

Predictive Models for the Victorian Water Market: A Case Study of the Murray-Darling Basin

Meredith Yates, Almaz Khalilov, Sultan Maharramli

June 2024

Contents

1	Introduction	<i>Meredith Yates</i>	1
1.1	Background		1
1.2	Premise		2
2	Inputs to Water Price Prediction Model		2
2.1	Scope		3
3	Machine Learning - Predicting Water Price		4
3.1	Anomaly Detection and Outlier Removal of Water Market Data		4
3.1.1	Density-based spatial clustering of applications with noise algorithm (DBSCAN) <i>Almaz Khalilov</i>		4
3.1.2	Rolling Interquartile Range <i>Almaz Khalilov and Meredith Yates</i>		5
3.2	Exogenous Data Analysis <i>Almaz Khalilov</i>		6
3.2.1	Visual Analysis of Exogenous Relationships Via PCA		6
3.2.2	Seasonality Analysis of Exogenous Variables		8
3.3	Temporary Water Allocation Price Prediction by Using Machine Learning Regression <i>Sultan Maharramli</i>		8
3.4	Data Preparation for Regression Analysis by A Deep Neural Network <i>Sultan Maharramli</i>		9
3.5	Further Data Processing <i>Sultan Maharramli</i>		9
3.6	Highly Non-linear Nature of New Datasets <i>Sultan Maharramli</i>		10
3.7	Time Series Estimation of Volume Weighted Average Prices via a Deep Neural Network <i>Sultan Maharramli</i>		11
3.8	Time Series Estimation of Volume Weighted Average Prices via an LSTM Network <i>Sultan Maharramli</i>		12
3.9	Time Series Estimation of Volume Weighted Average Prices via a Bayesian Neural Network <i>Sultan Maharramli</i>		13
3.10	Time Series Estimation of Volume Weighted Average Prices via a TensorFlow Structural Time Series Modelling <i>Sultan Maharramli</i>		15
3.11	Pruning a DNN Network using a Multi-Armed Bandit Algorithm <i>Sultan Maharramli</i>		17

3.12	ARIMA Model Parameter Selection	18
3.13	Basic ARMA	19
3.14	Rolling ARMA and SARIMAX Models	<i>Almaz Khalilov and Meredith Yates</i> 20
3.15	MOMENT Model Prediction	21
4	Optimisation - Timing of water purchases	<i>Meredith Yates</i> 22
4.1	Simple Formulation	22
4.2	Rolling Optimisation	23
4.2.1	Uncertainty Correction	23
4.3	Results	23
4.4	Continued Optimisation	25
4.4.1	Maximum Likelihood	25
4.4.2	Updating Volume Predictions	25
4.4.3	Improving Price Predictions	28
4.5	Applying to Decision Making	28
4.5.1	Further Work	30
5	Optimisation 2 - Optimising Portfolio Allocation	<i>Almaz Khalilov</i> 30
5.1	Problem Formulation	30
5.2	Primal Problem Formulation and data synthesis	31
5.2.1	Data Synthesis	31
5.3	Solving using Stochastic Gradient Descent (SGD) with vanishing gradient	31
5.4	Solving using Newton's Method	32
6	Appendix A1	a
6.1	Data Preprocessing for daily data	a

1 Introduction

1.1 Background

The Murray-Darling Basin is a basin crossing Queensland, New South Wales, Victoria and South Australia. Water is extracted from the basin for irrigation. The right to extract and use water is an **entitlement**.

Entitlements to extract water can be traded in the Murray-Darling Basin. Allocations (the amount of water you are entitled to in that year) can also be traded. This market is primarily used by irrigators (farmers), the government, and investors.

The market is largely independent of external market factors but is highly dependent on environmental conditions. In higher temperatures, more water is needed to be applied to crops. In periods of higher rainfall, less irrigation is needed and allocation volumes are higher. In droughts, prices rise.

Each year farmers need to determine if they will have enough water to support their crops. If their crops are annual, they can make the decision to skip the year and sell their allocation. If their crops are perennial, skipping would result in the loss of a multi-year investment e.g. orchard.

Selling an entitlement means that they no longer have annual allocations and must purchase them. Follow-

ing the millennium drought, many irrigators sold their entitlements to the government to pay debts accrued during the drought. They then needed to purchase water allocations each year to sustain their crops and were more subject to the whims of the market.

