Australian Government
Department of Agriculture and Water Resources
ABARES

Measuring water market prices

Statistical methods for interpreting water trade data

Orion Sanders, Neal Hughes and Mihir Gupta

Research by the Australian Bureau of Agricultural and Resource Economics and Sciences

Research Report 19.1
February 2019

Cataloguing data

This publication (and any material sourced from it) should be attributed as:
Sanders, O, Hughes, N, and Gupta, M 2019, *Measuring water market prices*, ABARES research report, report to client prepared for the Department of Agriculture and Water Resources, Canberra, March. CC BY 4.0.
https://doi.org/10.25814/5c5388a92260d

ISBN: 978-1-74323-414-3

ISSN: 1447-8358

Research report no: 19.1

This publication is available at agriculture.gov.au/publications.

Department of Agriculture and Water Resources
GPO Box 858 Canberra ACT 2601
Telephone 1800 900 090
Web agriculture.gov.au

Acknowledgements

The authors thank colleagues from Marsden Jacob Associates, the Bureau of Meteorology and the Department of Agriculture and Water Resources for their input and support in preparing this report.

Contents

Tables

Figures

Summary

Purpose

Data on Australian water trade recorded in state government registers contain significant measurement error or 'noise'. This noise can make it difficult to develop an accurate picture of prevailing market prices for water. Various private and public analysts (e.g., the Bureau of Meteorology, state government agencies, Marsden Jacob Associates, Aither) employ statistical methods to produce water market price estimates from register data. However, to date there has been limited assessment of the performance of these methods.

This report compares different methods for analysing water market price data. The goal is to identify a preferred method (which could be employed by all analysts) for producing accurate and robust estimates of historical and current water market prices, for both water allocations (temporary trade) and water entitlements (permanent trade). The report considers standard methods employed currently by analysts along with more sophisticated methods not currently in use.

Method

A range of statistical methods are tested empirically using Australian water trade data collated by the Bureau of Meteorology (BOM). This testing is undertaken for a variety of different water market types / locations, including those with a high frequency of trade—such as the southern Murray-Darling Basin (MDB)—and those with limited trade. Each method is applied to 27 specific water allocation markets and 21 entitlement markets across Australia for the period 2014-15 to 2016-17.

Four main methods are considered including:

- Discrete price methods which produce price estimates for specific time periods (e.g., months) as currently used by most analysts. In particular:
  - **2SD** – Monthly / quarterly mean price after applying two-standard deviation (2SD) based exclusion of outliers
  - **Median** - Monthly / quarterly median price
- 'Smoothing' methods, which provide continuous (e.g. daily) price estimates via a statistical model. In particular:
  - **LOESS** – Robust local quadratic regression using a LOESS framework (Locally Weighted Scatter-plot Smoother)
  - **GAM** – Robust penalised regression spline using a GAM framework (Generalised Additive Model)

The empirical performance of these methods is assessed using both qualitative (visual inspection of charts) and quantitative (i.e., cross-validated error metrics) evidence. Here good performance requires both accuracy – estimates adequately capture short-run variation in water market price levels – and robustness – estimates are not influenced by obvious outliers in the data. A complete set of results is available via an online dashboard.

Key Findings

The GAM smoothing method achieved the best overall performance

Overall, the GAM method achieves the best performance, with lower error metrics than alternatives in most markets, both for allocation and entitlement trade. The GAM method demonstrates superior accuracy – better capturing short run changes in market prices – and superior robustness – being less influenced by outliers.

The benefits of smoothing methods (e.g., GAM) depend on the type of market

Generally, the benefits of smoothing methods (i.e., increases in accuracy and robustness) were greater in entitlement markets than for allocations. For allocation markets, GAM achieves significant improvements (over 2SD and Median) within the sMDB, given the high frequency of trade and clear price signals. In these markets, smoothing methods provide a more accurate estimate of current market prices, accounting for changes in the most recent weeks and days of trading which can be missed by discrete methods. For allocation markets outside of the sMDB all methods achieve similar levels of performance.

Smoothing methods can struggle during market closures

Smoothing methods (including GAM) can produce unreliable results during brief periods of market inactivity in otherwise high frequency markets, such as when trading is temporally suspended by regulators. For this reason, smoothing methods need to be combined with a filter rule, to ensure that no prices are reported on days with no or limited trade activity.

Traditional methods (e.g., 2SD) performed the worst overall

2SD was outperformed by other methods in most allocation and entitlement markets. In some markets especially those with infrequent trade, 2SD can be influenced by low outliers. As a result, 2SD produces lower estimates of market price on average relative to other methods. In testing, median proved to be more robust, consistently outperforming 2SD. Monthly median prices worked best in allocation markets and quarterly prices were preferred in entitlement markets.

Next steps

At a minimum Median should be adopted in place of 2SD

In testing, Median consistently outperformed 2SD. Median is also easy for analysts to implement and easily understandable by stakeholders. Monthly median prices for allocation markets and quarterly median prices for entitlement markets provided a reasonable starting point when analysing water register data.

Effort could be put into further developing smoothing methods like GAM

While the GAM method achieved superior performance, it remains harder to implement and harder to explain to stakeholders than simple statistics like median. For these reasons, many analysts may be reluctant to adopt GAM. Further effort may be required to address these concerns. In particular, common software could be developed to fully implement the GAM method, including an appropriate filter rule. This software could then be used by all analysts with minimal effort and minimal risk of implementation error.

Effort could be put into improving water trade data at the source

Beyond the use of statistical methods, consideration should be given to improving the quality of government water register data at the source. There may also be opportunities to make greater use of alternative data sources, including price data recorded by private water exchanges and brokers.

Introduction

For water markets to function well, market participants need to be able to access timely, accurate and reliable information on prevailing market prices. For some time, concerns have been raised about the availability and quality of information on Australian water market prices available to participants.

These concerns have motivated a number of commonwealth responses including the development of the former National Water Market System (the NWMS), the current Bureau of Metrology (BoM) Water Market Information dashboard, and the ongoing Business Research and Innovation Initiative (BRII) project. In addition to these Commonwealth services, water price information is made available to market participants via state government websites, and via a range of private entities, including water brokers, exchanges and consultants.

The vast majority of the data presented in these platforms is obtained from same source: state government water trading registers. Data held in these registers, particularly the prices of recorded water trade transactions, are widely acknowledged to contain significant 'noise' including frequent outliers. As a result, some level of statistical analysis is required to generate reliable estimates of prevailing water market prices from raw register data.

Water market analysts (both government and non-government) have employed a variety of methods to produce market price statistics. However, to date there has been limited formal analysis of these methods or consideration of potential alternatives.

This report compares different statistical methods for analysing water market price data. The goal of this study is to identify a preferred method for deriving accurate and robust estimates of market prices on a daily basis, for both water allocations (temporary trade) and water entitlements (permanent trade).

The report considers both standard methods employed currently by analysts (discreet methods which estimate monthly or quarterly prices) along with some more sophisticated methods not currently in use (smoothing methods which produce continuous / daily price estimates). These methods are tested empirically using Australian water market data collated by the Bureau of Meteorology (BOM). This testing is undertaken for a variety of different water market types / locations, including those with a high frequency of trade—such as the southern Murray-Darling Basin (MDB)—and those with limited trade.

The report focuses solely on the application of statistical methods to existing water market data sources. Improving the collection, validation and auditing of water market data at the source (e.g. within state water registers) was considered out of scope.

This report begins by discussing the challenges in deriving water market price estimates from existing water register data (chapter 2). Next the alternative statistical methods are reviewed and compared (chapter 3) and then empirically tested against water market data (chapter 4).

1 Data issues in water markets

Within Australian water markets there are two main sources of price data: water exchanges/brokers and water registers.

- Water registers are mostly maintained by state governments, with some maintained by irrigation infrastructure operators (IIO). Water registers have evolved as a natural extension of land title registers. A water register's primary purpose is to record the ownership of water rights including any transfers of ownership arising from trade. While water registers contain data on prices associated with water trading, these prices are self-reported by market participants and are subjected to limited auditing.
- Water exchanges provide the role of a 'clearing house' matching buyers and sellers, not dissimilar to stock exchanges such as the Australian Stock Exchange (ASX). The advantage of water exchanges as a data source, is that actual prices of water trades are recorded electronically in near real-time. Brokers also maintain records of transactions which are sometimes made publicly available (for example, via bulletin boards). However, exchanges and broker datasets also have some disadvantages: there are many exchanges and brokers and each only records a sub-set of all transactions; they have no obligation to provide price information publically; they are private entities and there is no government regulation of the industry.

For the remainder of this report we focus on register data collated by the Bureau of Meteorology (BOM). In this section, we consider two main challenges with analysing price data: noise and market classification.

1.1 Noise

A fundamental concept in economics, is that a competitive market should arrive on a single equilibrium price. Many Australian water markets, particularly water allocation markets in the southern MDB, satisfy the economic notion of a competitive market: water is a homogenous good, water markets involve large numbers of buyers and sellers and are subject to relatively low transaction costs. While small differences in prices might be expected due to transaction costs, information problems and other market imperfections, it is reasonable to expect prices for water trades to be similar, at least for a given time period, location (i.e. water trading zone), water product (i.e. allocation / entitlement type) and transaction type (i.e. spot/forward/lease etc.).

In practice, water register data is typically highly dispersed (even for a given time period, location or product) due to measurement error / noise. Noise in water register data is readily apparent in Figure 1 and Figure 2 for the southern MDB. In these figures, prices above $1000/ML for allocation markets and above $5000/ML for entitlement markets, have been excluded.

