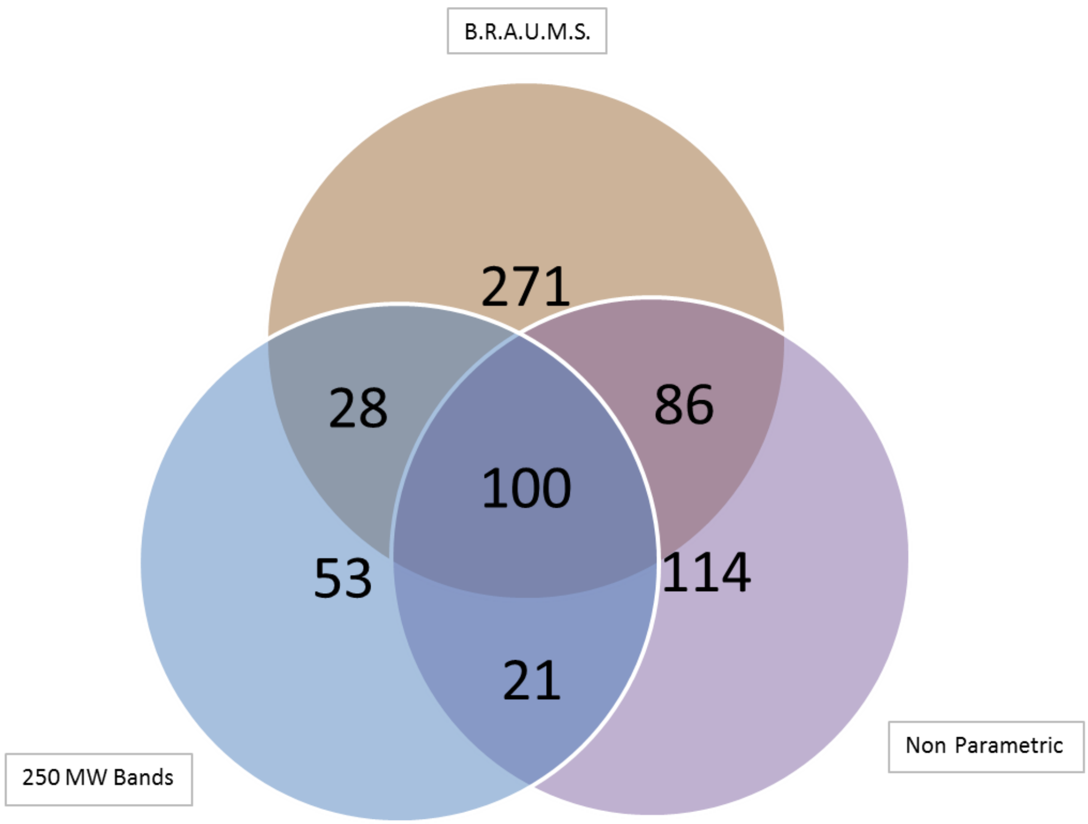
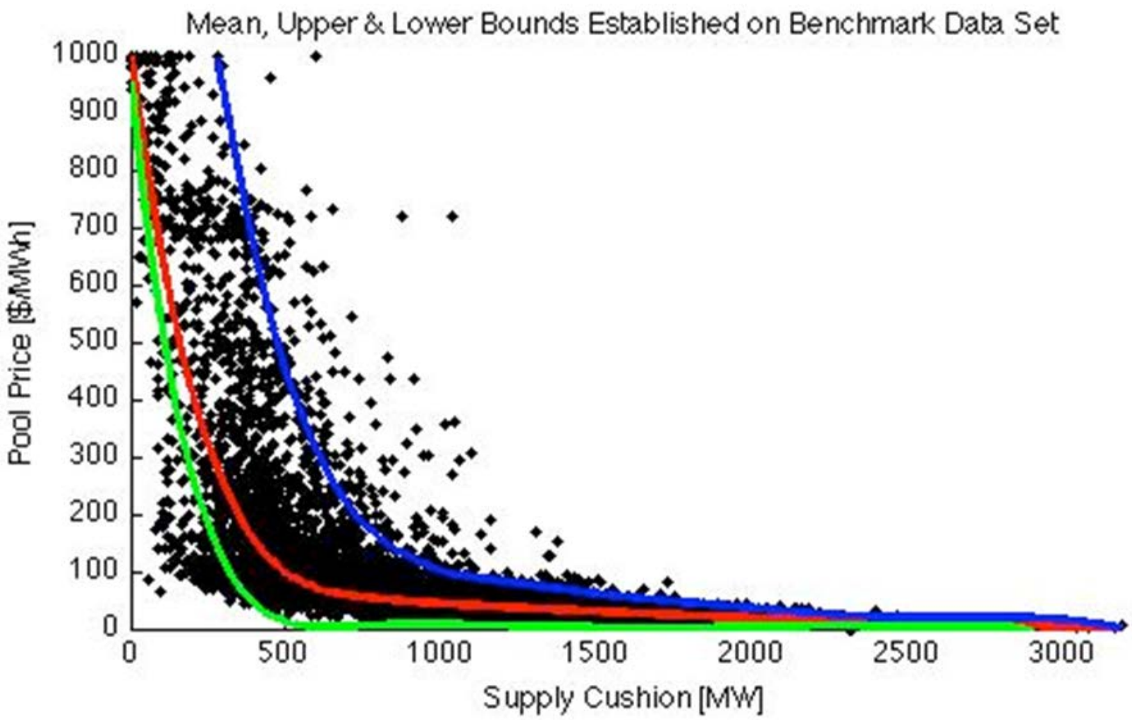


Figure 13.17 shows the BRAUMS results on the benchmark dataset. The MSA notes that the method identifies many downside outliers at low supply cushions that other methods do not.

Figure 3.17: Results of BRAUMS Analysis (Dataset February 2008 to June 2010 inclusive)



3.2.3. Bayesian Method

The supply cushion technical working group also discussed whether an approach based on Bayesian statistics might be appropriate. In simple terms, a Bayesian approach allows inferences to be made based on probabilities rather than relying upon confidence or prediction intervals. The disadvantage of the Bayesian approach is the complexity of the models involved and as a consequence it may not offer a significant enough improvement over the other methods considered. This approach was thus not considered further in this analysis.

3.3 SELECTION OF THE BENCHMARK DATA SET

When the MSA first proposed the supply cushion metric and its application to identify interesting events, a benchmark dataset was chosen from the period of Feb. 2008 to Jun. 2010. Some stakeholders have suggested the MSA should update the benchmark dataset in order to take into account changes in market structure and market dynamics. With this in mind the MSA believes it is helpful to articulate some attractive features of a benchmark dataset.

3.3.1. Attractive Features of a Benchmark Dataset

The following list is not an attempt to exhaustively describe what might make good criteria for evaluating a benchmark dataset; rather it lists some attractive features. Ultimately, the selection involves tradeoffs between competing priorities.

- **Contains a variety of supply cushion levels** – In selecting the original benchmark dataset used in section 2 the MSA wanted to make sure it contained a variety of supply cushion levels. The period included very tight (low supply cushion) periods in 2008 and periods in 2010 when transmission constraints resulted in tight markets conditions. It also contained a number of months of relative surplus that resulted from the economic downturn during the period.
- **Sufficient observations to produce robust estimates** – larger numbers of observations are likely to produce more robust estimates of means and standard deviations. Some methods may require more information than others (for example, a fixed band size method may require more data than a non parametric method that includes a Kernel function that draws information from a wide range of neighboring observations).
- **Fundamentals not included in the supply cushion are not likely to have a large impact on price outcomes** – In November 2010 when the initial supply cushion was published the MSA was concerned that if gas prices changed the benchmark dataset would not be representative of fundamentals. We consider this point in more detail in section 3.4. To some extent more recent data may be preferable to data from many years ago.
- **Historical** – Given that the MSA uses the supply cushion analysis to detect events as they happen it is important that it relies upon historical data, i.e. data that is known, rather than data for the current calendar year**.**
- **Appropriate to the analysis –** In some instances the choice of benchmark data set may be driven by a desire to test whether a particular rule or regulatory change has had an impact on price formation at a given level of supply cushion i.e. conducting a before and after analysis.