If an allocation holder does not use their allocation, the volume is kept in their account as **carryover**. This volume can then be used in the future. However, if there is a high flow event that requires the release of water from storages e.g. a spill, this volume lost to spill is first taken from any carryover. Thus there is some risk in leaving it to carryover.

1.2 Premise

The government (state, federal) has been buying entitlements and allocations for **environmental water**, water delivered specifically for environmental outcomes. Reports stated that they would require a total volume gigalitres in entitlements and continuing to acquire this. Due to the large volume of water to be acquired, the costs could be increased by millions of dollars if purchases are made at inopportune times. Thus it is useful to have informed decision making.

This paper aims to provide a comprehensive analysis of the relationships between the water market and two key stakeholder groups: irrigators and government purchasers. The paper will focus on three main areas:

1. Development of a model to predict prices and volumes using previous values and exogenous variables. This process is described in [3](#).
2. Development of an algorithm to decide on when to purchase entitlements given the output of the model. This process is described in [4](#).
3. Development of an algorithm to optimise and contract the portfolio allocation weighting of each of the three trading zones vs ASX stock market [5](#)

With these models, a comprehensive drought mitigation strategy and action plan can be developed for times of drought. The strategy will provide a framework for the government to make informed decisions about purchasing environmental water entitlements, ensuring that purchases are made at optimal times to minimize costs and maximize benefits. The action plan will also offer guidance to irrigators and farmers on how to optimize their water usage during drought conditions, helping them to minimize the impact of water scarcity on their operations. Moreover, the water allocation model will enable the efficient distribution of available water resources among different agricultural and urban zones during drought periods.

2 Inputs to Water Price Prediction Model

For the development of our water price prediction model, the following data is used, sourced from the Victorian water register [[11](#)]. Our main dataset consists of trades from 2010-2024 in three main trading zones:

1. Greater Goulbourn (Trading zone source 1A)
2. Murray River - Barmath to SA section (Trading zone source 7)
3. Murray River - Dart to Barmah section (Trading zone section 6)

These were chosen due to the fact that they have the most organised and available data, compared to other zones. Additionally, Low reliability data, as well as data with rejected, expired or refused were not considered.

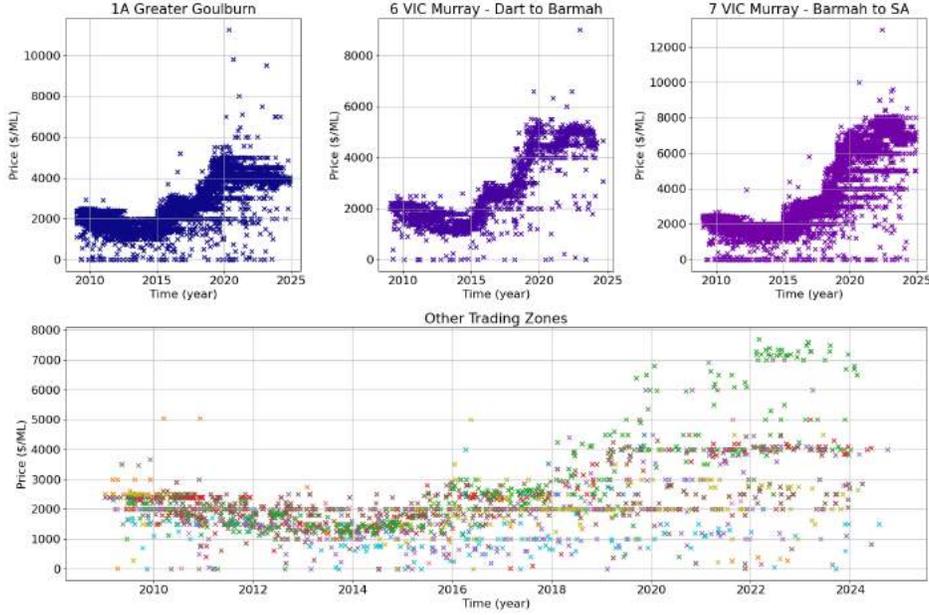


Figure 1: Full Water Market Dataset, Separated By Trading Zones

To Enhance the model, the following exogenous variables were analysed both on the daily and monthly timescale. The parameters are compiled across all weather stations in Victoria from 2010-present. Details of data fetching are described in Appendix A:

1. Historic Water Course Discharge [5]
2. Historic Rainfall [5]
3. Historic Storage Volume [5]
4. Historic Electrical Conductivity [5]
5. Historic Temperature [5]

The selection of relevant inputs is important for the model performance as irrelevant variables increase the likelihood of spurious relationships becoming included in the model.