Figure 1 Allocation transactions, southern connected regulated surface water

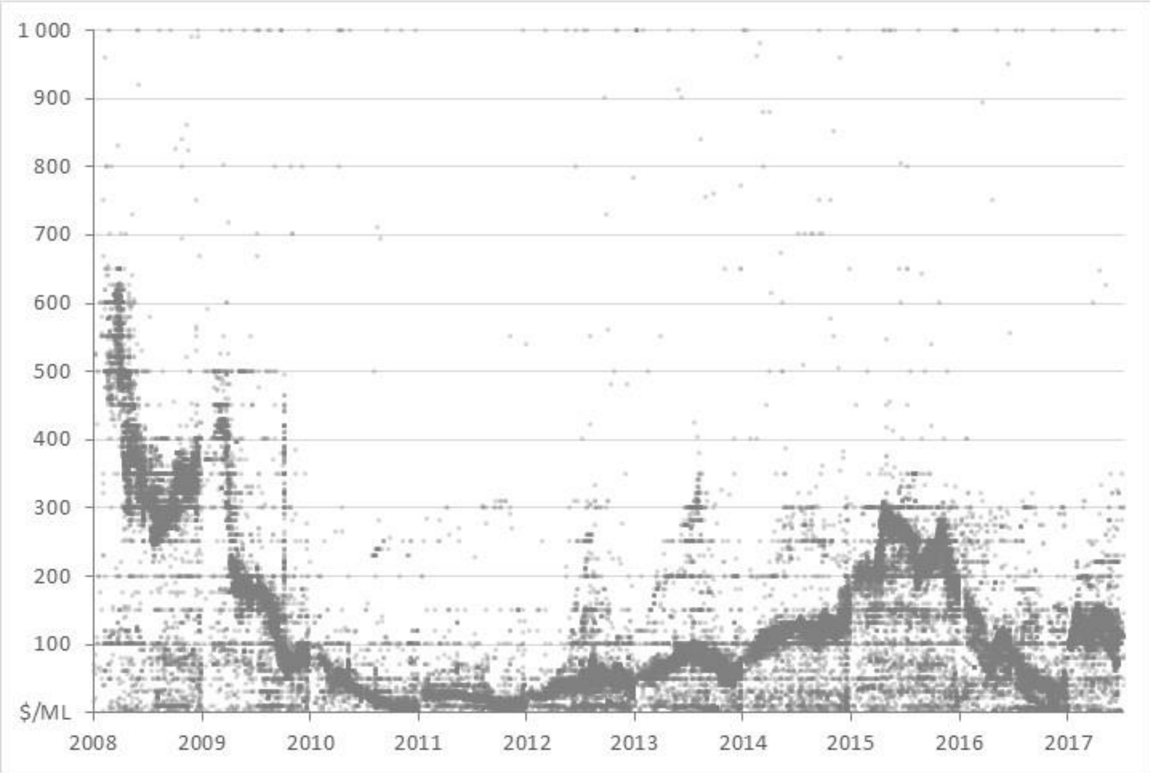
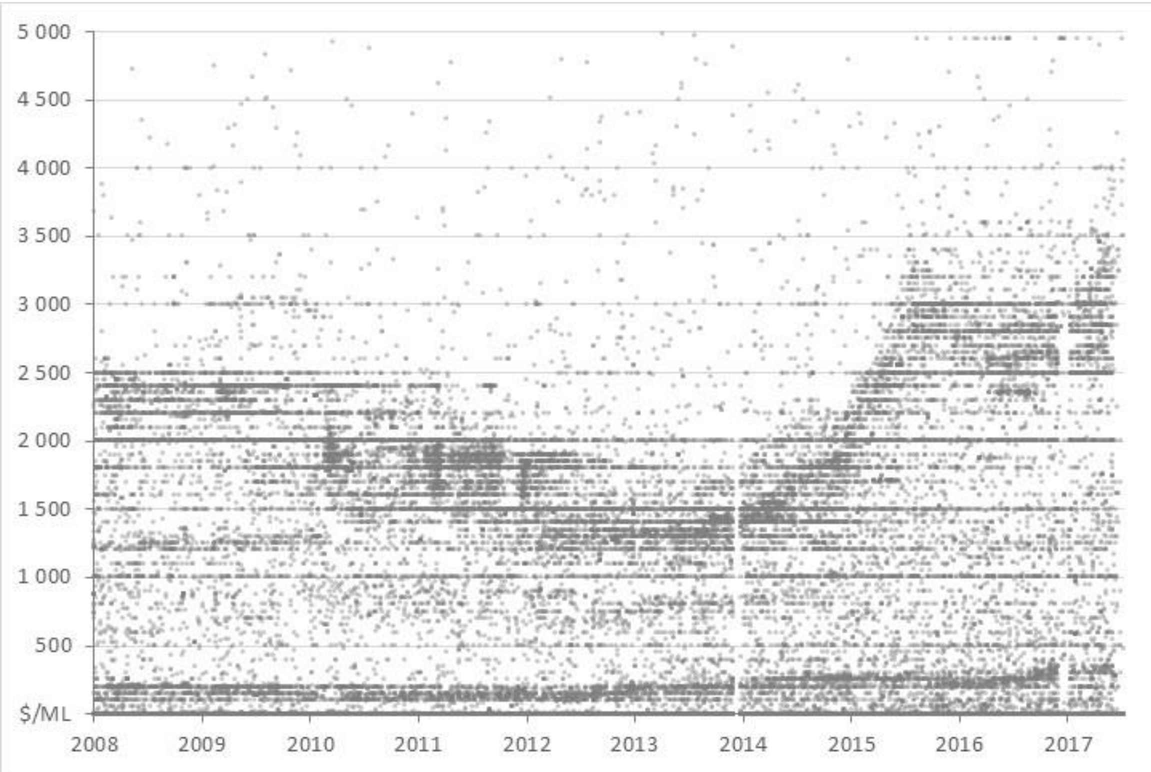


Figure 2 Entitlement transactions, southern connected regulated surface water



While these figures show clear market price 'signals' around which large numbers of transactions are clustered, there remains much noise. There are more outliers below the market price than outliers above the market price, but those outliers above can be extreme (up to almost $1 million per megalitre). There are also a very large number of unreported or $0 prices, and a number of unrealistically low price trades from $0.01 to $10.

There are a range of potential sources for noise in register data, including:

- Incorrect price data entry: transposition, mislabelling of price and price per megalitre, rounding errors.
- Misleading / false reporting (including non-reported prices and $1 per megalitre trades etc.).
- Incorrect date captured: date of registration rather than transaction / settlement date (particularly a problem for entitlement trades)
- Duplicated transactions
- Mislabelled regions or product types
- Inclusion on water registers of 'non-market' transactions:
  - gifts and related-party transfers;
  - environmental allocation transfers and government water recovery infrastructure projects;
- No separation of heterogeneous transaction types:
  - forward contracts
  - carryover 'parking' (trade in carryover capacity)
  - 'wet' / 'dry' water entitlements (entitlements with or without current annual allocations attached).
  - heterogeneous reliability classes (supplementary / other), and
  - combined land and water sales.
- Prices incorrectly distributed between individual products for combined (multi-product) transactions.

1.2 Market classification

Across Australia there are a large number of mostly disconnected water markets. While a single market price for water is expected to emerge within each of these markets, differences in prices are to be expected between markets, given differences in climate conditions and water availability between regions along with differences in the types of water products being traded.

One key exception is the southern Murray-Darling Basin (southern MDB) where significant inter-region allocation trade occurs. When inter-regional allocation trade is permitted between locations (i.e., trading zones) these regions can be viewed as a single market, with a single equilibrium price. However, within the southern MDB, limits on trade between regions are frequently imposed, temporarily disconnecting trading zones and leading to differences in prices (see ABARES 2018).

For simplicity, we assume that each allocation trading zone is a separate market throughout this report. In practice, a more accurate estimate of price could be obtained by pooling data from connected trading zones, this remains a potential subject for future research.

Water allocation markets can be classified by the following characteristics:

- Location (water trading zone)
- Resource type (surface / ground water)

For water entitlement markets, additional characteristics are required:

- Reliability type (e.g. high / low reliability)
- Regulated / unregulated

Based on this classification, there were 146 allocation markets and 718 entitlement markets in which water trading activity was recorded in Australia during 2016–17. However, the majority of these 'markets' contained very few transactions, while others (such as those in Queensland) have no price data recorded.

In this report, we limit our attention to allocation markets with at least 40 priced transactions (Table 1) and entitlement markets with more than 15 priced transactions (Table 2). This leaves 27 allocation markets and 21 entitlement markets. Even then, a majority of trading activity occurs in a small number of very active markets particularly in the southern MDB. Entitlement markets are noticeably less active than allocation markets, in terms of number of transactions per year.

Table 1 Active allocation markets

Water system	Resource type	Regulated	Transactions
7 VIC Murray - Barmah to SA	Surface water	Regulated	2748
1A Greater Goulburn	Surface water	Regulated	2317
Murray Irrigation	Surface water	Regulated	2001
New South Wales Murray Regulated River Water Source / that part of the water source downstream of the River Murray at Picnic Point	Surface water	Regulated	739
6 VIC Murray - Dart to Barmah	Surface water	Regulated	650
Murrumbidgee Regulated River Water Source	Surface water	Regulated	615
River Murray Prescribed Watercourse	Surface water	Regulated	614
Lachlan Regulated River Source / that part of the water source upstream of Lake Cargelligo Weir	Surface water	Regulated	183
Macquarie And Cudgegong Regulated Rivers Water Source / that part of the water source downstream of the upper limit of Lake Burrendong	Surface water	Regulated	173
New South Wales Murray Regulated River Water Source / that part of the water source upstream of the River Murray at Picnic Point	Surface water	Regulated	135
Harvey Water-Harvey	Surface water	Regulated	124
6B Lower Broken Creek	Surface water	Regulated	111
4A Campaspe - Eppalock to WWC	Surface water	Regulated	99

Water system	Resource type	Regulated	Transactions
3 Lower Goulburn	Surface water	Regulated	83
5A Loddon - Coliban Channel to Lower Weir Pool	Surface water	Regulated	82
Lower Murrumbidgee Deep Groundwater Source	Groundwater	Unregulated	80
Border Rivers Regulated River Water Source	Surface water	Regulated	74
Coleambally Irrigation Co-op Ltd	Surface water	Regulated	69
41 Macalister	Surface water	Regulated	64
Upper Namoi Regulated River Source	Surface water	Regulated	63
1B Boort	Surface water	Regulated	60
Gwydir Regulated River Water Source	Surface water	Regulated	58
Macquarie And Cudgegong Regulated Rivers Water Source / that part of the water source upstream of the upper limit of Lake Burrendong	Surface water	Regulated	47
Lachlan Regulated River Source / that part of the water source downstream of Lake Cargelligo Weir	Surface water	Regulated	44
Lower Darling Regulated Rivers Water Source	Surface water	Regulated	44
Lower Namoi Regulated River Source	Surface water	Regulated	44
Lower Murray Groundwater Source	Groundwater	Unregulated	42

Table 2 Active entitlement markets

Water system	Resource type	Regulated	Reliability	Transactions
7 VIC Murray - Barmah to SA	Surface water	Regulated	High	604
1A Greater Goulburn	Surface water	Regulated	High	564
1A Greater Goulburn	Surface water	Regulated	Low	247
6 VIC Murray - Dart to Barmah	Surface water	Regulated	High	159
River Murray Prescribed Watercourse	Surface water	Regulated	General	131
7 VIC Murray - Barmah to SA	Surface water	Regulated	Low	94
New South Wales Murray Regulated River Water Source	Surface water	Regulated	General	94
6 VIC Murray - Dart to Barmah	Surface water	Regulated	Low	69
New South Wales Murray Regulated River Water Source	Surface water	Regulated	High	63
Murrumbidgee Regulated River Water Source	Surface water	Regulated	General	54
Macquarie And Cudgegong Regulated Rivers Water Source	Surface water	Regulated	General	47
Lachlan Regulated River Source	Surface water	Regulated	General	43
Border Rivers Regulated River Water Source	Surface water	Regulated	General	30
41 Macalister	Surface water	Regulated	High	29
41 Macalister	Surface water	Regulated	Low	24
Lower Namoi Regulated River Source	Surface water	Regulated	General	22
River Murray Prescribed Watercourse	Surface water	Regulated	High	20

Water system	Resource type	Regulated	Reliability	Transactions
6B Lower Broken Creek	Surface water	Regulated	High	19
Border Rivers Regulated River Water Source	Surface water	Regulated	Other (Supplementary Water)	17
Lachlan Fold Belt Mdb Groundwater Source	Groundwater	Unregulated	Aquifer	17
Lower Namoi Groundwater Source	Groundwater	Unregulated	Aquifer	16

The above water markets are highly diverse, varying both in their activity volume (number of observations) and level of measurement error. This variation in sample size and noise creates some challenges for developing an appropriate statistical methods to 'clean' the data. In general, a more accurate price signal will be possible in markets with a higher trade volumes and lower levels of noise.

2 Statistical methods

2.1 Discrete methods

Currently, most analysis of water market data involves the estimation of price statistics for discrete time windows, such as the annual or monthly average price. With this approach, emphasis is placed on data cleaning: identifying outlier observations, excluding them from the analysis and using the remaining data to compute average prices for each time period.