3.3.2. Comparison of Alternate Benchmarks

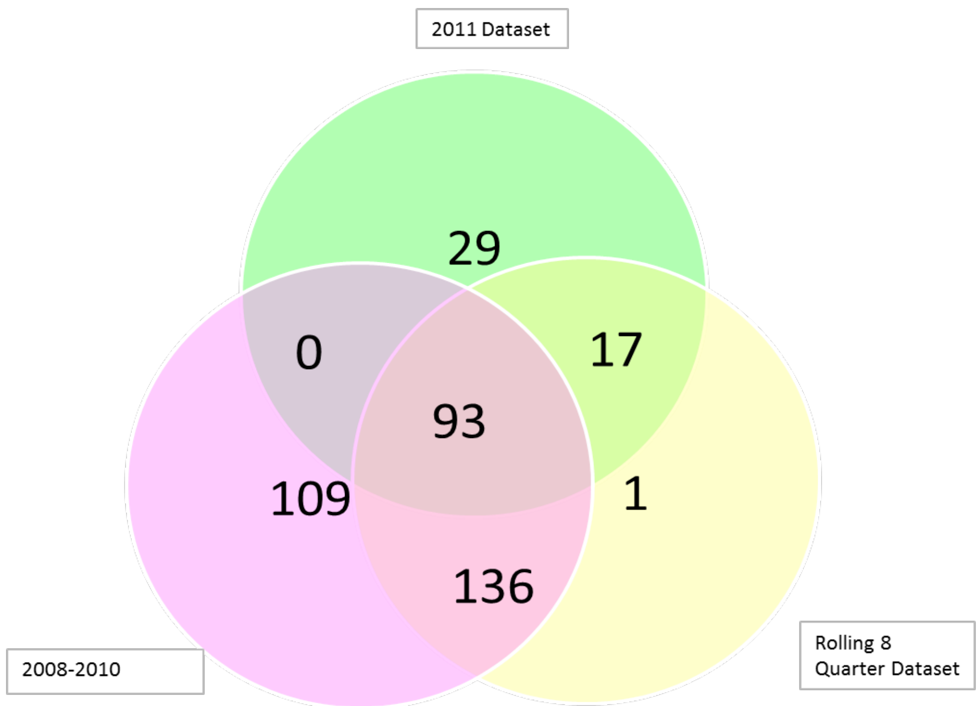
To illustrate the impact of benchmark dataset selection we consider the detection of outliers in 2011 using the 250 MW band approach outlined in Section 2 with three different benchmark datasets:

a) Feb 2008 – Jun 2010 benchmark
b) Rolling 8-quarter benchmark

c) 2011 benchmark[14]

Figure 3.18 shows the intersection outliers detected using the three alternate benchmarks.

Figure 3.18: Venn Diagram of the 2011 Outliers Identified by Estimates Produced by Various Benchmark Datasets

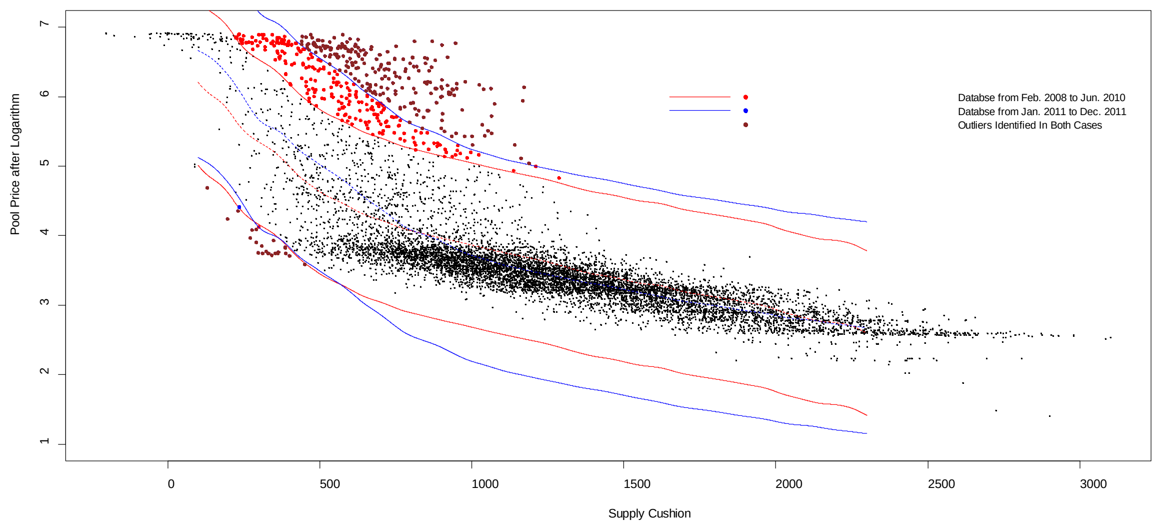


When analyzed using the 2008-2010 benchmark dataset 2011, 338 events of interest are detected. Using a 2011 benchmark the number drops to 139. The estimates based rolling 8-quarter benchmark dataset came somewhere in the middle at 247. The number of events detected is not of primary concern. For example, a method or benchmark that identifies few events outside 3 standard deviations of the mean would allow the MSA to focus on events outside 2 standard deviations without increasing the resources required to analyze the events. The number of events, given a fixed standard, may be interesting only in that it might reveal a trend or change in pricing dynamics. The MSA believes that the rolling 8 quarter dataset, while marginally more computationally complex, is not without merit. Additional analysis might shed further light on the 'quality' of events identified by both options. Further, the MSA notes that the choice of 8 quarters is arbitrary and that 12 or 16 quarters are equally feasible as a benchmark

[14] A 2011 benchmark is included just for illustration. Obviously, a same-year benchmark dataset is only available at the end of the year. It would thus be unsuitable for the detection of outliers during the period.

For further illustration in Figure 3.19 we show the results from the non parametric method using two of the benchmarks: Feb 2008 – Jun 2010 and the 2011 benchmark. The 2011 benchmark dataset predicts a higher expected pool price at a supply cushion less than about 900 MW but a slightly lower expected value at a supply cushion greater than about 900 MW. The new dataset also generates a wider standard error (note that the 2011 benchmark contains significantly fewer observations). The combination of these two effects results in significantly less upside outliers at a low level of supply cushion.

Figure 3.19: Comparison of Non Parametric Method with Different Benchmark Datasets



3.4 ADDITIONAL FUNDAMENTALS

3.4.1. Fuel Costs

As noted above the supply cushion methodology does not account for changing fuel costs during the benchmark period or during the period of study. The most significant of these would be the influence of natural gas prices. In the February 2008 to December 2011 benchmark dataset gas prices fluctuated from a low of $2.16/GJs to a high of $11.47/GJs. It is reasonable to assume that while the correlation between pool price and natural gas price is not strong it is a significant factor in the determination of market outcomes at high levels of supply cushion where the marginal offer is more likely to reflect cost. The MSA has also observed the TMR Reference Price, which is linked to the gas price, has an impact on price setting. The MSA has considered some alternatives for modifying supply cushion analysis to account for gas price but has not pursued these given that natural gas prices remain very low and relatively stable.

3.4.2. Other Factors

The supply cushion estimate is defined as the difference between available and dispatched MW in the merit order. The estimate includes availability of wind (i.e. increased wind generation causes a reduction in dispatch of other generation in the merit order, all else equal), actual load response, impact of transmission constraints on generation availability along with derates and outages. As the supply cushion is a simple one dimensional metric of market tightness, by relying on it we ignore the extra information that would be conveyed if one examined the composition and context of a given supply cushion. For example, a given level of supply cushion that results from high wind availability rather than high availability at thermal generators results in an observable difference in the expected pool price. Similarly, a given level of supply cushion appears to result in different price outcomes between the on-peak rather and off-peak periods. To some extent these results may shed light on limitations in the supply cushion analysis but they may also represent interesting results worthy of further study. (For example, they may reveal a weakness in the current market design or operation). For this reason the MSA has not pursued methods that might 'correct' the estimate of the mean / stand deviation as a result of these other factors.

It's worth reiterating that the supply cushion analysis is only a tool to identify hours of interest that will then be further analyzed, not the be-all and end-all of the MSA's oversight. The MSA believes this further scrutiny is currently the best approach to taking into account the impact of the 'other factors' identified in this section.

3.5 MORE DISCRETE DATA

In this section we consider a variety of issues related to data availability and in particular whether the fact that MSA's current supply cushion estimate based on a snapshot taken during the middle of each hour limits the usefulness of the metric.

3.5.1. Historical Data on the Link Between Pool Price and Supply Cushion

If the data were available the MSA believes it would be an interesting exercise to examine how the relationship between supply cushion and pool price has evolved since the market opened. Unfortunately data prior to February 2008 is limited to the point that that we do not believe comparable estimates could be constructed.

The supply cushion is estimated using merit order snapshots that are taken every 5 minutes. The snapshot data has only been recorded by the AESO since mid-January 2008. More fundamentally the supply cushion estimate relies upon the "must offer" obligation introduced in the Quick Hits Rule amendments that were introduced on December 3, 2007. Prior to this time available MW did not have to be offered, although were recorded in different categories of Total Declared Energy (TDE). In addition market participants were required to manually restate the availability of generating units when dispatched for reserves. The MSA does not believe that there is a reasonably efficient way of accounting for these differences and therefore has not been able to construct robust supply cushion estimates prior to February 2008.