2.1 Scope

This paper explores the application of various machine learning techniques to predict water allocation prices and optimize water purchasing decisions in the Murray-Darling Basin water market. The key areas of interest and the core analysis techniques are as follows:

1. Data preprocessing and anomaly detection
 - DBSCAN clustering is used to identify and remove outliers and noise in the raw water market price and volume data.
 - rolling interquartile range is employed to detect and remove anomalies, and IQR is calculated to convert to monthly data
2. Analysis of exogenous variables

- Principal Component Analysis (PCA) is applied to reduce dimensionality, analyze correlations, and better understand the relationships that impact prices.
 - Additive seasonal decomposition is performed on each exogenous variable to examine and remove seasonal effects before PCA.
3. Water allocation price prediction
 - Linear regression and deep neural networks are used to predict water allocation prices based on previous price, volume, and exogenous data.
 - LSTM networks are employed to predict prices.
 - ARIMA, ARIMAX, and MOMENT models are also explored for price trend forecasting.
 4. Optimization of water purchasing decisions
 - A rolling optimization approach using the water price prediction models is developed to decide the optimal timing and volumes of water allocations to purchase.
 - Uncertainty in the price predictions is accounted for through regression and maximum likelihood estimation.
 5. Portfolio allocation optimization
 - Stochastic gradient descent and Newton's method are used to optimize the allocation of funds between water allocations in different trading zones and the stock market to maximize the Sharpe ratio.

3 Machine Learning - Predicting Water Price

Prediction of $\text{Price}(t, x)$ and $\text{Volume}(t, x)$ based on contextual variables in x , e.g rainfall, temperature, and auto-regression.

Our aim for this section is to simply predict the price trends via multiple different methods, along with exogenous variables to find any cyclical, seasonal or directional trends or correlations.

With the assumption that future exogenous data is available further down in the future, the goal of our models is to find that "edge", ensuring better than random performance and hence model a function that can be used in optimisation.

3.1 Anomaly Detection and Outlier Removal of Water Market Data

The data must be "cleaned" for further processing, we exclude irrelevant data and outliers.

3.1.1 Density-based spatial clustering of applications with noise algorithm (DBSCAN)

This algorithm - the Density-based spatial clustering of applications with noise algorithm was used to cluster the two dense regions, as well as separate outliers. This algorithm was selected due to the usefulness of separating outliers along a dense trend.

While similar clustering algorithms such as K-means and GMM's are often used to cluster and group data, they present a number of challenges for this specific dataset. Mainly due to the fact that their objective function clusters in a concentric way, and hence is unsuited for time-series data.