Trimming

Typically, any analysis of register water price data involves an initial 'trimming' procedure to exclude transactions above or below predetermined thresholds. These thresholds are chosen based on market knowledge of normal price movements and allow obvious extreme outliers to be quickly removed.

Zero dollar transactions are most easily excluded and comprise almost half of all allocation and entitlement transactions. The significant number of reported $1 per megalitre transactions are likely to relate to 'default' price submissions or related-party transfers more than actual market trades and so are also typically excluded.

In all of the subsequent analysis in this report water register data is first trimmed. For allocation markets, observations are only included where the price was between $1 and $5,000 per ML. For entitlement markets, observations are only included where price was between $1 and $10,000 per ML.

Two-standard deviation exclusion

Given water prices vary dramatically across and within years, simple trimming based on fixed thresholds is insufficient to remove all outliers. A commonly employed next step, is to use specific exclusion thresholds for each time period of interest. Typically, this has involved excluding trades with prices above and below two standard deviations from the mean on an annual, quarterly or monthly basis.

The use of monthly or quarterly two-standard deviation exclusion is one of the most commonly employed methods to date. However the approach remains subject to some obvious limitations. Firstly, mean and standard deviation are themselves not robust statistics (they are heavily affected by outliers). Further, two standard deviation exclusion is commonly applied when the error is assumed to be normally distributed. However, water market register data is commonly not symmetric and is skewed towards lower observations. As such, this method could potentially retain low outliers. Finally, in cases where water market prices change significantly within a given period (month or quarter) there is a risk the method could exclude legitimate transactions (see Figure 3, Figure 4, Figure 5).

Median

The field of robust statistics is concerned with methods which emulate more common statistical measures but which are more robust to outliers. For example, median is considered a more robust statistic than mean. Median price estimates are reported by some analysts (including for example the Victorian water register). The use of median rather than mean avoids the need to eliminate outliers, however it still relies on 'windowing' the data to a particular time-step.

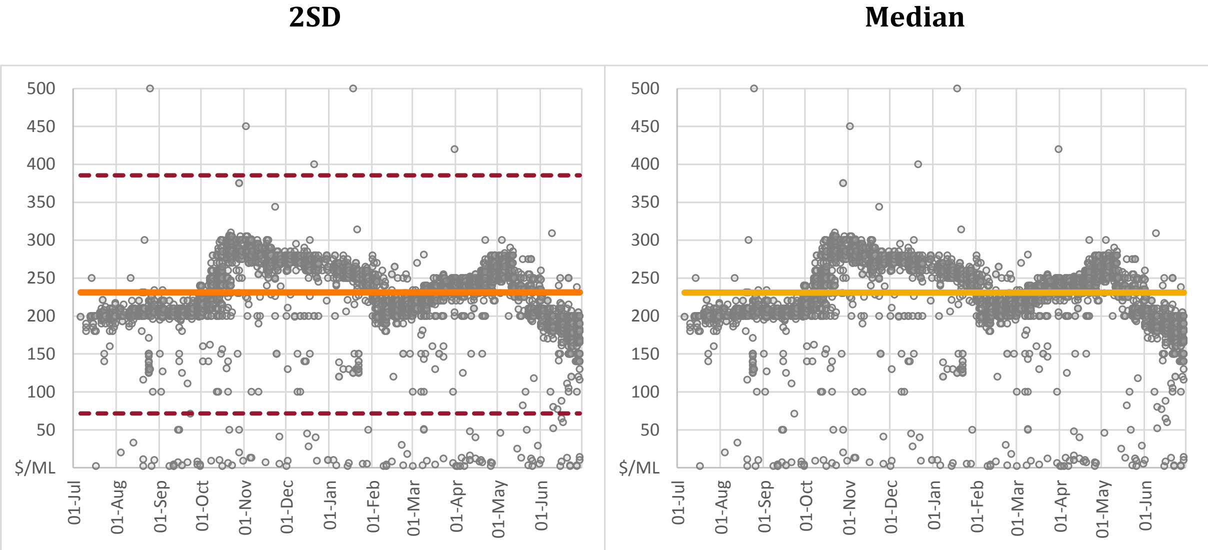
The bias-variance trade-off

Beyond criticisms of robustness, all discrete methods involve a difficult problem of selecting the appropriate time step to use. If the window is too large (for instance, a year), the single statistic fails to adequately convey the variation and trends within the window, leading to biased results (Figure 3). Conversely, if the window is too small, the statistic can easily suffer from excessive variance (Figure 5).

This is a classic example of the standard 'bias-variance' trade off. Here annual statistics (Figure 3) 'underfit' the data resulting in bias, while daily statistics (Figure 5) 'overfit' the data leading to high variance (noise). In practice, the appropriate window size will depend greatly on the market type, especially the frequency of trade activity. In the subsequent analysis, monthly and quarterly time steps are considered, which are both commonly employed by analysts.

Another related concern with discrete methods is that they can introduce 'end point bias'. That is, if prices increase or decrease significantly within a given time period, reported (average) prices will significantly under or overestimate observed values at the end and or start of the time period (see Figure 3). This is important given that the end point (the most recently observed price) is of high importance to market participants.

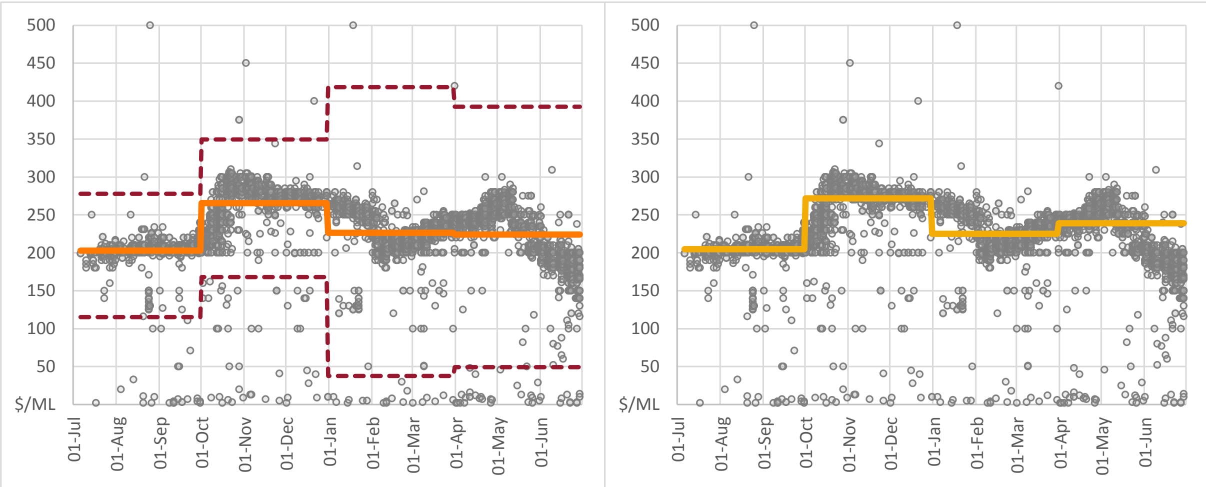
Figure 3 Two standard deviation adjusted mean (2SD) and median, using an annual window (1A Greater Goulburn 2015–16)



Note: The dotted red lines represent two standard deviations from the mean. The y-axis on these charts has been limited to show the majority of observations (there are numerous observations outside this limit). The mean estimate is presented after removing outliers determined by the 2SD method.

Figure 4 Two standard deviation adjusted mean (2SD) and median, using a quarterly window (1A Greater Goulburn 2015–16)

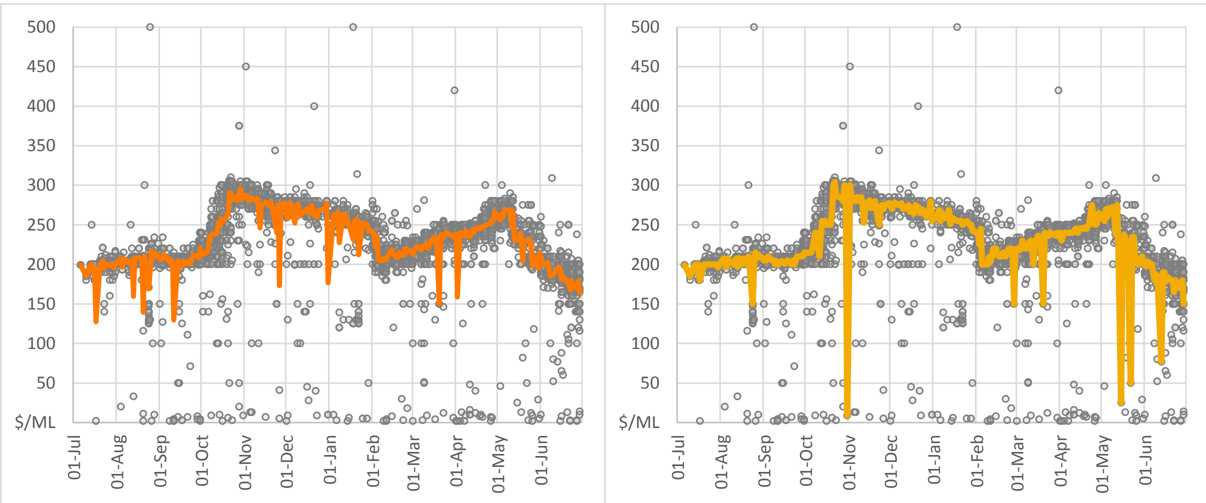
2SD **Median**



Note: The dotted red lines represent two standard deviations from the mean. The y-axis on these charts has been limited to show the majority of observations (there are numerous observations outside this limit). The mean estimate is presented after removing outliers determined by the 2SD method.

Figure 5 Two standard deviation adjusted mean and median, using a daily window (1A Greater Goulburn 2016–17)

2SD **Median**



Note: The dotted red lines represent two standard deviations from the mean. The y-axis on these charts has been limited to show the majority of observations (there are numerous observations outside this limit). The mean estimate is presented after removing outliers determined by the 2SD method.

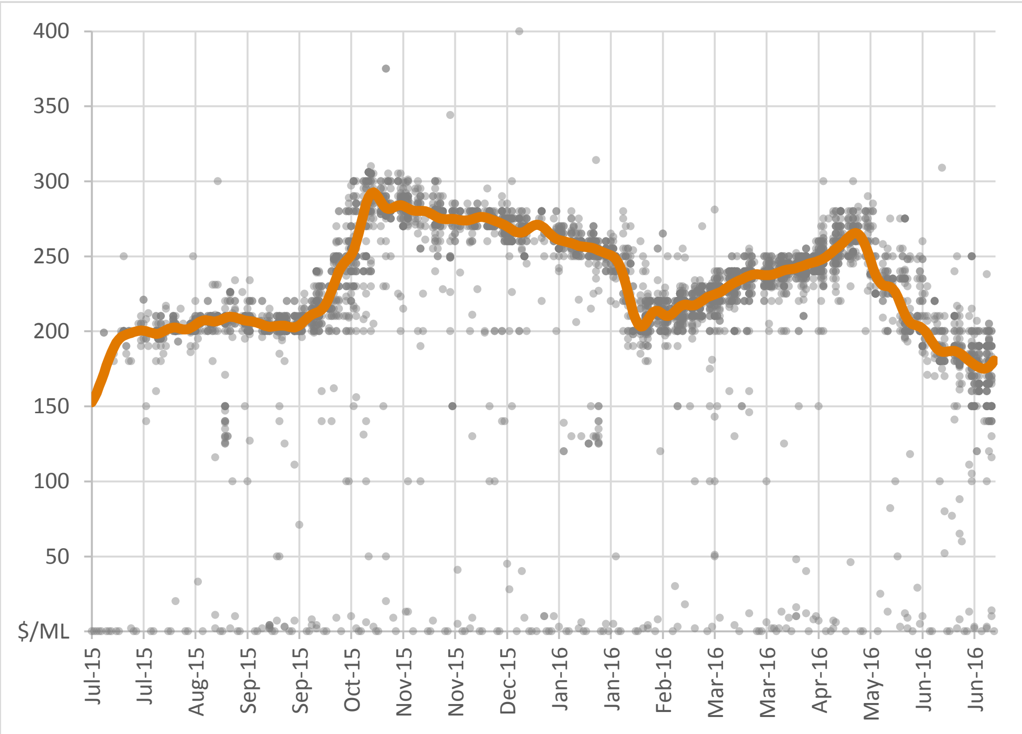
2.2 Smoothing methods

To date, analysis of water market price data has focused almost exclusively on discrete methods, where price statistics are reported over a specified time window. This approach has a number of limitations as detailed in the previous section. While the use of robust statistical measures (such as median) may represent an improvement over more common methods (mean and univariate outlier detection) this still involves the use of static time windows.