3.5.2. More Discrete Observations of Supply Cushion

As noted above, snapshots of the merit order are taken every five minutes.[15] Theoretically, each of these snapshots could be used to construct an estimate of the supply cushion that could in turn be compared with pool price (or the prevailing SMP, see discussion in section 3.5.3). Alternatively, the five minute snapshots could be averaged in a given hour to give an average hourly supply cushion rather than a snapshot in time. Both offer possible refinements to the method currently used by the MSA at the expense of transparency (only the snapshots from the middle of the hour are made public on the AESO website) and computational effort. For these reasons the MSA has not examined whether more discrete observations on supply cushion would represent a genuine refinement. Instead, when examining detected outliers the MSA conducts a quick check that the middle of the hour was representative of the hour as a whole. If not, the hour is unlikely to receive further scrutiny.

From March 2012 onwards additional merit order data (above and beyond the five minute snapshots) has become available. Theoretically, this would allow a minute by minute estimate of supply cushion during an hour. However, this data is not available historically so it would be sometime until there was sufficient data to construct a robust benchmark. For the reasons stated above the MSA does not intend to pursue this approach at present, but acknowledges that it may be worthwhile to investigate in the future.

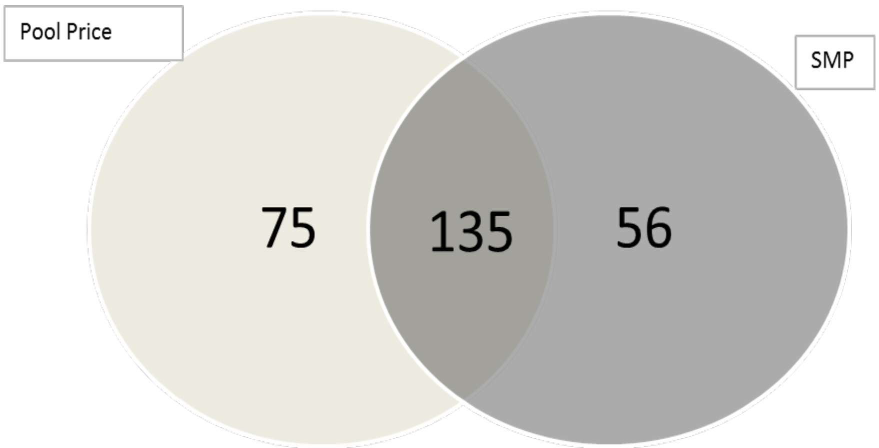
3.5.3. Pool Price vs. SMP

The MSA's original supply cushion approach and the variants consider in this report all examine the relationship between supply cushion and pool price. Given the supply cushion estimates are based on a middle of the hour merit order a case could be made to compare to the prevailing system marginal price (SMP) at that time, rather than the pool price for the hour.

To illustrate the impact of pool price v. SMP in Figure 3.20 we show a Venn diagram with a comparison of the results with a 250 MW band size.

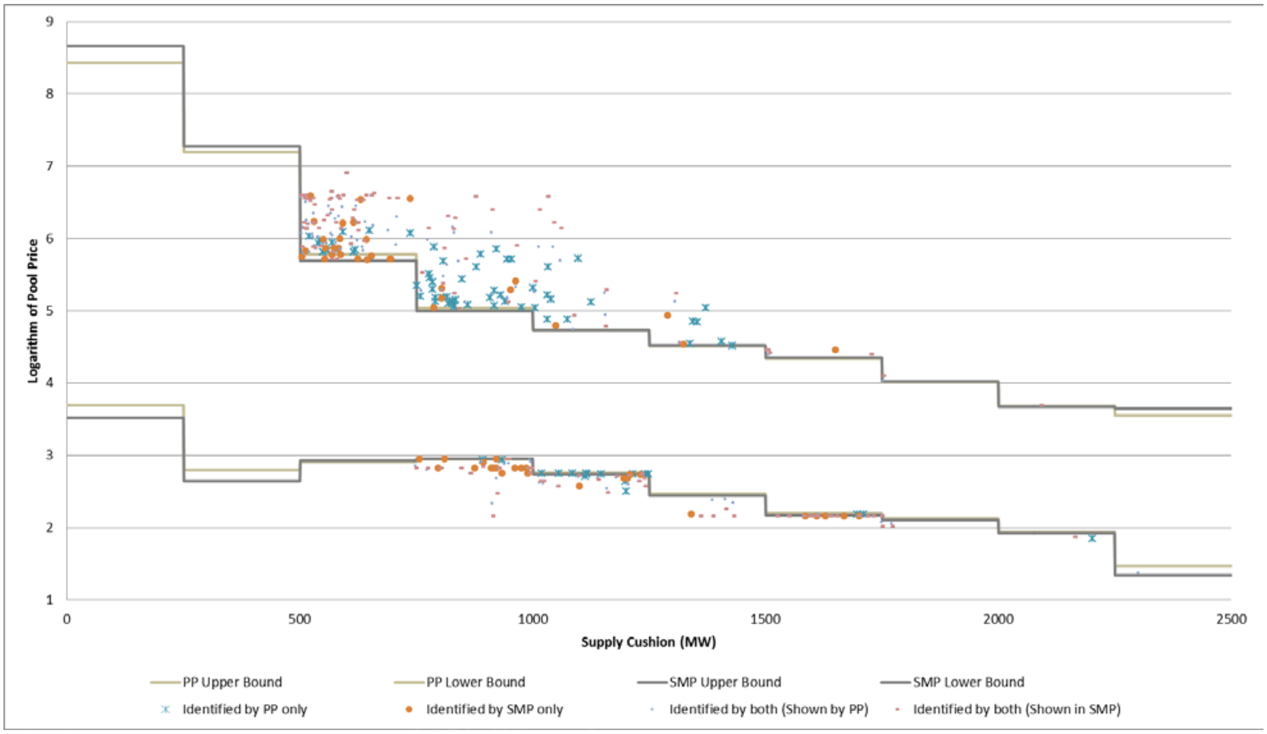
[15] Some of the 5 minute observations are missing corresponding during IT outages or other issues affecting AESO systems.

Figure 3.20: Venn Diagram Showing Intersection Pool Price and SMP Outliers Using a 250 MW Band Size (Dataset: Feb 2008 to Jun 2010 Inclusive)



The two approaches identify many of the same hours. The differences result from two causes: differences in the estimates of the mean and standard deviation and differences between pool price and SMP. The second cause is the more significant - the average difference between SMP and pool price when only one metric flagged an hour was $91.79. Figure 3.21 shows the identified hours (note that as hours identified by both metrics have two prices associated with them, both have been shown)

Figure 3.21: Comparison of Events of Interest identified by Pool Price and SMP (Dataset February 2008 to June 2010 inclusive)



Using pool price rather than SMP has a number of advantages. The MSA is concerned with pool price in relation to market tightness not SMP per se; it is pool price that settles the market and ultimately speaks to competition and/or manipulation. Further, pool price already includes information on abnormal SMPs, conveniently weighted based on their impact. In contrast, instantaneous SMP gives relatively little information about the pool price, and may not be representative of the hour on the whole as by its nature SMP is more volatile than pool price. Pool price is the fundamental variable of interest and has some benefits while there is little clear advantage to switching to SMP, leading the MSA to conclude that pool price is more appropriate in this analysis.

This is not to say that the MSA believes that measures based on the SMP have no merit in identifying events of interest. For instance, a complementary metric could be constructed based on variation between the SMP and pool price that could flag abnormally volatile market conditions. Alternatively, should the MSA start employing multiple observations of the supply cushion in the given hour it may be appropriate to pair those with the SMP. The MSA may investigate such measure in the future, but does not currently propose to do so.

4. Conclusions

This paper reviews the original methodology proposed by the MSA for the use of a supply cushion metric in detecting price outliers. The MSA continues to believe that this approach has merits over a manual process for detecting these hours – namely that it is less time consuming, more consistent, and more transparent.

The paper also considers a number of refinements to the existing methodology that have originated from the MSA or the supply cushion technical working group. We have described what we believe to be the relative merits of each. In some cases there is a real tradeoff between elegance and computational effort. In other cases the tradeoff is between a methodology that is (relatively) simple to understand and one that relies on fewer or different assumptions about the underlying data. Throughout the report we have included a number of Venn diagrams indicating the intersection between events of interest (more than 3 standard deviations from the mean). The detection of some events is common across many or all methods considered, in other cases an event that fails to meet the 3 standard deviation threshold would still be more than two standard deviations away from the mean.

The MSA believes the non parametric method (subsection 3.1.2.2) overcomes some of the issues related to band size selection (subsection 2.1.2.3) and avoids the assumptions about functional form associated with the OLS projection method. Some methods based on alternative distributions (subsection 3.2) appear to have attractive features but are the most computationally complex. In the MSA's view, the original 250 MW band size method remains the most easily understood and most transparent of the alternatives considered.