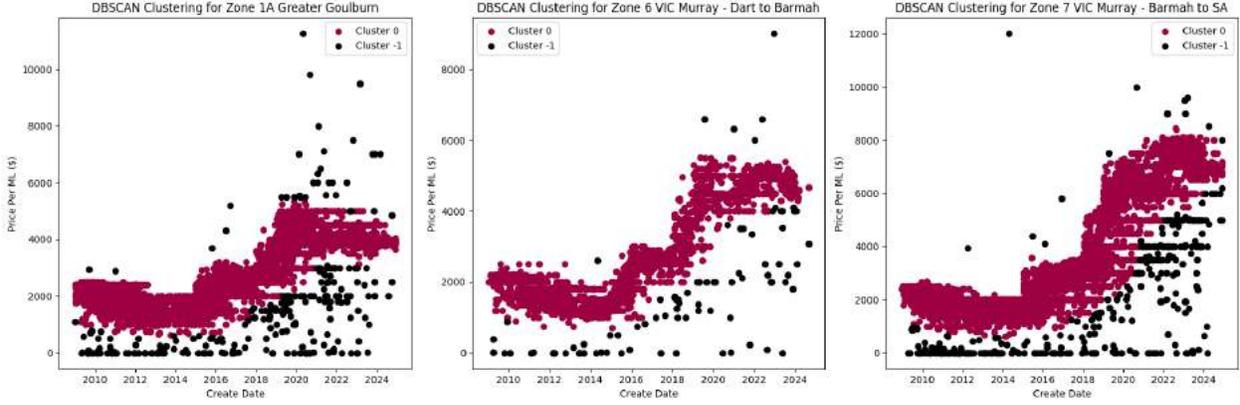


Figure 2: DBSCAN Clustering removing noise

3.1.2 Rolling Interquartile Range

The interquartile range is defined by equation 1 where Q_1 and Q_3 represent the first and third quartiles of the data. It is used with equation 2 to define a typical range of values, and thus allowing for values outside of this range to be identified as outliers. We performed in prior to DBSCAN to separate out the extreme outliers. A rolling approach was used where the window of values within the days of each data point was used to construct bounds for each timestamp.

$$IQR = Q_3 - Q_1 \quad (1)$$

$$x \in (Q_1 - 1.5IQR, Q_3 + 1.5IQR) \quad (2)$$

Additionally, we used IQR for each monthly span to segment and find the IQR and mean for each month, to transform a series of trades occurring within the month to a monthly mean with a corresponding IQR, for both price and volume. We notice that for volume, there is no significant trend.

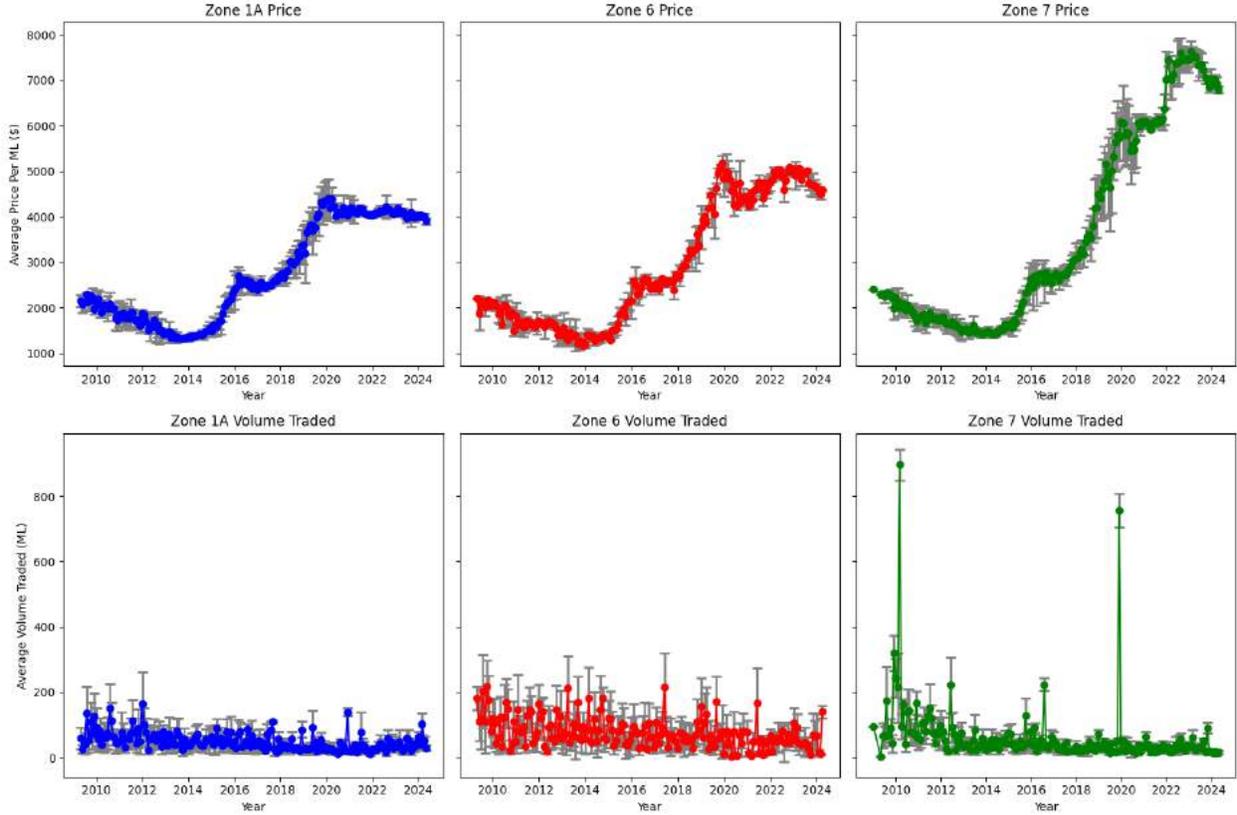


Figure 3: Monthly resampled price and Data, with IQR as error bars

3.2 Exogenous Data Analysis

This section examines the exogenous variables in our model using data [5] for stations in Victoria from 2010 to 2024. Principal Component Analysis (PCA) is used to reduce the data’s dimensionality, analyze correlations, and better understand the variables.