Ideally we would like to relax these constraints and identify more continuous methods for data analysis, capable of providing accurate and robust price statistics on a more frequent (i.e., daily basis). Fortunately, there are many well researched methods that fall under the general banner of *statistical smoothing*.

Statistical smoothing—similarly, non-parametric regression or curve fitting—is a banner for a range of methods that attempt to approximate a dataset by a smooth function that best fits the underlying patterns in the data. That is, instead of estimating an average price for a given time period, the problem is recast as fitting a continuous curve to the data which captures changes in price over time (as shown in Figure 6).

Figure 6 Example of a statistical smoothing applied to water allocation price data (1A Greater Goulburn 2015–16)



More formally the smoothing task can be viewed as single variable regression problem:

$$y_{ij} = f\left(x_{ij}\right) + e_{ij}$$

Where:

- y_{ij} is the recorded price of transaction *i* in water market (i.e., region) *j*
- x_{ij} is the recorded date of transaction *i* in water market *j*
- e_{ij} is a residual (noise) term

Here the price variable *y* is continuous, while the date *x* is discrete, as it is reported to the nearest day. It is common to have both multiple observations per time period, and data gaps; that is, time periods with no price observation.

Potentially other data beyond just transaction date could be used in this type of analysis. One option would be to explore interactions between markets, such as between different entitlement reliabilities within the same region, or between connected water allocation trading zones. Beyond this, it might be possible to combine register data with other data sets on water supply and demand to develop more detailed economic models of prices. These options are left as potential subjects for future research.

Three broad class of smoothing methods relevant to this problem are: nearest neighbour methods, local regression, and regression splines. Nearest neighbour methods – such as moving averages or moving medians – are a natural extension of the discrete methods described above. However, these methods are unlikely to offer much performance improvement over discrete methods in practice. In particular, they remain subject to end point bias and would offer no improvement in estimating current water market prices.

Within each of these classes, the possibilities available for fitting and ensuring robustness are very large. For tractability, it is necessary to identify the key parameters that influence behaviour before detailed evaluation of these methods against market data can be undertaken. In addition, to satisfy ease of use and reproducibility, focus is limited to accessible implementations of these methods in the open source statistical software *R*. Below we summarise the local regression and penalised regression splines approaches and for each define a specific candidate method using *R*.

Local regression

The idea behind local regression is to non-parametrically estimate a smoothing function by a locally weighted regression at each point of interest. The most common approach is known as LOESS or locally weighted scatterplot smoothing (see Cleveland, Grosse and Shyu 1992). LOESS builds on standard regression methods such as Ordinary Least Squares (OLS) regression. LOESS involves fitting a simple linear or low-degree polynomial model by OLS, multiple times for localised subsets of data, in order to produce a more complex non-linear fitted curve.

Robust variants of these methods are also available, including:

- estimating by robust M or S-estimators rather than the sum of squared residuals, or
- estimating via quantile (median) regression, or
- using iterative reweighted least squares

Key defining characteristics of methods within this class are the functional form of the local regression (linear, quadratic etc.), the estimation method, and smoothing parameters (which determine the size of the local data subsets and the weights applied to each sample point).

Given the wide range of options available, some initial testing was undertaken of various local regression packages available in the *R* software environment to arrive on single candidate approach. Based on this testing, the *R stats::loess* package was selected with the following set-up:

- Robust fitting via iteratively reweighted least squares (*family = 'symmetric'*)
- Smoothing parameter selection via k-folds cross-validation (with k = 2 and sum of absolute residuals '*sar*' as the evaluation metric)
- Quadratic polynomial model (*degree = 2*)

Penalized regression splines

Penalized regression splines are another powerful smoothing method for fitting arbitrary data. They are commonly implemented as generic smoothers for a wider variety of statistical models, the General Additive Model (GAM) (see Hastie 1992). Most significantly, penalised regression splines allow freedom to model error terms more flexibly with non-normal distribution families.

A penalized regression spline model involves a data model:

$$y = \sum_{k=1}^{n} w_k \varphi_k(x) + e$$

where:

e is an error term with a given distribution family

φ_k are a set of n spline basis functions

w_k are a set of weights (parameters) to be estimated

Fitting a regression spline model (that is, estimation of the weights w_i) involves minimising both a measure of fit and a penalisation term which rewards 'smoothness'. This approach usually requires a smoothing parameter to be calibrated. While it is clear that there are numerous implementation details that could be adjusted, the key choices are:

- Distribution family (of the error term e)
- The type and number spline basis functions
- Method for smoothing parameter selection

Experimentation with spline regression methods led to the *R gam::mgcv* implementation as the preferred candidate. Key set-up details include:

- Number of basis spline 'knots' *k = min(unique(x), 200)* where *unique(x)* is the number of dates contained in the data set.
- Thin plate basis splines (*bs='ts'*)
- Scaled t-distribution (*family = scat()*)
- Restricted maximum likelihood estimation (*method='REML'*)

3 Empirical analysis

3.1 Method

Four candidate methods were evaluated based on the discussion in Chapters 3:

- **LOESS** - Robust local quadratic regression
- **GAM** - Penalised regression spline (with *t* error distribution)
- **2SD** - Monthly / quarterly mean with two-standard deviation exclusion
- **Median** - Monthly / quarterly median

These four methods were used to estimate daily prices for each of the allocation and entitlement markets listed in Table 1 and Table 2 over the period 2014-15 to 2016-17. Prior to applying the above methods, a simple min-max cut-off was applied to the water price data, to exclude obvious outliers. For allocation markets, observations are only included where price, was between $1 and $5,000 per ML. For entitlement markets, observations are only included where price was between $1 and $10,000 per ML.

The above methods are evaluated both qualitatively (via visual inspection of charts) and quantitatively (based on how well they fit register data). The key challenge with quantitatively evaluating the models using register data, is that the data are known to contain errors and outliers. Standard measures of model fit such as squared error impose a high penalty on large residuals, which in our case are mostly likely to be caused by erroneous outliers in the data rather than poor model performance. To address this we consider a number of robust performance metrics, which place lower weights on large residuals (listed below).

These performance indicators are measured using a 10-fold cross validation ('out-of-sample' testing) procedure: where each model is repeatedly fit to sub-samples of the data and evaluated on withheld data. Cross validation helps address the problem of 'over-fitting', where the predicted prices are highly variable and sensitive to noise in the data. Methods which overfit tend to have good in-sample performance but poor out-of-sample performance.

Performance metrics

Mean Absolute Error (MAE)

The mean absolute error is a measure of the absolute difference between the predicted and observed data. It is defined by:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

where y_i denotes the predicted value and x_i denotes the observed data.

Median Absolute Error (MEDAE)

The median deviation is defined as the median of the absolute difference between predicted and observed data.

$$Median\ deviation = Median(|y_i - x_i|)$$

Symmetric Mean Absolute Percentage Error (SMAPE)

This is an accuracy measure based on the relative percentage error between predicted and observed data. It is easy to interpret, having a value between 0 and 100 per cent. It is defined by:

$$SMAPE = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|y_i - x_i|}{|y_i| + |x_i|}$$

3.2 Results

A complete set of results for each allocation and entitlement market and each method is available via an online dashboard. A summary of the results is provided below.

Allocation markets

Table 3 shows the number of markets where a price cleaning method is preferred according to each performance metric while Figure 7 shows the average prediction error for each price cleaning method (here lower values indicate better performance) across all allocation markets. Example charts are also presented comparing the daily predictions for each method for the Goulburn (Figure 8), Lower Namoi (Figure 9) and Boort (Figure 10) markets.

Table 3 Number of allocation markets where each method is ranked highest by performance metric

Method	MAE	MEDAE	SMAPE
Monthly median price	3	2	4
2SD: Monthly average price	2	2	1
GAM	19	12	14
LOESS	1	9	6

Figure 7 Prediction error by method and metric, average across all allocation markets

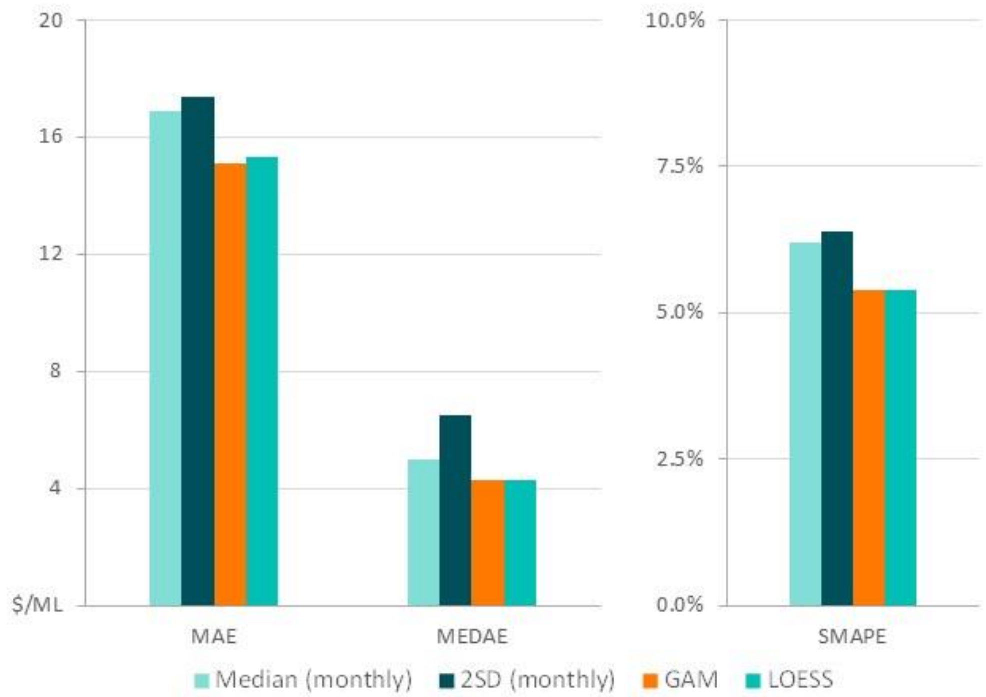


Figure 8 Price predictions by method, 1A Greater Goulburn, surface water allocation market, 2014–15 to 2016–17