In addition to the methodological refinements we have considered the choice of benchmark dataset and whether more discrete data should be employed (more frequent estimate of supply cushion or use of SMP rather than pool price). While it is possible to describe some attractive characteristics of a benchmark dataset the MSA believes that ultimately the decision about what benchmark is appropriate may come to the question it is seeking to answer. For example, evaluation of a rule change may obviously suggest a benchmark data set before the change came into effect. With respect to more discrete data the MSA is not yet convinced this is would be an area where the potential gains are large.

The MSA appreciates the participation of stakeholders in the technical working group and the interest in the techniques the MSA has developed. With respect to the supply cushion methodology the MSA intends to consider more efficient ways of classifying events of interest over time so that trends can be observed. To the extent that stakeholders remain interested in the MSA's approach to monitoring we will continue to share those results as they become available.

Appendix A: B.R.A.U.M.S. (Bounds by Regression Analysis)

A.1 INTRODUCTION

This appendix introduces an alternative method, Bounds by Regression Analysis Upon Modeled Statistics (BRAUMS), for finding events of interest. This methodology has the following characteristics.

Continuous upper/lower bounds are established by regression fitting

Upper/lower bounds are determined without using a log transformation on price data

Price distribution about the mean is not assumed to be normal (i.e. symmetrical)

The BRAUMS methodology was developed by JCTC Solutions Inc., sponsored by Industrial Power Consumers Association of Alberta. JCTC Solutions Inc. was one of the participants in the supply cushion technical working group. A summary of the results is included in the data file accompanying this report.

A.2 METHODOLOGY DESCRIPTION

BRAUMS is the term chosen to describe the proposed methodology because it provides an overall description of the process. This method achieves continuous upper/lower bounds on market price through performing least square regression fitting on the series of upper and lower acceptable price values established at different supply cushion levels. This method can be broken down into the following four sequential steps:

Divide supply cushion and price domains into discrete bins.

For each supply cushion bin, examine the corresponding price distribution to determine which Probability Density Function (PDF) provides the closest approximation.

Use the PDF model to calculate the mean and upper/lower bound price values.

Perform regression fitting on the set of mean, upper and lower prices over the entire supply cushion range to establish continuous bounds.

The following subsections provide a more in-depth description of each of the four aforementioned steps.

A.2.1 Step 1 – Discretizing the Data Domain

The Market Surveillance Administration (MSA) has provided the benchmark dataset to members of the TWG for the purpose of exploring alternative SCA methods. The following figure provides a visual illustration of the supply cushion vs. hourly market pool price relationship in the benchmark dataset.

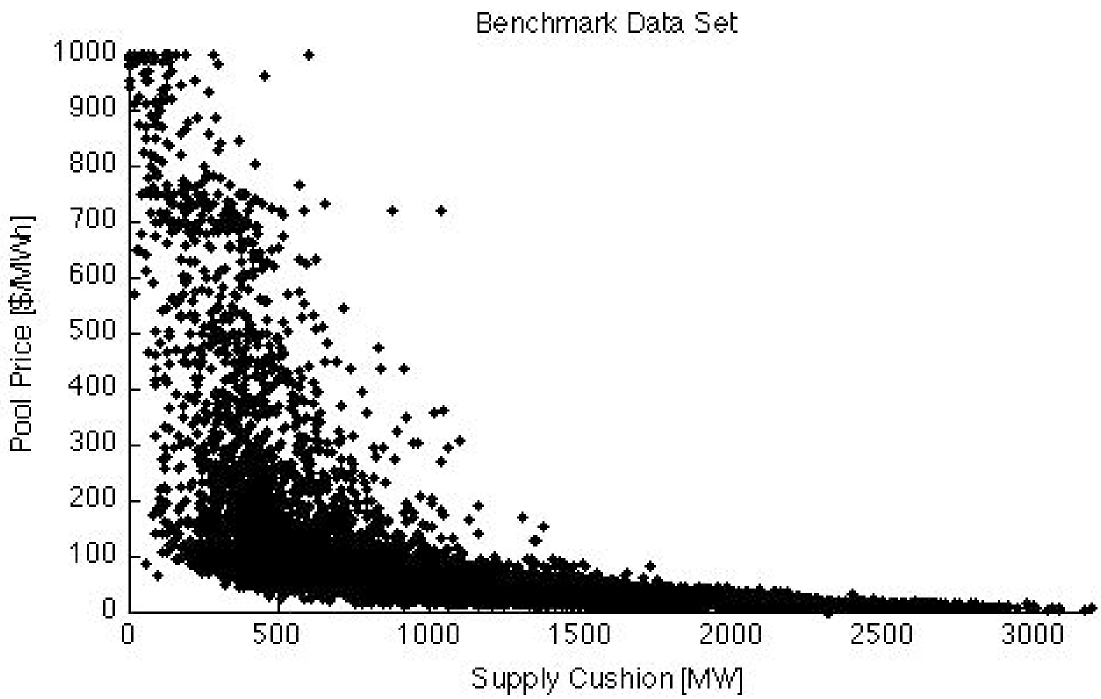


Figure 22: Benchmark Dataset Illustration

A 5MW supply cushion bin size was chosen for discretizing the independent X-axis. This bin size was chosen because it provides a very fine resolution, while still retains sufficient number of data points in the majority of supply cushion bins. Next, a $5/MWh price bin size was also chosen for discretizing the dependent Y-axis. The pool price also needs to be discretized in order to sort the price data (within the same supply cushion bin) into histograms – which provides the basis to determine the type of PDF that provides the best approximation.

A.2.2 Step 2 – Modeling the Benchmark Dataset

In order to properly model the benchmark dataset, it was helpful to visualize it in three-dimensional space. This is illustrated below:

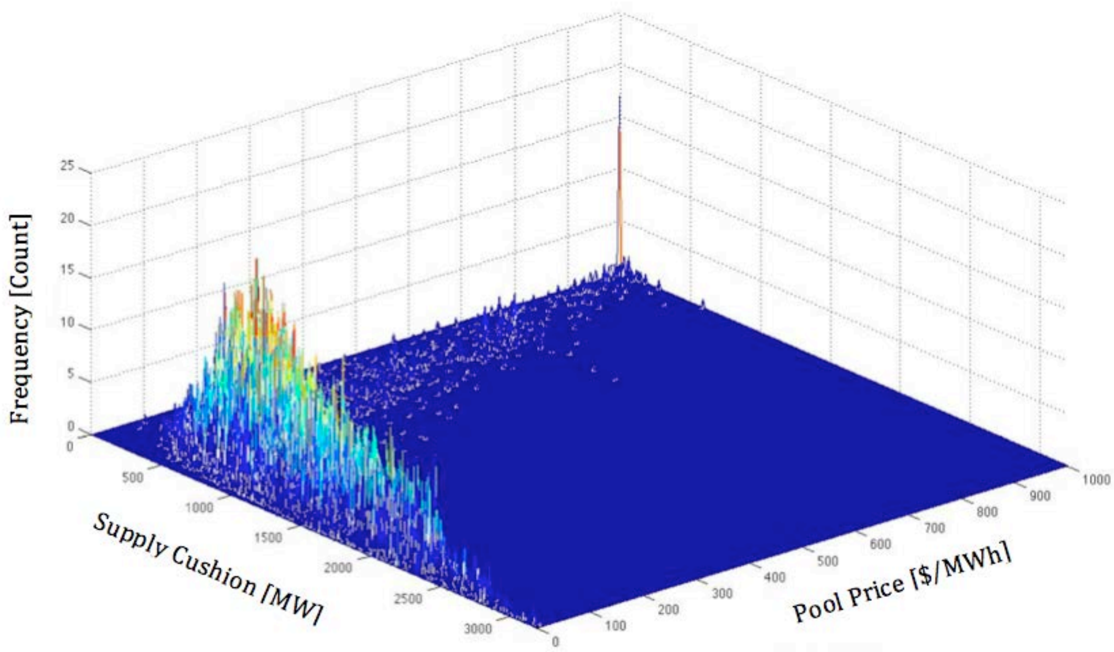


Figure 23: Benchmark Dataset 3-D Visualization (View from Bottom Right Corner)