A trade-off exists between temporal and spatial patterns in the data, as PCA requires valid values for every row and column. The data is messy, with many stations lacking complete daily data from 2010-2024. To avoid imputation, we use the method from Appendix 1 to remove missing data.

Additionally, because we aim to analyse the effect of each exogenous variable through time, and not spatially, each monthly exogenous variable is compiled to a monthly mean from each weather station to capture the behaviour through time.

3.2.1 Visual Analysis of Exogenous Relationships Via PCA

Many trends can be noticed and inspected visually using PCA. This Analysis allows inspection of what exogenous variables to use and the relationships between them. We use daily data, with method outlined in Appendix 1A to delete missing rows and columns.

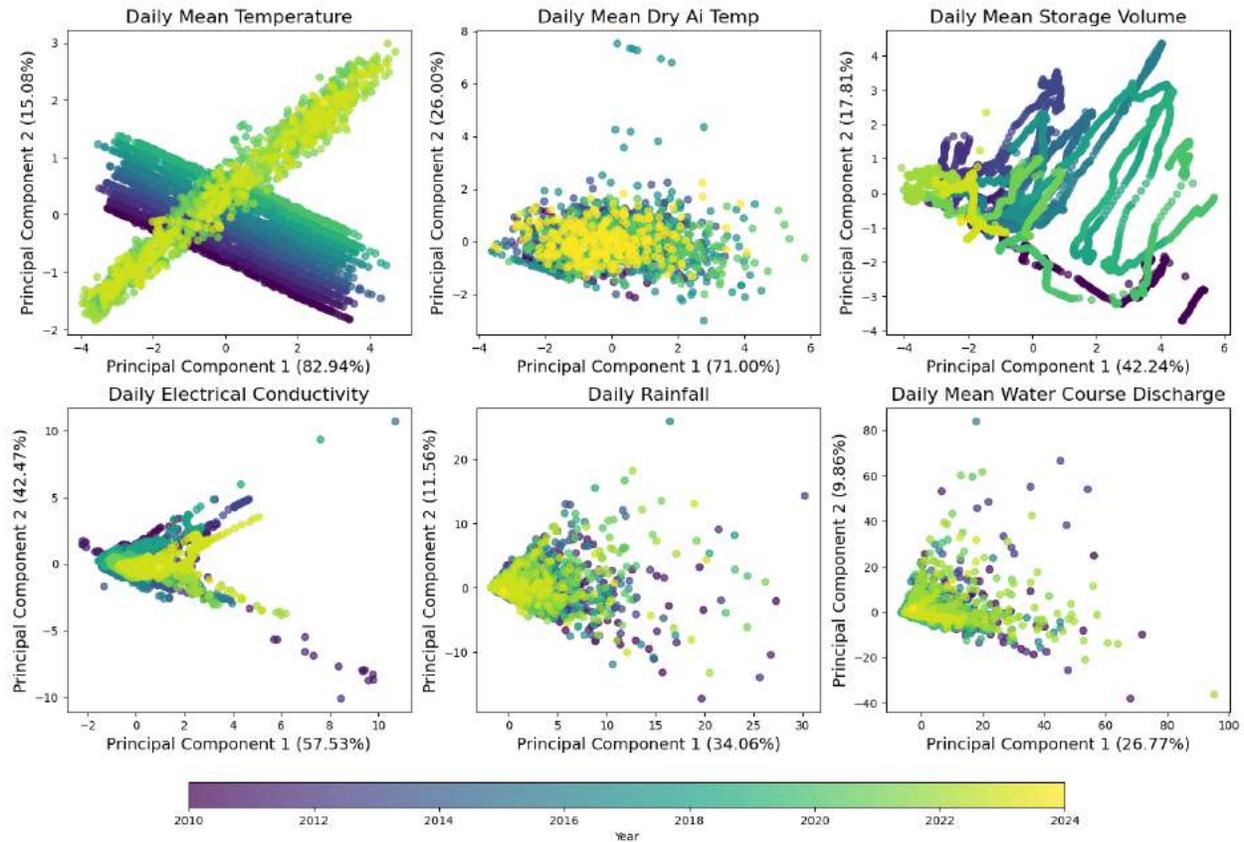


Figure 4: Exogenous daily Data PCA

We notice quite a few important trends we can use in the data. For the temperatures, we can see that only two principal components can explain a very large variation in data, with PC1 explaining 82.94% of the variance, and PC2 explaining 15.08% of the variance. A pattern in the data can be seen. While the earlier data forms shapes in the form of sloping lines downwards, points at later times form relationships between the two principal components upwards.

What this means is that in 2010, there seemed to be two inversely correlated factors, where as the Principal component moves upwards, the principal component 2 moves downwards, additionally it can be seen that from 2010 to 2017, the points trend upwards along PC2.

A hypothetical interpretation of this PCA data up to this point is that PC1 may explain the variance due to seasonality (as it is the primary factor), while PC2 may explain some of the variance due to global warming or related factors, as temperatures along that dimension trend upwards through time. After 2020, PC1 and PC2 became directly correlated however.

Other data such as dry air temperature are hard to interpret, with largely no uniform correlation and a concentric uniform cluster for dry air temperature around its mean. Thus it may mean that there is no clear underlying structure in the data and no strong relationships or dependencies.

Daily mean Storage Volume has many lines on its plot, strongly suggesting that it is dependant on the previous time value, as expected. The rest of the data has a "cone shape", suggesting that variance in either PC1 or PC2 results in variance in the other.

It's important to note however, that each variable has a different number of weather stations as input, and hence direct comparison of variance between them is not able to be compared between them.

3.2.2 Seasonality Analysis of Exogenous Variables

Firstly, it is best to analyse the seasonality of variables before performing PCA analysis, as excess variance may be captured with the seasonality, and hence PCA analysis may not be as effective. We opt for Additive seasonal decomposition of each variable, according to:

$$Y_t = T_t + S_t + R_t \quad (3)$$

where: T_t , S_t , R_t represents the trend, seasonal component, and residual through time respectively.

We can see that while monthly mean temperature, dry air temperature and watercourse discharge has a strong seasonal effect, the other variables do not exhibit strong seasonality.