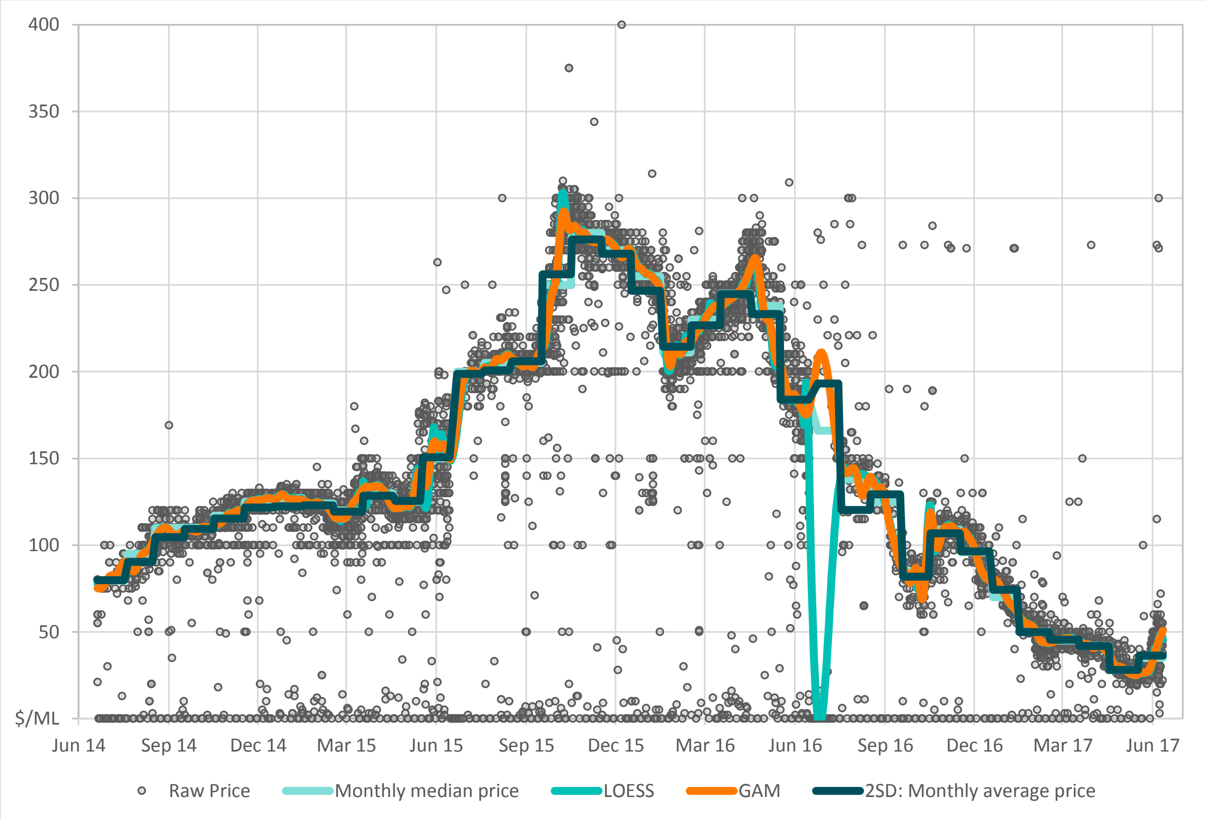


Figure 9 Price predictions by method, Lower Namoi, surface water allocation market, 2014–15 to 2016–17

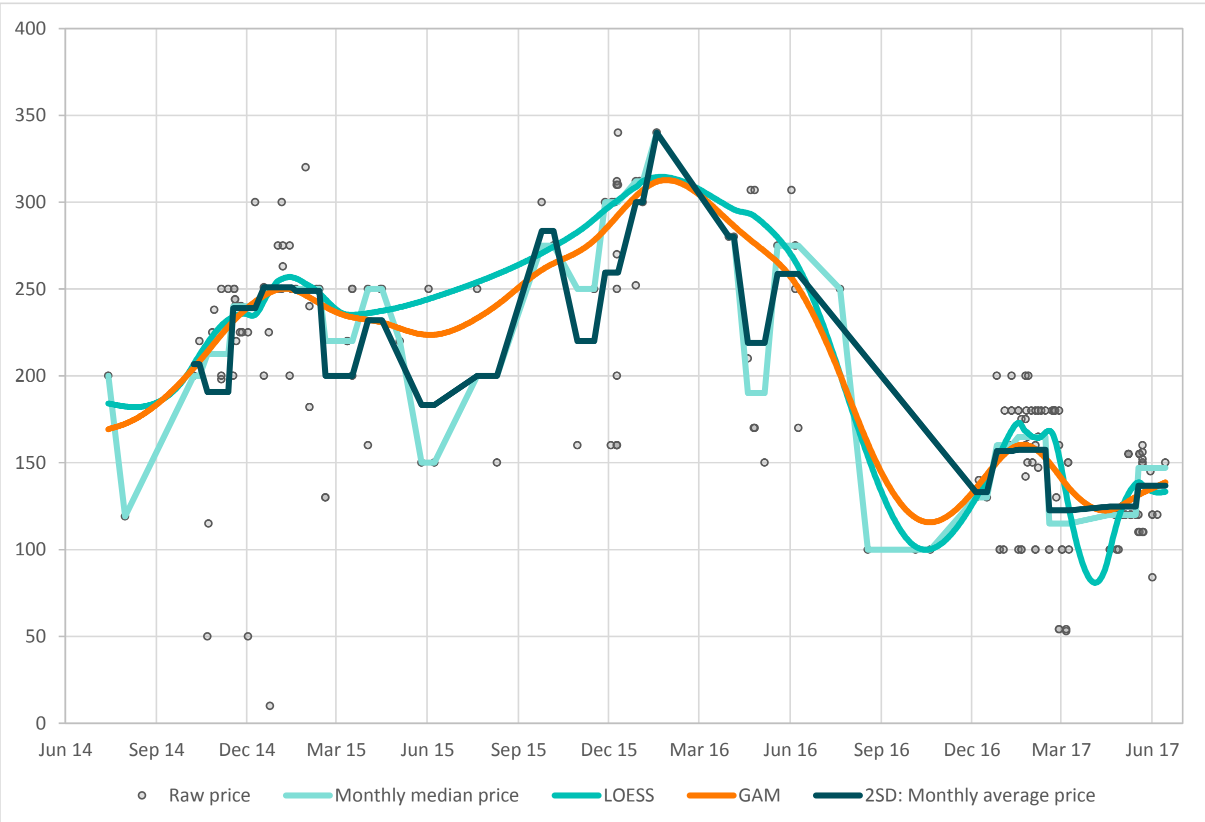
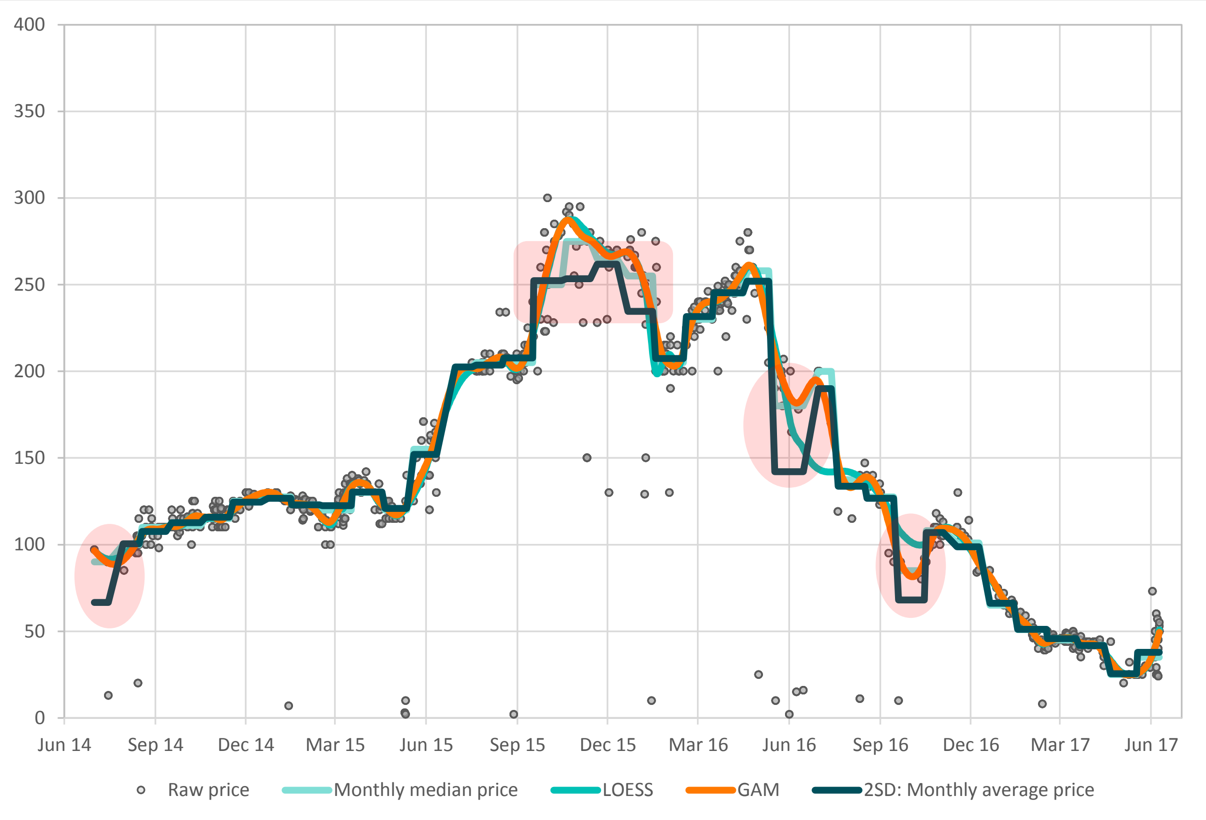


Figure 10 Price predictions by method, 1B Boort, surface water allocation market, 2014–15 to 2016–17



For allocation markets, the GAM method provides the most robust results overall. In aggregate across all allocation markets the GAM method was ranked the highest preforming method by each metric (Figure 7). While all methods were preferred in at least some allocation markets, GAM was preferred in the majority of markets for each metric (Table 3).

Smoothing methods (LOESS and GAM) achieved larger performance gains relative to traditional measures (median and 2SD) within the sMDB. Outside of the sMDB smoothing methods offer less of an improvement, with GAM, LOESS and median often achieving similar performance scores. This is not surprising given the lower frequency of trade outside of the sMDB. In high frequency markets, LOESS and GAM produce precise results capturing changes in price on a weekly and daily basis (Figure 8). For example, at the end of the 2016-17 water year, there is a rapid increase in price in the final weeks of trading which is captured more accurately by smoothing methods than by monthly methods (Figure 8). In contrast, when less data is available GAM and LOESS (by design) produce a smoother price series capturing only longer-term changes in price and therefore offer less of an improvement over monthly prices (see Figure 9).

Inspection of charts suggests that the results produced by GAM are robust to noise and outliers. In contrast, the 2SD method can be influenced by outliers in some cases, particularity lower outliers in lower frequency markets. As a result, 2SD tends to produce lower average estimates of allocation prices overall. This is particularly illustrated by the allocation market in the 1B Boort trading zone. As shown in Figure 10, there are numerous periods (marked by red circles) when 2SD is affected by low prices that appear to be outliers, leading to lower estimates of market prices compared with other methods.