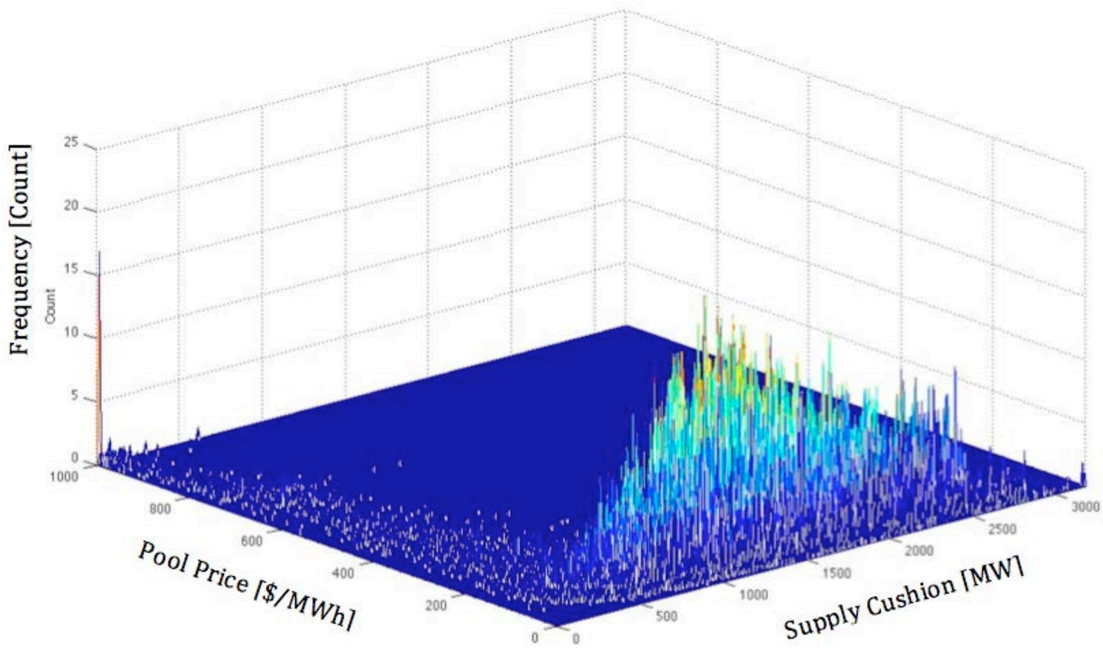


Figure 24: Benchmark Dataset 3-D Visualization (View from Bottom Left Corner)

The aim of Step 2 is to examine the set of price data in each supply cushion bin to determine what is the most appropriate PDF to describe the price distribution. In this preliminary version of BRAUMS, the price dataset is first tested for Gaussian distribution (the Lilliefors test is adopted). If the hypothesis test passes, a Gaussian distribution is applied. If the hypothesis is rejected, a Gumbel PDF is used instead. The Gumbel PDF is appropriate in this application since it resembles a skewed Gaussian PDF. This property captures the fact that price distribution tends to skew towards the high or the low side depending on the supply cushion level (for example, it is expected that the price distributions at low supply cushion levels tend to skew towards the high side). More importantly, the Gumbel distribution is typically known for its usage in Type 1 Extreme Value Analysis and is particularly good for estimating the maxima and minima extremes. This property is useful for calculating the upper and lower acceptance bounds.

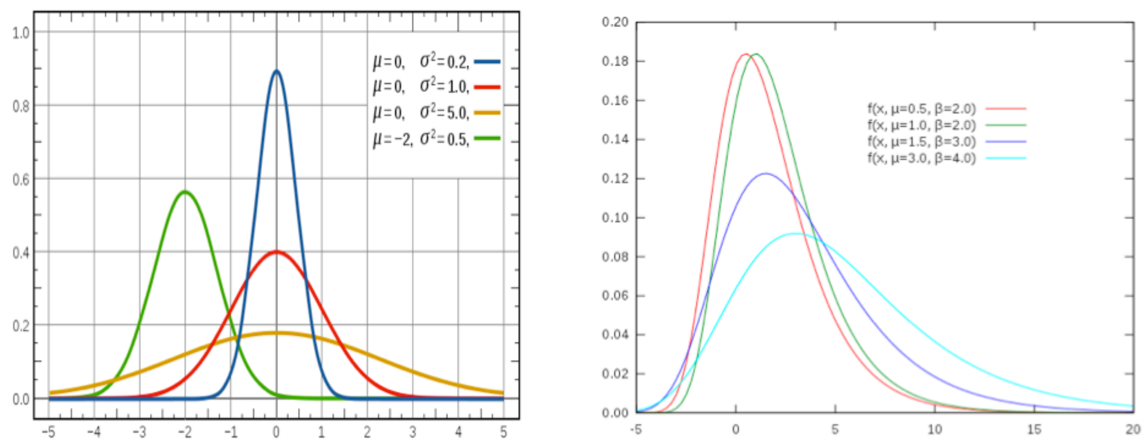


Figure 25: Gaussian (Above Left) and Gumbel (Above Right) PDF

The PDF approximation process continues from the first supply cushion bin to the last until every set of price data is fitted with either a Gaussian or Gumbel PDF. The complete set of PDFs then become the model used to describe the benchmark dataset. It is from this set of statistical models that the mean, upper and lower bounds will be calculated.

The figures on the next page provide an illustration of the resulting model created from this step. Note that these figures appear apparently different from and because of the difference in the vertical axis. The vertical axes in and describe the frequency of occurrence, while the vertical axes in and represent probability.

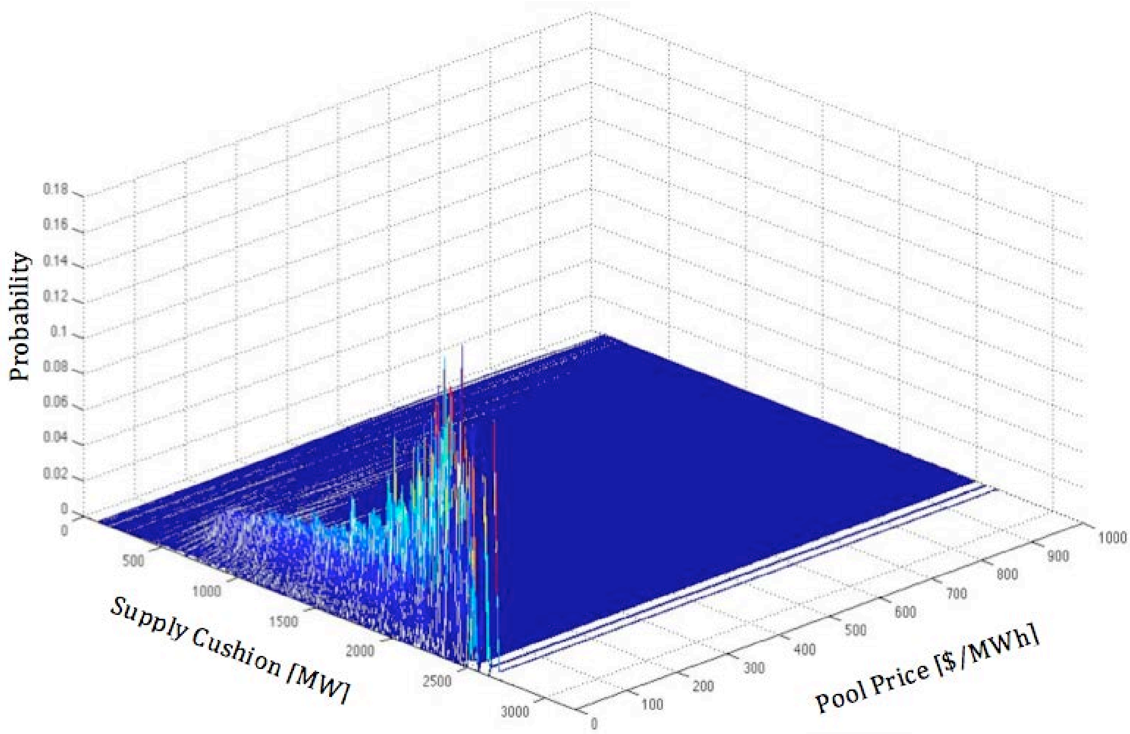


Figure 26: Benchmark Dataset Model Visualization (View from Bottom Right Corner)

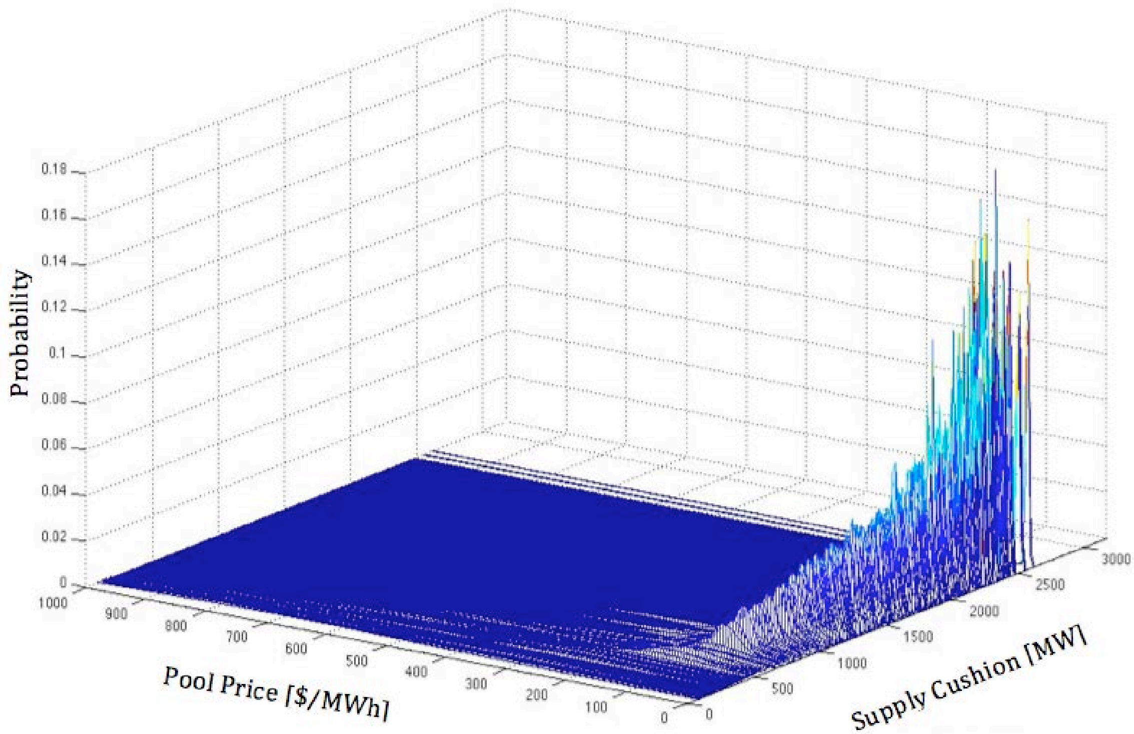


Figure 27: Benchmark Dataset Model Visualization (View from Bottom Left Corner)