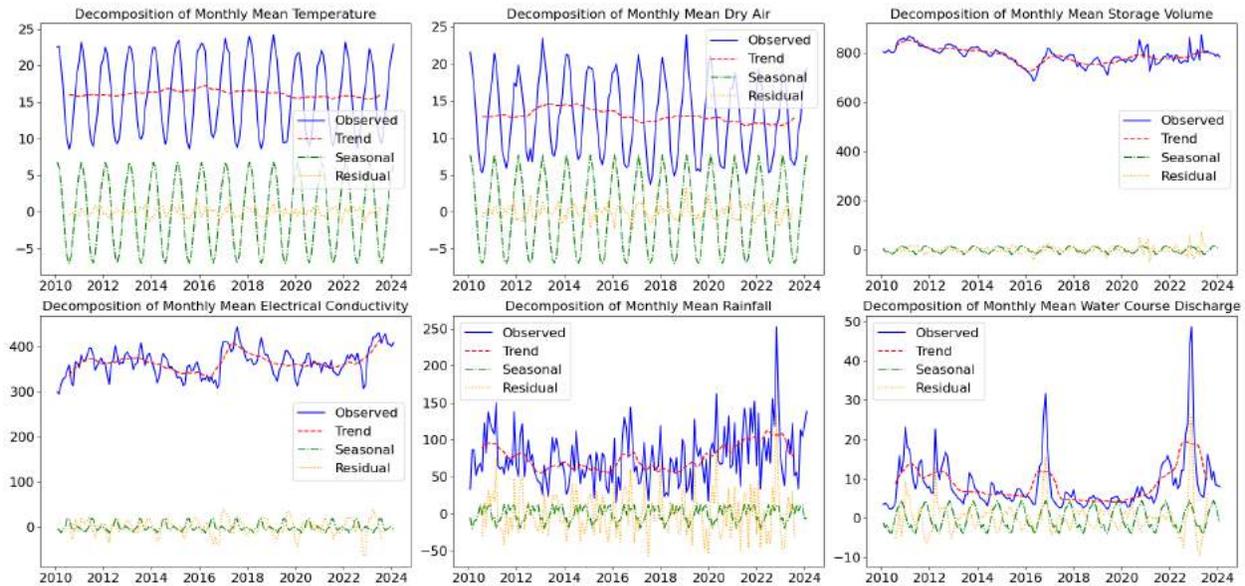


Figure 5: Seasonal Decomposition of Exogenous Variables

Table 1: Seasonality Effect on Various Monthly Mean Variables

Variable	Seasonality Effect (%)
Monthly Mean Temperature	8.232955
Monthly Mean Dry Air	13.120156
Monthly Mean Storage Volume	0.019220
Monthly Mean Electrical Conductivity	0.058204
Monthly Mean Rainfall	1.433109
Monthly Mean Water Course Discharge	6.270994

3.3 Temporary Water Allocation Price Prediction by Using Machine Learning Regression

The raw dataset regarding temporary water allocation prices and traded allocation volumes has several non-numerical features with the potential to have a considerable influence on the allocation prices. These data features include different water system sources, different seller authorities, different trading zones, etc.

Certainly, there are other essential factors as well, as discussed in the previous section dedicated to exogenous variables, including but not limited to monthly mean temperatures, water turbidity, average monthly rainfall, etc. However, considering the fact that the raw data obtained does not integrate any of those specific exogenous variables, a simple thus reasonable approach to take is to utilize the easily accessible data for regression analysis. As mentioned, the raw water market data has only two numerical features, namely the volume of traded volume at each transaction and associated allocation price, while all the remaining ones are non-numerical.

3.4 Data Preparation for Regression Analysis by A Deep Neural Network

An initial look at the raw data makes it very clear that it does require a certain level of processing. Firstly, each water transaction made has a corresponding status according to the data, and there are four different statuses, namely 'Recorded', 'Rejected', 'Refused', and 'Expired'. As a matter of fact, transactions only with the 'Recorded' status are kept, while all the remaining data points are excluded from the data. Furthermore, the obtained data is further processed to only contain those characterized as highly reliable, while all the remaining ones are excluded.

Additionally, the data indicates that there are several recorded transactions in each month between 2009 and 2024. The analysis of several research articles, intending to develop price prediction models reveals that all the models attempt to predict allocation prices on a monthly basis. Based on this information, it is essential to organize the data so that monthly volume weighted average allocation prices can be computed. In order to compute monthly prices, the data is divided into sub-sets for each representative year and its months. Prior to this step, the volume of