Comparing the smoothing methods, while the LOESS method performs well on most performance measures, visual results suggest the method is slightly less robust than GAM and is still somewhat influenced by outliers. One weakness of both smoothing methods is that they can produce poor results during brief periods of limited or no trade activity within high frequency markets (as seen during the period of July 2016 in the Goulburn – see Figure 8). However, this problem can be addressed by combining the method with a filter rule, which suppresses estimated prices when trade activity is insufficient (Figure 11).

Two example filter rules that accept a wide range of input values are included in the online dashboard. The first is a sample size rule, which imposes a minimum number of trades over a specified time period (as demonstrated in Figure 11). The second makes use of the error-bounds estimated by the GAM model, suppressing fitted prices when the bounds are too large. In practice, filter rules may not be required if there is external knowledge available on temporary water market closures.

Figure 11 Price predictions by method, applying filtering, 1A Greater Goulburn, surface water allocation market, 2014–15 to 2016–17

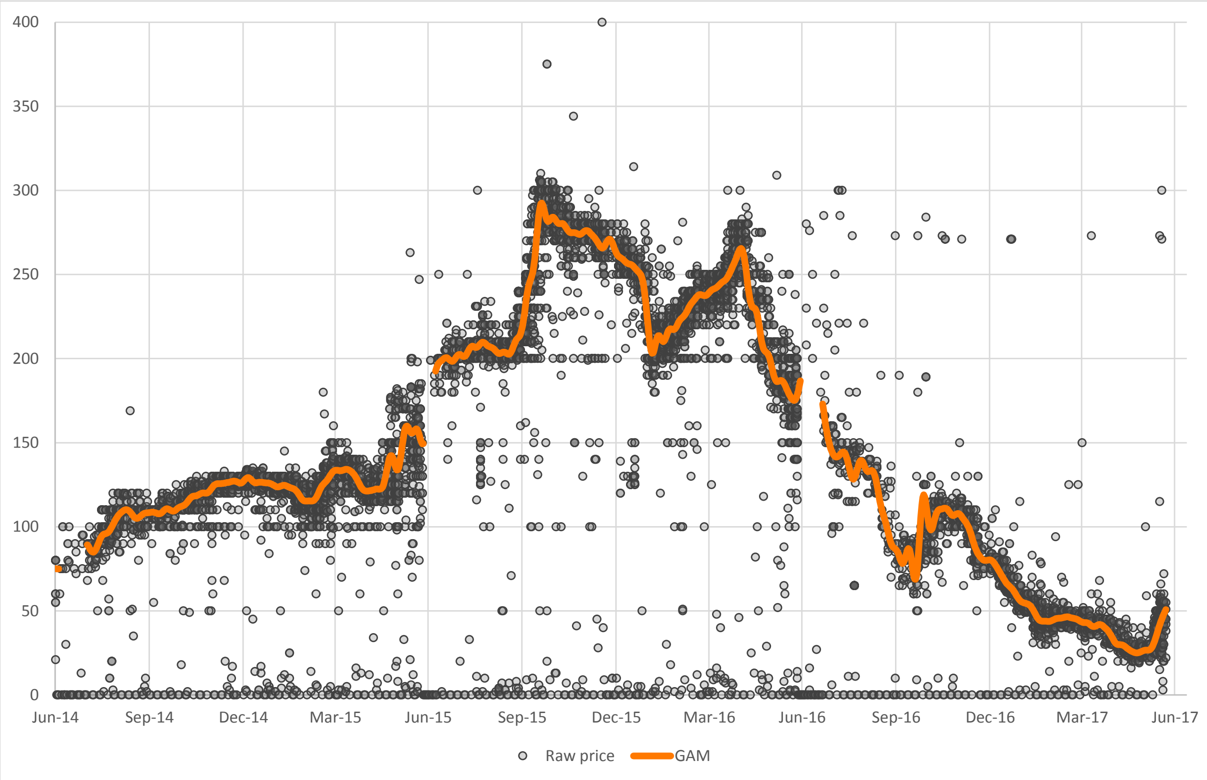


Table 4 shows the difference in predicted values for price, comparing GAM with the 2SD method. Estimated prices on the last day of trade in the dataset (30th June 2017) provide an example of how the methods compare in estimating current (most recently observed) market prices. In many markets significant differences emerge in predicted values on the last day. For example, in the Goulburn and Victorian Murray prices for the 30th June 2017 are around $15 per ML or 30 per cent lower under 2SD versus GAM. There are also markets, particularly outside the southern basin (such as, Border Rivers, Gwydir and Lower Namoi), where the difference is substantial even when averaging over the entire period (2008–09 to 2016–17). In most markets, 2SD produces lower average water market prices relative to GAM. Similar comparisons for the predicted values from LOESS and Median are presented in Appendix A.

Table 4 Difference in predicted values GAM versus 2SD (GAM price less 2SD price)

Region	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Units	**$/ML**	**$/ML**	**$/ML**	**$/ML**
1A Greater Goulburn	14.4	1.4	36.4	-42.8
1B Boort	11.6	3.5	68.0	-27.3
3 Lower Goulburn	7.9	1.6	74.3	-24.4
41 Macalister	-37.8	-0.1	47.3	-37.8
4A Campaspe - Eppalock to WWC	-0.9	2.9	120.3	-16.7
5A Loddon - Coliban Channel to Lower Weir Pool	5.3	6.7	91.8	-16.7
6 VIC Murray - Dart to Barmah	15.3	1.2	37.8	-36.5
6B Lower Broken Creek	13.2	-1.0	43.2	-357.9
7 VIC Murray - Barmah to SA	16.9	-0.4	36.0	-39.9
Border Rivers Regulated River Water Source	-4.9	11.2	132.0	-24.7
Coleambally Irrigation Co-op Ltd	-2.1	-0.8	31.2	-41.9
Gwydir Regulated River Water Source	-6.4	35.5	274.7	-37.8
Harvey Water-Harvey	-3.6	0.4	24.0	-8.2
Lachlan Regulated River Source / that part of the water source downstream of Lake Cargelligo Weir	-19.6	-0.6	45.0	-41.5
Lachlan Regulated River Source / that part of the water source upstream of Lake Cargelligo Weir	-4.6	0.2	47.5	-23.4
Lower Darling Regulated Rivers Water Source	-1.0	0.4	29.0	-27.9
Lower Namoi Regulated River Source	1.9	10.0	60.4	-28.3
Macquarie And Cudgegong Regulated Rivers Water Source / that part of the water source downstream of the upper limit of Lake Burrendong	4.9	2.0	61.2	-28.5
Macquarie And Cudgegong Regulated Rivers Water Source / that part of the water source upstream of the upper limit of Lake Burrendong	-0.1	2.2	24.7	-46.9
Murray Irrigation	-0.5	1.2	42.8	-34.5
Murrumbidgee Regulated River Water Source	4.3	-0.9	34.5	-41.9
New South Wales Murray Regulated River Water Source / that part of the water source downstream of the River Murray at Picnic Point	3.9	0.6	39.8	-39.7
New South Wales Murray Regulated River Water Source / that part of the water source upstream of the River Murray at Picnic Point	4.5	1.3	56.2	-41.3
River Murray Prescribed Watercourse	9.4	-1.0	48.1	-52.0
Upper Namoi Regulated River Source	-9.8	1.9	69.0	-57.8

Entitlement markets

Based on performance metric results Table 5 shows the number of entitlement markets where a particular price cleaning method is preferred. Figure 12 shows the average estimation error for each price cleaning method (where lower values are considered better) across all entitlement markets in the sMDB. Table 6 summarises the differences in predicted entitlement prices between GAM and 2SD. Example charts are also presented for a high frequency market for Goulburn - High reliability (Figure 13), Lower Namoi - General security (Figure 14) and Murrumbidgee – General security (Figure 15). Once again, a full set of results for all entitlement markets is available via the online dashboard accompanying the report.

Table 5 Number of entitlement markets where each method is ranked highest by performance metric

Method	MAE	MEDAE	SMAPE
Median (quarterly)	0	0	0
2SD (quarterly)	0	0	0
GAM	5	7	8
LOESS	7	5	4

Figure 12 Prediction error by method and metric, average across all entitlement markets

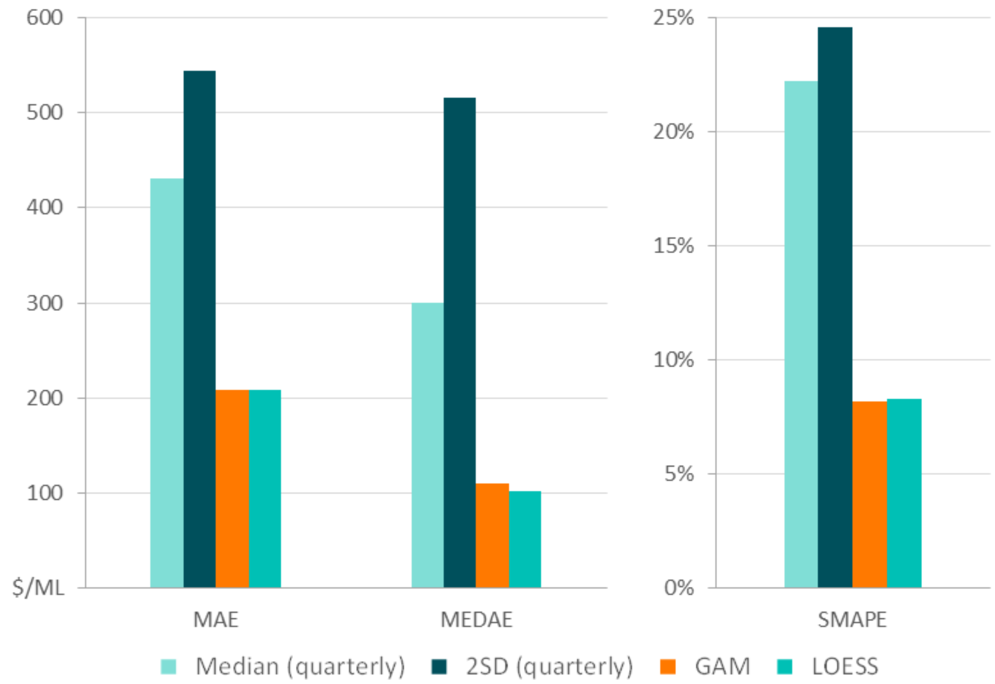


Figure 13 Price predictions by method, Greater Goulburn, high reliability, surface water entitlement market, 2014–15 to 2016–17