Despite the differences, one can still use the benchmark data histograms (and) to confirm the validity of the statistical model created. For example, if one examines the high supply cushion region of the histogram, one realizes that the number of price samples is relatively fewer (reflected by the lower height in the histogram), but they are very concentrated within a small price range. Correspondingly, this feature is reflected in the probability diagrams in and , where the distributions are tightly packed into a small price range, resulting in the sharp peaks that denote high probabilities. As one moves into the low supply cushion region, the probability distributions become more and more spread out over a large price range. This feature reflects the wide range of price sample points in the histograms. These quick sanity checks provide confirmation that the probability models produced in Step 2 is descriptive of the underlying benchmark dataset.

A.2.3 Step 3 – Calculate the Mean, Upper and Lower Bounds

Once the probability model for each supply cushion bin has been established. The mean, upper and lower bounds can be calculated based on the Cumulative Distribution Function (CDF) of the underlying PDF. The CDF approach provides the users with the flexibility to adjust the statistical confidence levels, allowing the bounds to be made wider or narrower. Furthermore, the BRAUMS methodology inherently features asymmetrical upper/lower bounds due to that fact that asymmetrical PDFs (i.e. Gumbel distribution) are incorporated in the underlying statistical model. The asymmetrical bounds more closely reflect the reality of the problem since there is no reason to believe that the upper and lower acceptance bounds should be equidistance from the mean.

A.2.4 Step 4 – Regression Fitting

The last step in the BRAUMS methodology utilizes regression analysis to determine the set of mathematical expressions that best describe the mean, upper, and lower bound datasets obtained in Step 3. Iterations are typically required to determine the type of regression that achieves the most reasonable fit. The following figures illustrate the results of this fitting process. The fits superimposed on top of the datasets are obtained by using robust, least-square, regression fitting technique with polynomial splines. Note that any negative 2-sigma and 3-sigma price values are filtered out from the dataset before the regression analysis is performed. This is due to the fact that no negative pricing is allowed in the current market framework.

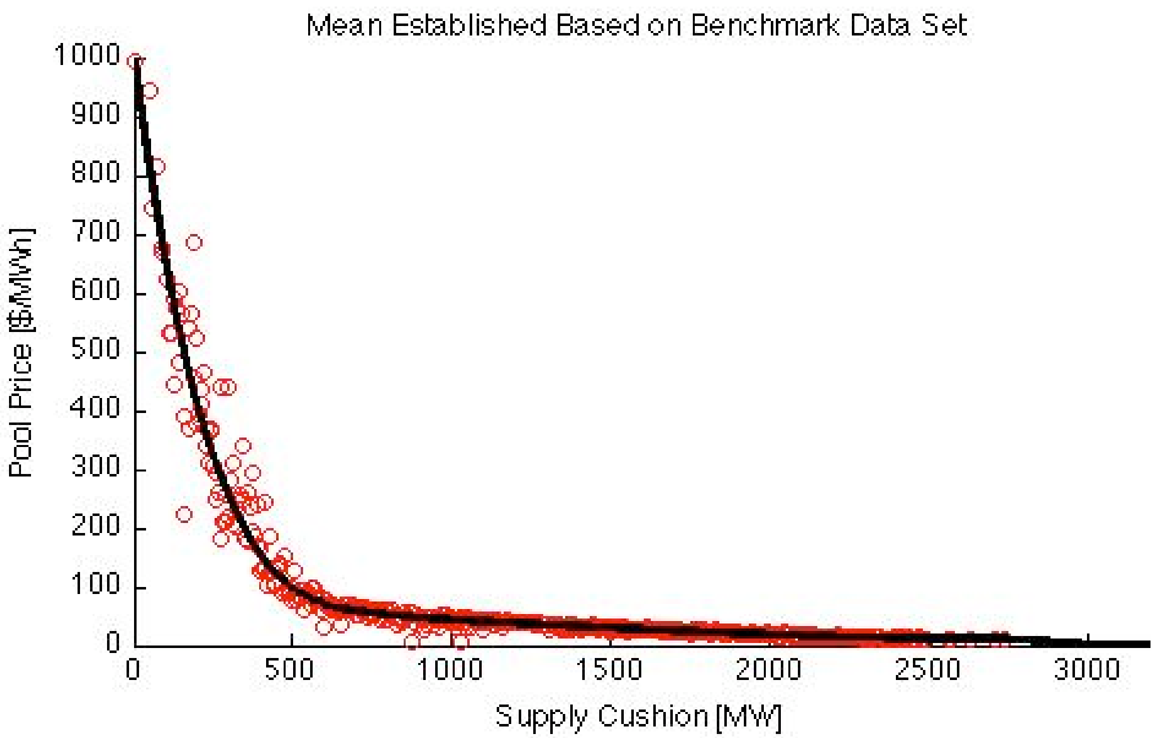


Figure 28: Mean Value Dataset Regression Fitting Result

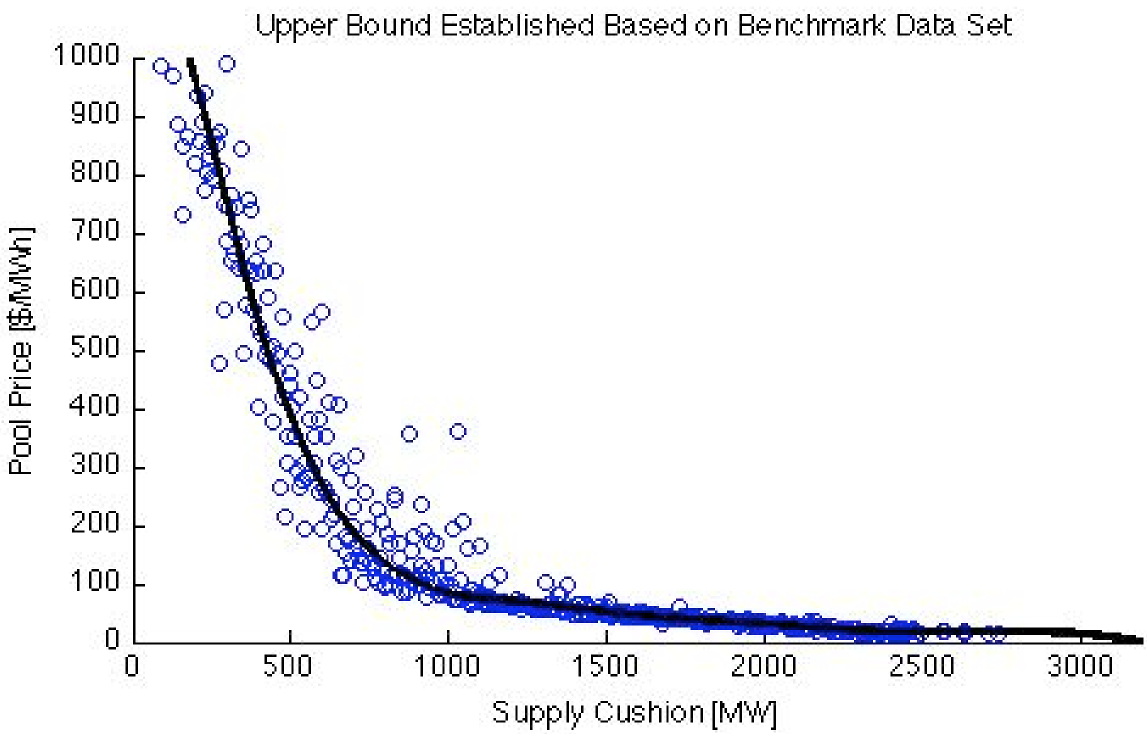


Figure 29: Upper Bound Dataset Regression Fitting Result (2-sigma)