traded water and its corresponding price is multiplied for each recorded transaction and incorporated into the data as another feature. Having undertaken this step and divided the data into each allocation year, volume weighted average prices are calculated by dividing the total sum of (traded water volume * allocation price) by the total sum of traded water volume for each month of each representative year from 2009 to 2024. Finally, after joining each year with its corresponding months, the dataset displaying total traded water volumes and volume weighted average prices on a monthly basis is created.

3.5 Further Data Processing

As was stated previously, there are different water system and trading zone sources in the raw water market data, possibly implying that the co-existence of these various sources may be a potential factor contributing to the non-linearity of the relationship between the variables of interest. Considering this possibility, a reasonable step to take is to further process the water market data by categorizing it based on each water system source, trading zone source, and selling water authority. Upon inspection of these non-numeric features of the data, it is apparent that out of nine water system sources, only two of them (Murray and Goulburn) govern most of the transactions made, with total transaction counts of 15886 and 12818, respectively. Therefore, owing to the possibility of different exogenous factors contributing to the highly non-linear relationship between the variables of interest, it is reasonable to divide the raw transaction data based on these two water system sources. Having performed this categorization and further data inspection, it becomes clear that the transaction data for Murray is characterized by only three trading zone sources, of which the governing one is '7 VIC Murray – Barmah to SA'. Further processing of this dataset according to the 'Seller Water Authority' feature reveals that there are two governing selling authorities, namely 'Lower Murray Water' and 'Goulburn-Murray Water'. Considering this, there are two distinct datasets obtained from the transaction data for Murray.

The same processing steps are performed on the transaction data for Goulburn. Its analysis reveals that the primary governing trading zone source is '1A Greater Goulburn'. However, in comparison with the transaction data of Murray, this dataset is characterized as having only one major selling water authority, namely 'Goulburn-Murray Water'. Therefore, only one transaction dataset is obtained for Goulburn. Overall, from the raw water market data obtained from the Internet, only three datasets were obtained as a result of data processing attempts.