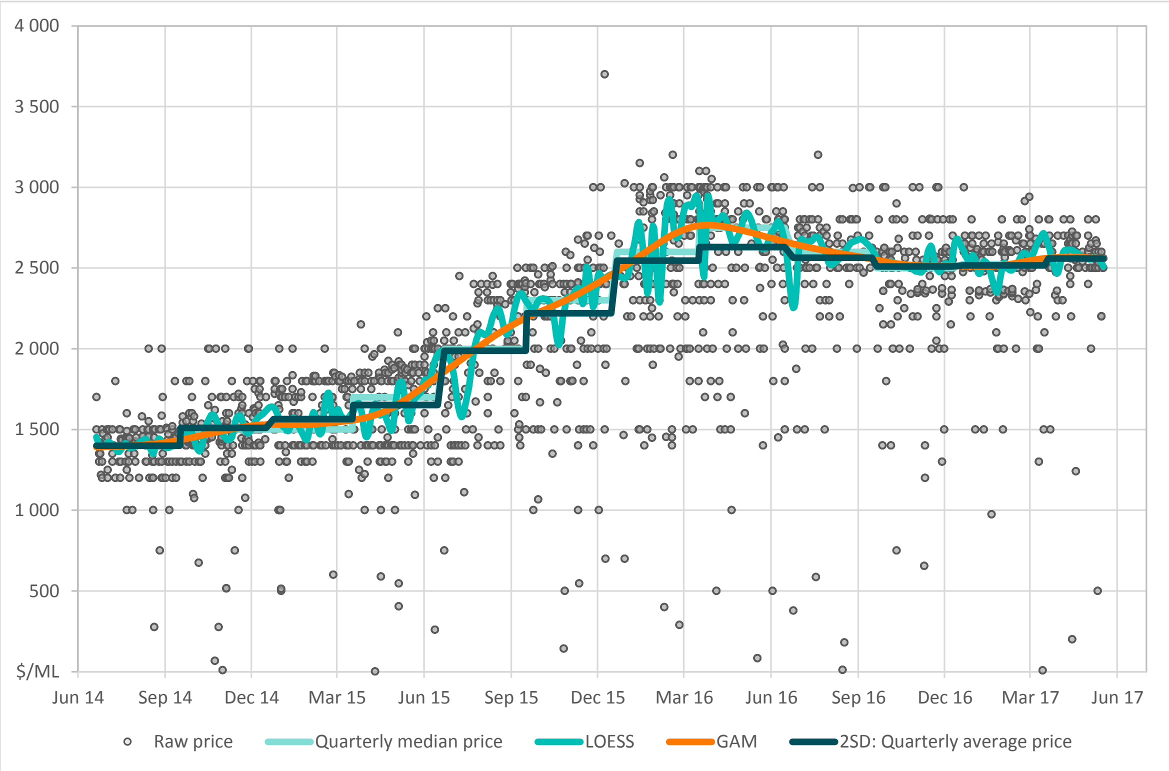


Figure 14 Price predictions by method, Lower Namoi, general reliability, surface water entitlement market, 2014–15 to 2016–17

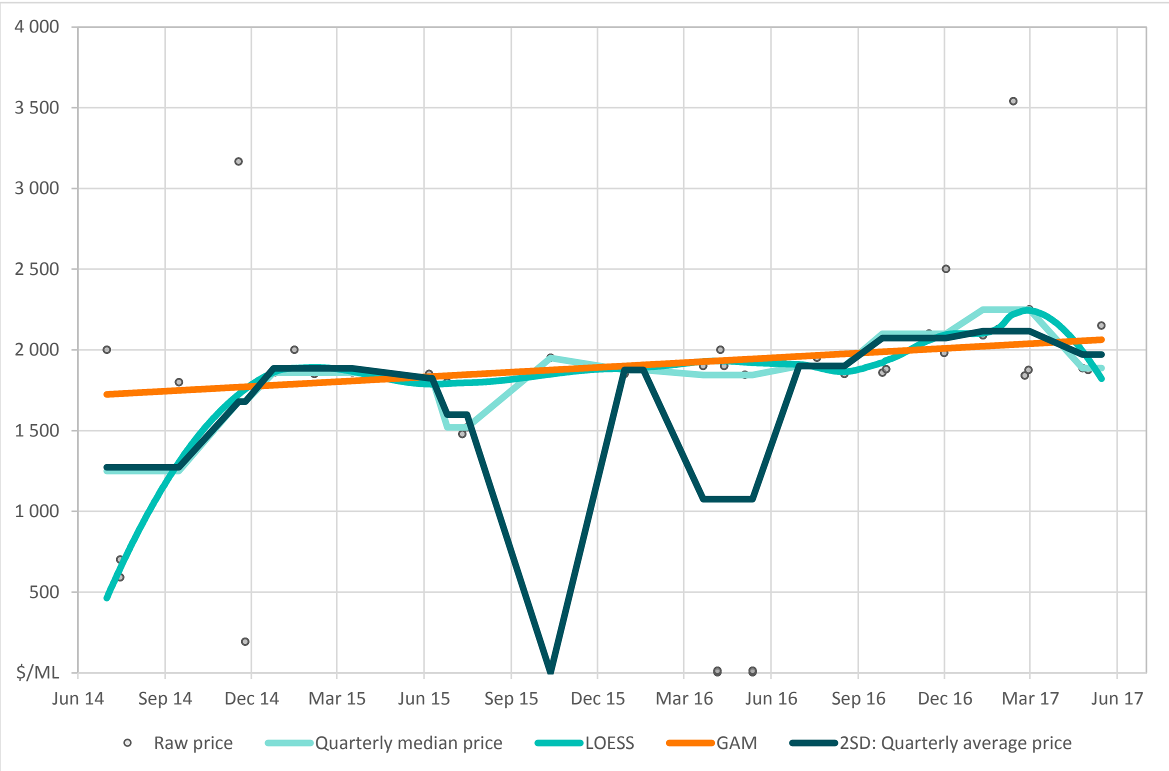
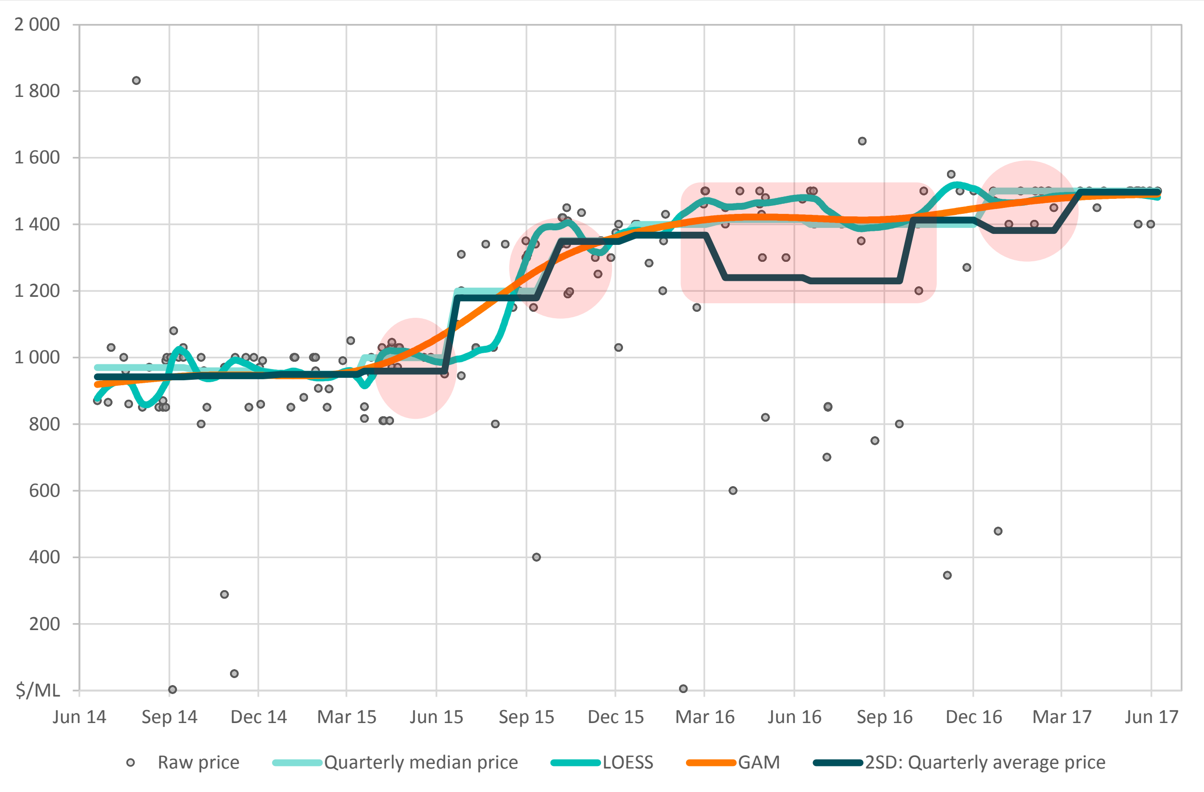


Figure 15 Price predictions by method, Murrumbidgee, general reliability, surface water entitlement market, 2014–15 to 2016–17



Based on the performance metrics and inspection of charts, the GAM method provides the best results for entitlement markets overall. For most metrics GAM and LOESS achieve similar performance. GAM is preferred in more markets based on SMAPE and MEDAE and LOESS is preferred in more markets based on MAE (although MAE is the least robust of the chosen metrics).

Further, the visual results suggest that LOESS lacks robustness relative to GAM and has a tendency to overfit the data at least in some markets (see Figure 13, Figure 14 and Figure 15). This can also be seen by comparing in and out-of-sample metrics (which are both available in the dashboard). For LOESS there tends to be a larger gap between in-sample and out-of-sample performance indicative of overfitting.

In general, the smoothing methods offer significantly more accurate estimates than the quarterly Median and 2SD methods. Note that the performance of monthly median and 2SD methods is in most cases inferior to quarterly, due to the relatively infrequent trading activity (and high level of noise) within entitlement markets (see online dashboard). Once again the 2SD method is shown to be influenced by lower outliers (see for example the period of April-May 2016 in Figure 14 and the period March-September 2016 in Figure 15, marked by red circles).

The results for the Lower Namoi general reliability water entitlement market are shown in Figure 13. This chart is a good example of the robustness of the GAM method in comparison with alternatives, with 2SD, Median and LOESS all influenced to some extent by outliers. While Median is more robust than 2SD it can still be affected by outliers in low frequency markets, especially in periods with only two or three trades. In contrast, the GAM method adapts to the smaller sample size, producing a smooth price series.

Table 6 shows the difference in predicted values for entitlement price, comparing GAM with the 2SD method. If these methods are used to estimate the price on the last day of trade, the difference in predicted values is notable in some markets. In contrast to allocation trade, markets in the southern basin (such as, Goulburn, Vic Murray and NSW Murray), show substantial differences even when averaging over the entire period (2008–09 to 2016–17). Similar comparisons for the predicted values from LOESS and Median are presented in Appendix A.

Table 6 Difference in predicted values GAM versus 2SD (GAM price less 2SD price)

Region	Reliability	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Units		**$/ML**	**$/ML**	**$/ML**	**$/ML**
1A Greater Goulburn	High	-1.4	29.5	238.6	-128.9
1A Greater Goulburn	Low	24.1	6.8	24.1	-17.0
41 Macalister	High	-8.4	30.6	152.3	-73.4
41 Macalister	Low	1.3	-18.9	49.7	-98.7
6 VIC Murray - Dart to Barmah	High	42.4	21.1	263.1	-146.6
6 VIC Murray - Dart to Barmah	Low	5.4	1.9	36.0	-53.9
6B Lower Broken Creek	High	172.5	34.7	578.4	-340.9
7 VIC Murray - Barmah to SA	High	63.6	56.4	236.4	-183.8
7 VIC Murray - Barmah to SA	Low	13.5	8.5	51.6	-21.3
Border Rivers Regulated River Water Source	General	-605.1	-81.5	492.6	-1011.8
Border Rivers Regulated River Water Source	Other	-42.2	150.8	237.9	-42.2
Lachlan Regulated River Source	General	-57.1	-45.9	97.2	-639.9
Lower Namoi Regulated River Source	General	92.2	166.5	869.3	-105.0
Macquarie And Cudgegong Regulated Rivers Water Source	General	-49.8	-23.3	356.3	-351.3
Murrumbidgee Regulated River Water Source	General	-5.8	37.9	188.3	-83.2
New South Wales Murray Regulated River Water Source	General	20.5	22.5	117.7	-35.7
New South Wales Murray Regulated River Water Source	High	-0.5	27.3	438.2	-83.4
River Murray Prescribed Watercourse	General	48.4	-18.0	565.2	-619.4
River Murray Prescribed Watercourse	High	356.9	50.1	604.9	-984.7

Appendix A: Additional results

Tables A1 and A2 show the results of aggregate performance metrics for each price cleaning method for allocation and entitlement markets. Tables A3 and A4 show the difference in predicted values comparing LOESS and median with 2SD, for allocation markets. Tables A5 and A6 show this comparison for entitlement markets.