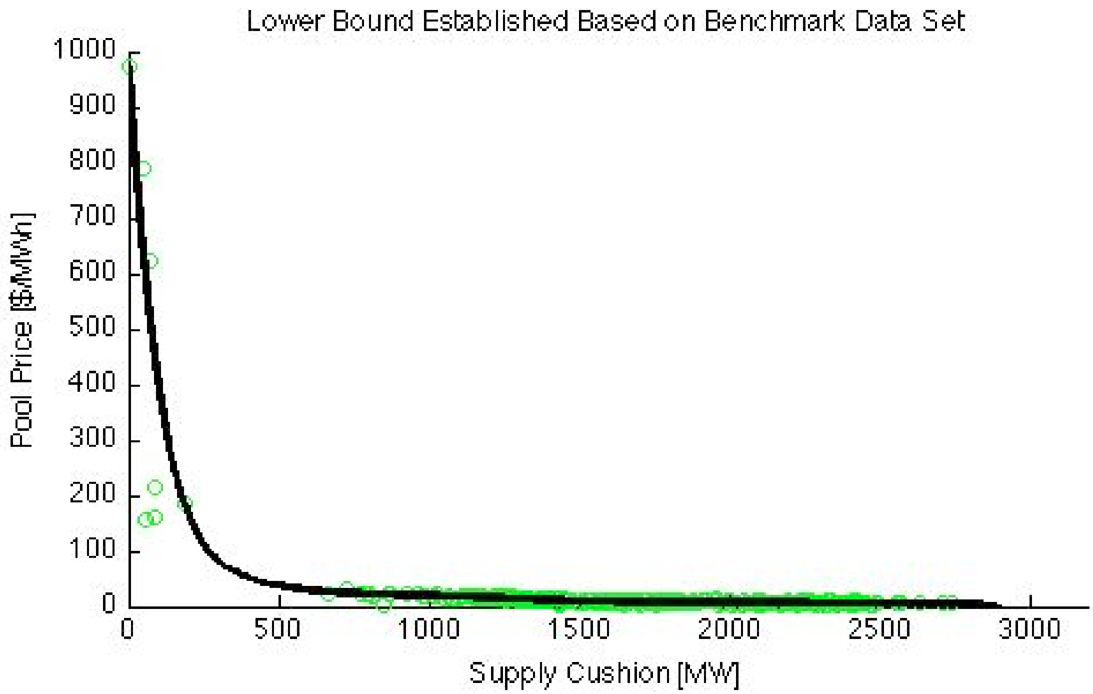


Figure 30: Lower Bound Dataset Regression Fitting Result (2-sigma)

The following figures illustrate the mean, upper and lower bounds applied on top of the benchmark dataset.

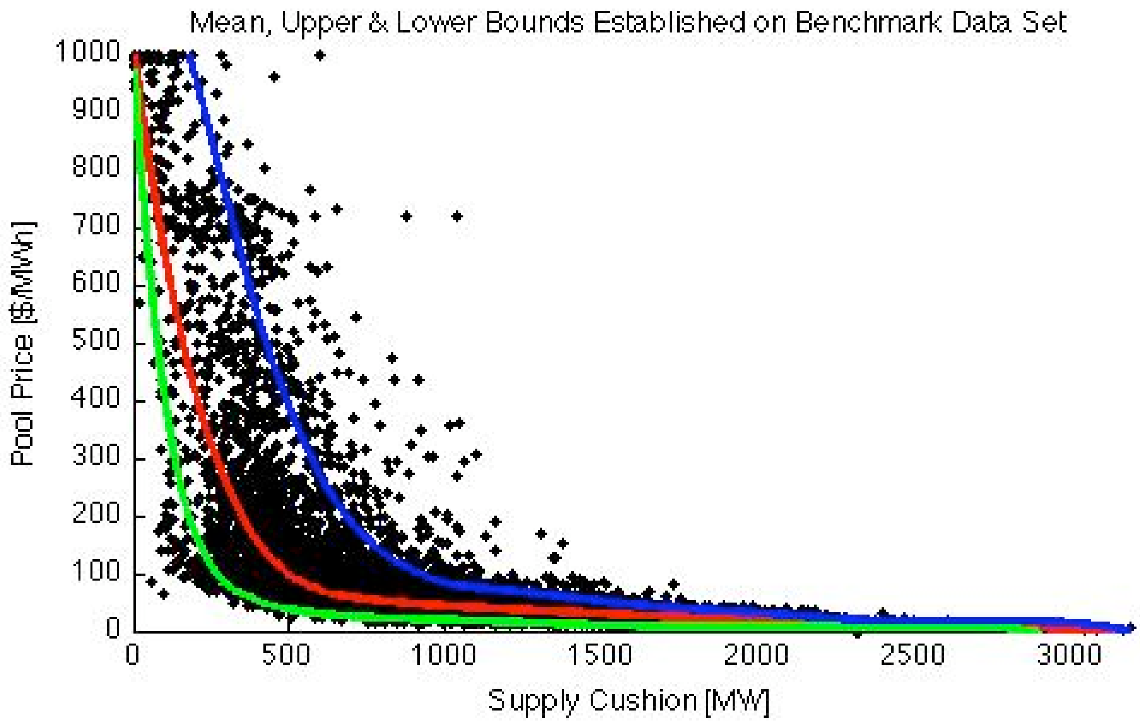


Figure 31: Mean, Upper and Lower Bounds (2-sigma) Established on Benchmark Dataset

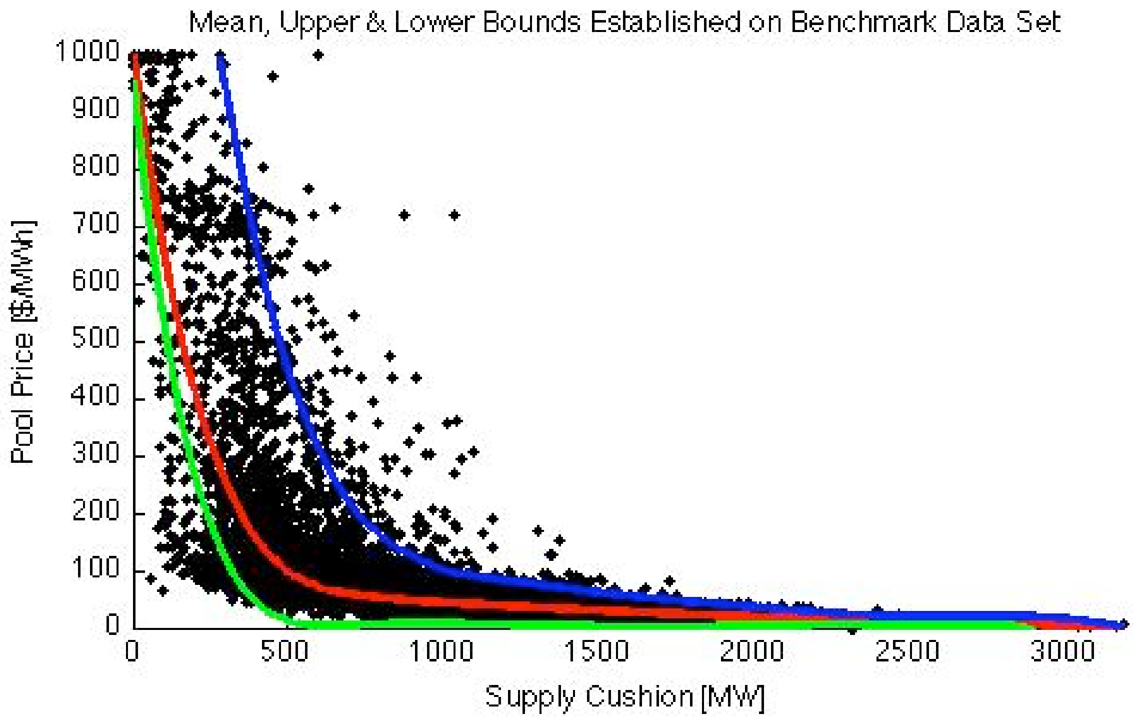


Figure 32: Mean, Upper and Lower Bounds (3-sigma) Established on Benchmark Dataset

presents the upper/lower bounds established based on 2-sigma confidence levels (statistically speaking, 2-sigma bounds capture 95.6% of the sample population) and presents the bounds established based on 3-sigma confidence levels (statistically speaking, 3-sigma bounds capture 99.7% of the sample population).

A.3 SUMMARY

The proposed BRAUMS methodology provides a set of continuous upper/lower bounds that are established based on a set of statistical models created from the underlying benchmark price data (as opposed to the logarithm of price). The implementation of hypothesis test and asymmetrical PDFs ensure that the resulting statistical models provide a close approximation to the underlying benchmark dataset. This also means that the methodology no longer relies on any normality assumption in the price distribution.

A.4 RESULTS: APPLYING THE BOUNDS TO HISTORICAL QUARTERLY DATA

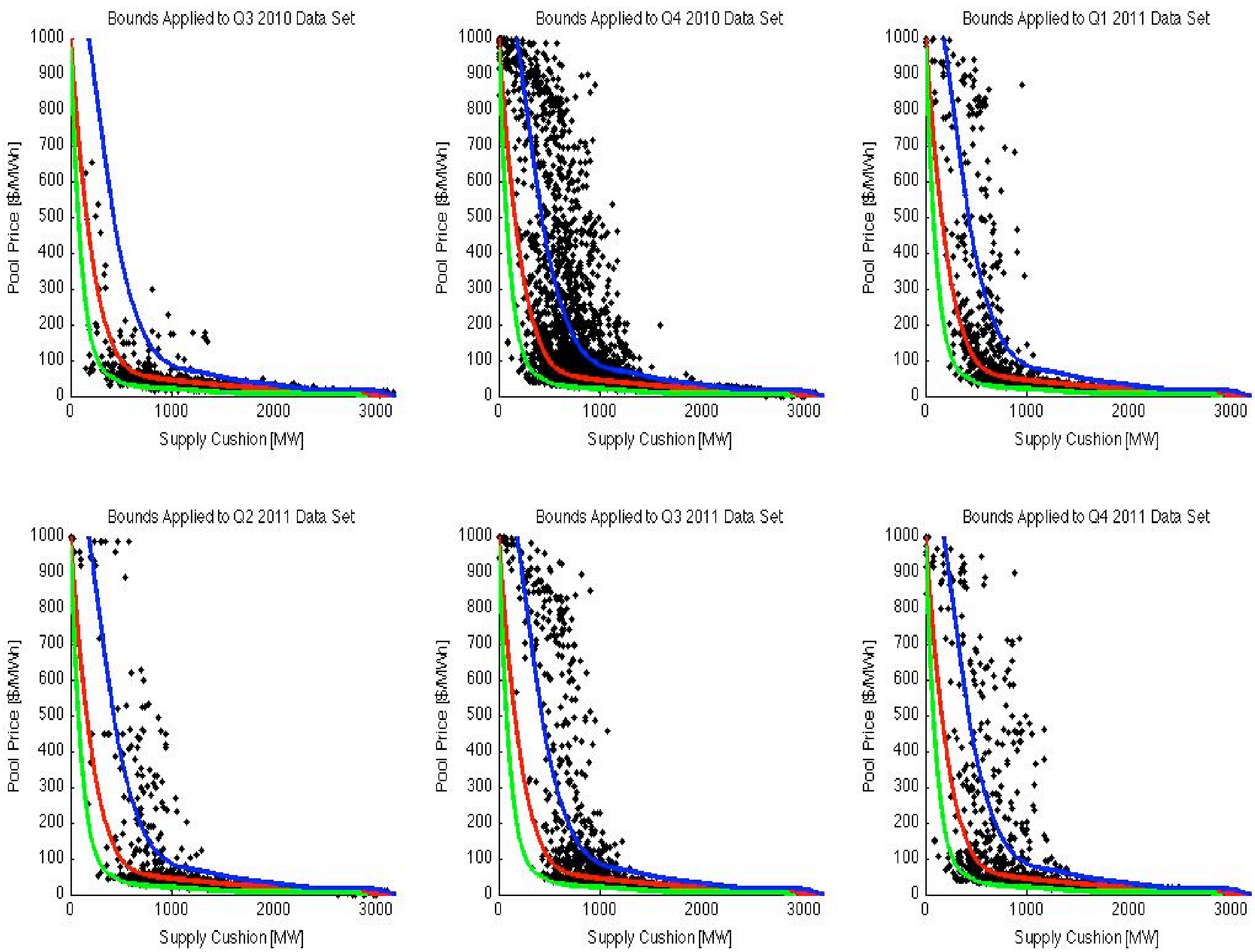


Figure 33: 2-sigma Bounds Applied to Historical Quarterly Data

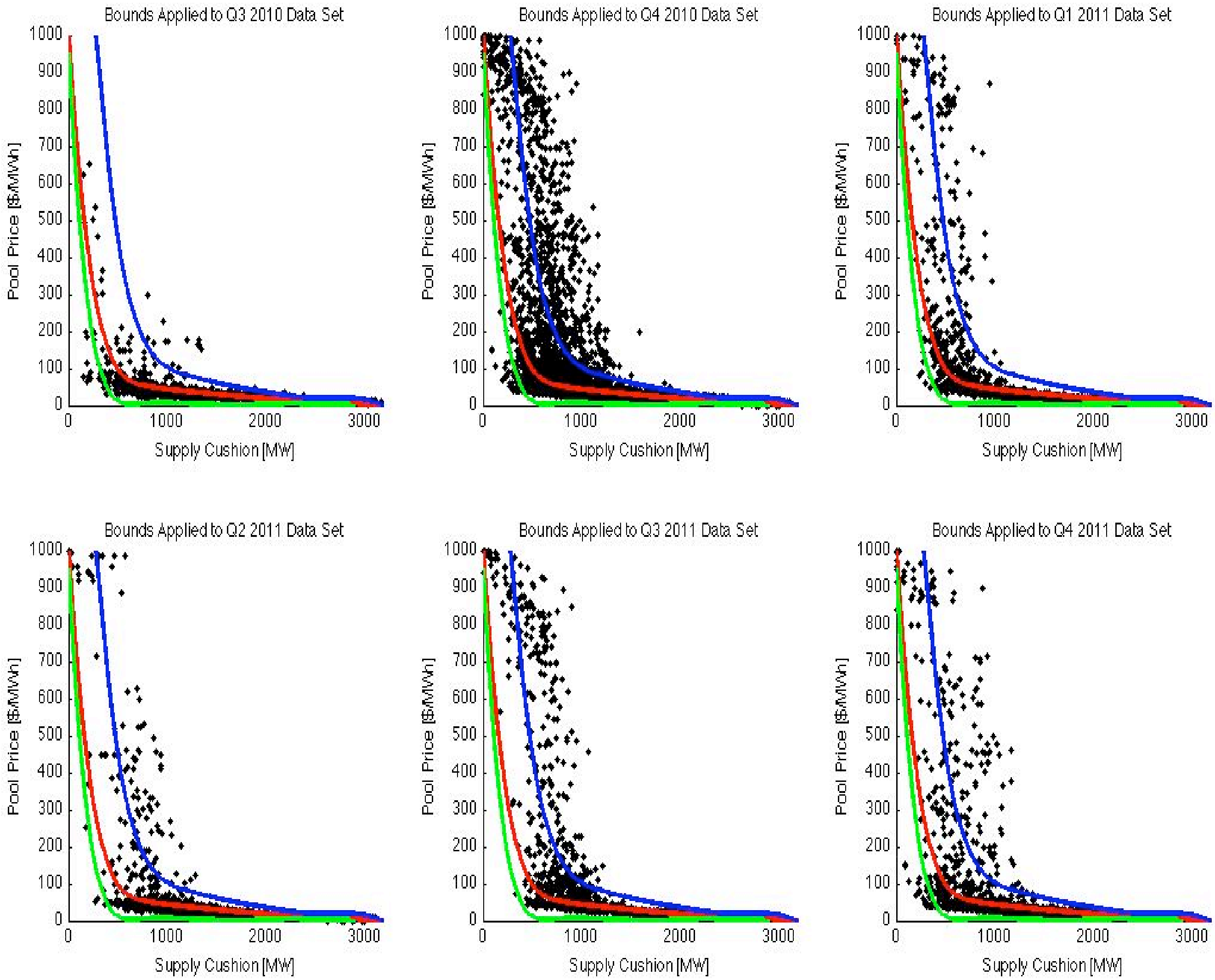


Figure 34: 3-sigma Bounds Applied to Historical Quarterly Data

References

Market Surveillance Administrator

MSA Q3 2010 Report

http://albertamsa.ca/index.php?page=quarterly-reports

Notice RE: Supply Cushion Data and Technical Working Group

http://albertamsa.ca/uploads/pdf/Archive/2012/Notice%20re%20Supply%20Cushion%20Data%20030612-1.pdf

Other

Racine, J.S. (2008), "Nonparametric Econometrics: A Primer," *Foundations and Trends in Econometrics*: Vol. 3: No 1, pp 1-88.

Hardle, W (1990) *Applied Nonparametric Regression,* Cambridge University Press.

Li, Q. and J.S. Racine (2007), *Nonparametric Econometrics: Theory and Practice*, Princeton University Press

The Market Surveillance Administrator is an independent enforcement agency that protects and promotes the fair, efficient and openly competitive operation of Alberta's wholesale electricity markets and its retail electricity and natural gas markets. The MSA also works to ensure that market participants comply with the Alberta Reliability Standards and the Independent System Operator's rules.