Table A1 Aggregate performance metrics for allocation markets (with cross validation)

Method	MAE	MEDAE	SMAPE
All markets			
Median (monthly)	16.9	5.0	6.2%
2SD (monthly)	17.4	6.5	6.4%
GAM	15.1*	4.3*	5.4%*
LOESS	15.3	4.3	5.4%
sMDB markets			
Median (monthly)	16.8	5.0	6.1%
2SD (monthly)	17.1	6.4	6.3%
GAM	14.8*	4.2*	5.2%*
LOESS	14.9	4.2	5.2%
non-sMDB markets			
Median (monthly)	18.8*	5.0*	7.5%
2SD (monthly)	20.1	7.3	8.0%
GAM	18.9	5.8	7.4%*
LOESS	18.8	5.7	7.7%

Note: * indicates preferred method according to each metric.

Table A2 Aggregate performance metrics for entitlement markets (with cross validation)

Method	MAE	MEDAE	SMAPE
All markets			
Median (quarterly)	430.7	300.0	22.2%
2SD (quarterly)	544.5	515.6	24.6%
GAM	208.2	110.3	8.2%*
LOESS	207.9*	102.3*	8.3%
sMDB markets			
Median (quarterly)	421.8	300.0	21.3%
2SD (quarterly)	538.4	502.9	23.8%
GAM	203.7*	110.1	7.4%*
LOESS	203.7	103.0*	7.5%
Non sMDB markets			
Median (quarterly)	584.8	655.1	37.7%
2SD (quarterly)	648.8	735.4	37.8%
GAM	275.2	127.4	19.7%*
LOESS	270.7*	89.3*	21.4%

Note: * indicates preferred method according to each metric.

Table A3 Difference in predicted allocation prices, LOESS versus 2SD (LOESS price less 2SD price)

Region	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Units	**$/ML**	**$/ML**	**$/ML**	**$/ML**
1A Greater Goulburn	9.2	0.0	46.9	-195.5
1B Boort	13.0	2.8	72.9	-48.1
3 Lower Goulburn	9.3	0.4	67.0	-106.5
41 Macalister	-47.4	0.4	44.5	-47.4
4A Campaspe - Eppalock to WWC	2.2	3.0	116.3	-16.4
5A Loddon - Coliban Channel to Lower Weir Pool	7.0	6.0	87.1	-89.6
6 VIC Murray - Dart to Barmah	14.8	0.9	38.7	-47.3
6B Lower Broken Creek	8.9	-3.3	30.9	-390.7
7 VIC Murray - Barmah to SA	14.5	-0.3	36.7	-40.9
Border Rivers Regulated River Water Source	1.0	15.7	151.7	-26.4
Coleambally Irrigation Co-op Ltd	-1.5	-0.3	31.6	-44.4
Gwydir Regulated River Water Source	5.7	21.3	193.0	-188.5
Harvey Water-Harvey	-4.4	-0.1	38.7	-36.7
Lachlan Regulated River Source / that part of the water source downstream of Lake Cargelligo Weir	0.1	0.7	41.8	-32.6
Lachlan Regulated River Source / that part of the water source upstream of Lake Cargelligo Weir	1.4	1.2	53.6	-15.0

Region	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Lower Darling Regulated Rivers Water Source	-1.8	-0.2	20.2	-23.3
Lower Namoi Regulated River Source	-3.4	15.7	74.7	-25.9
Macquarie And Cudgegong Regulated Rivers Water Source / that part of the water source downstream of the upper limit of Lake Burrendong	3.0	0.8	27.5	-28.0
Macquarie And Cudgegong Regulated Rivers Water Source / that part of the water source upstream of the upper limit of Lake Burrendong	1.1	4.8	161.2	-43.9
Murray Irrigation	-0.5	1.5	42.1	-33.7
Murrumbidgee Regulated River Water Source	0.0	-0.4	34.9	-31.1
New South Wales Murray Regulated River Water Source / that part of the water source downstream of the River Murray at Picnic Point	4.4	0.4	40.3	-127.2
New South Wales Murray Regulated River Water Source / that part of the water source upstream of the River Murray at Picnic Point	-0.4	1.1	56.2	-89.4
River Murray Prescribed Watercourse	7.3	-0.4	57.4	-51.6
Upper Namoi Regulated River Source	13.6	7.4	71.5	-42.5

Table A4 Difference in predicted allocation prices, median versus 2SD (median price less 2SD price)

Region	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Units	**$/ML**	**$/ML**	**$/ML**	**$/ML**
1A Greater Goulburn	-1.4	1.1	17.8	-27.1
1B Boort	-2.8	2.7	38.0	-2.8
3 Lower Goulburn	3.4	0.7	14.1	-12.7
41 Macalister	-16.0	-0.7	16.4	-16.0
4A Campaspe - Eppalock to WWC	-1.0	1.0	21.8	-6.9
5A Loddon - Coliban Channel to Lower Weir Pool	2.9	4.1	56.0	-2.8
6 VIC Murray - Dart to Barmah	-1.5	0.6	14.8	-23.3
6B Lower Broken Creek	-6.5	-2.1	17.0	-350.0
7 VIC Murray - Barmah to SA	-5.5	0.4	5.6	-18.5
Border Rivers Regulated River Water Source	8.1	2.2	21.4	-33.3
Coleambally Irrigation Co-op Ltd	-0.2	-0.2	20.0	-22.1
Gwydir Regulated River Water Source	-5.0	4.4	82.0	-70.0
Harvey Water-Harvey	0.3	-0.1	9.4	-15.0

Region	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Lachlan Regulated River Source / that part of the water source downstream of Lake Cargelligo Weir	-9.0	-0.8	12.5	-23.3
Lachlan Regulated River Source / that part of the water source upstream of Lake Cargelligo Weir	-0.2	-0.7	10.0	-14.1
Lower Darling Regulated Rivers Water Source	0.2	-1.0	2.6	-8.3
Lower Namoi Regulated River Source	10.2	3.9	40.5	-33.3
Macquarie And Cudgegong Regulated Rivers Water Source / that part of the water source downstream of the upper limit of Lake Burrendong	7.7	0.8	27.2	-28.0
Macquarie And Cudgegong Regulated Rivers Water Source / that part of the water source upstream of the upper limit of Lake Burrendong	0.6	1.7	22.0	-12.0
Murray Irrigation	0.1	0.1	7.4	-5.2
Murrumbidgee Regulated River Water Source	0.3	-0.9	17.4	-19.9
New South Wales Murray Regulated River Water Source / that part of the water source downstream of the River Murray at Picnic Point	-0.8	-0.3	12.2	-18.2
New South Wales Murray Regulated River Water Source / that part of the water source upstream of the River Murray at Picnic Point	-2.4	-2.1	10.9	-34.0
River Murray Prescribed Watercourse	-0.2	-0.9	14.0	-23.5
Upper Namoi Regulated River Source	11.4	5.0	35.8	-20.0

Table A5 Difference in predicted entitlement prices, LOESS versus 2SD (LOESS price less 2SD price)

Region	Reliability	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Units		**$/ML**	**$/ML**	**$/ML**	**$/ML**
1A Greater Goulburn	High	-50.1	36.9	393.4	-410.5
1A Greater Goulburn	Low	41.6	9.4	41.6	-11.4
41 Macalister	High	1.5	20.1	173.9	-155.6
41 Macalister	Low	-15.5	-31.4	47.9	-148.7
6 VIC Murray - Dart to Barmah	High	40.1	39.7	329.5	-131.1
6 VIC Murray - Dart to Barmah	Low	2.6	1.8	27.7	-63.1
6B Lower Broken Creek	High	-105.9	44.5	806.5	-497.7
7 VIC Murray - Barmah to SA	High	82.9	88.6	382.8	-157.2
7 VIC Murray - Barmah to SA	Low	14.2	8.8	54.6	-20.2
Border Rivers Regulated River Water Source	General	-31.1	-82.6	1715.2	-1664.7
Border Rivers Regulated River Water Source	Other	79.0	179.5	259.2	25.0

Region	Reliability	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Lachlan Regulated River Source	General	-31.5	-46.2	94.6	-648.3
Lower Namoi Regulated River Source	General	-148.9	108.7	856.7	-809.3
Macquarie And Cudgegong Regulated Rivers Water Source	General	-1378.5	81.8	599.2	-1378.5
Murrumbidgee Regulated River Water Source	General	-15.3	46.8	248.7	-183.7
New South Wales Murray Regulated River Water Source	General	35.3	24.7	126.7	-43.2
New South Wales Murray Regulated River Water Source	High	-30.2	37.3	496.1	-70.4
River Murray Prescribed Watercourse	General	46.3	-38.9	512.2	-604.8
River Murray Prescribed Watercourse	High	356.9	50.1	604.9	-984.7

Table A6 Difference in predicted entitlement prices, median versus 2SD (median price less 2SD price)

Region	Reliability	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
Units		**$/ML**	**$/ML**	**$/ML**	**$/ML**
1A Greater Goulburn	High	0.3	22.6	119.4	-65.3
1A Greater Goulburn	Low	15.6	12.3	27.0	2.9
41 Macalister	High	40.0	33.6	126.8	-128.6
41 Macalister	Low	-37.5	-16.5	46.7	-100.0
6 VIC Murray - Dart to Barmah	High	27.4	11.0	129.3	-49.9
6 VIC Murray - Dart to Barmah	Low	11.0	3.9	24.4	-63.4
6B Lower Broken Creek	High	0.0	36.9	507.8	-100.0
7 VIC Murray - Barmah to SA	High	117.8	71.2	213.6	-28.4
7 VIC Murray - Barmah to SA	Low	19.7	9.0	38.8	-18.1
Border Rivers Regulated River Water Source	General	-8.5	119.7	668.1	-198.0
Border Rivers Regulated River Water Source	Other	0.0	140.7	169.7	0.0
Lachlan Regulated River Source	General	-27.0	-53.6	41.1	-652.3
Lower Namoi Regulated River Source	General	-83.8	121.6	770.3	-83.8
Macquarie And Cudgegong Regulated Rivers Water Source	General	453.6	68.5	453.6	-277.1
Murrumbidgee Regulated River Water Source	General	3.3	45.2	175.4	-13.3

Region	Reliability	Price difference on the last day of modelling period	Average difference	Maximum difference	Minimum difference
New South Wales Murray Regulated River Water Source	General	21.0	27.9	157.9	-15.8
New South Wales Murray Regulated River Water Source	High	4.9	8.9	105.6	-79.2
River Murray Prescribed Watercourse	General	-41.4	40.6	339.6	-289.9
River Murray Prescribed Watercourse	High	0.0	85.2	487.5	0.0

References

ABARES 2015, Australian water markets report 2013–14, Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra.

Cleveland, W., Grosse, E. and Shyu, W. (1992) Local regression models. Chapter 8 of *Statistical Models in S* eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.

Hastie, T. (1992) Generalised Additive Models Chapter 7 of *Statistical Models in S* eds J.M. Chambers and T.J. Hastie, Wadsworth & Brooks/Cole.