3.6 Highly Non-linear Nature of New Datasets

As was previously discussed, the transaction data for Goulburn is significantly larger than the two datasets obtained for Murray. Therefore, this dataset is reasonably used for new analysis of the non-linear relationship between the variables of interest. To determine if there is any relationship between total monthly traded volume of water and associated volume weighted average price, a deep neural network with five Dense hidden layers, each with a unit size of 128, is created. All Dense layers of the model are designed to use the ‘relu’ activation function, while the output layer with a single unit size employs a ‘linear’ activation function owing to the fact that a regression problem is under consideration. For the model to learn all the potential underlying relationships in the training dataset, an epoch size of 1000 and a batch size of 1 is used, with the intention of creating a favorable condition for over-fitting. Furthermore, to serve this purpose, a very small training dataset (test size = 0.9) is created so as to enable the model easily learn the relationship. After training the model, predictions are made for the training dataset, with the intention of determining the performance of the training of the model. However, even with five Dense layers and favorable conditions for over-fitting, the model still cannot make satisfactory predictions for the training dataset, implying the significantly non-linear nature of the relationship between the variables of interest.

To further simplify the training process for the model, the transaction data corresponding to the summer season of each representative year is extracted from the dataset and used for regression analysis. Despite the fact that the same model structure and environment is used to train the model, there is not much improvement in the model’s training performance. As a last step, limits are imposed on the minimum and maximum values of volume weighted average prices for the transaction data of Goulburn for the summer season to bring more sense to the data. This step is taken based on the fact that calculated prices are comparatively higher than those reported elsewhere, implying the potential existence of outliers in the data. It could also be supported by the fact that some data points regarding allocation prices are unreasonably high. To account for this fact, lower and upper limits are posed on the calculated monthly allocation prices, the lower and upper limits being at 0.5 and 1575 AUD per ML of water, respectively. In spite of the fact that limitations posed on the seasonal dataset decreased the size of the training dataset further, the deep neural network of the same structure still under-performed. All these steps taken once again proved the highly non-linear nature of the relationship between the variables, emphasizing the importance of more numerical features for an effective prediction model. The below figure represents predictions made by the DNN model of five Dense layers on the training dataset. The figure clearly shows that the model’s predictions even for the training dataset are not satisfactory.

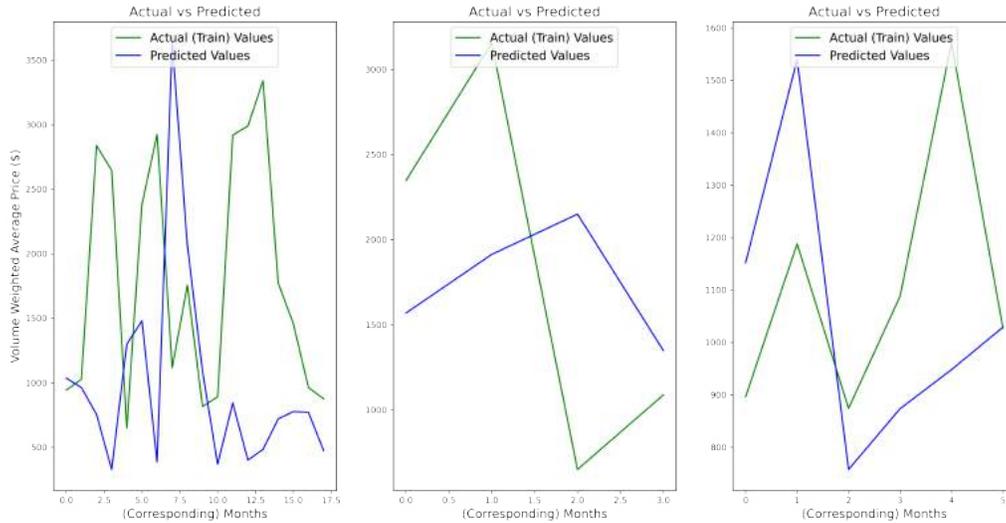


Figure 6: Non-linearity of Processed Datasets

3.7 Time Series Estimation of Volume Weighted Average Prices via a Deep Neural Network

From the previous analysis, it is apparent that the prediction of volume weighted average allocation prices via regression is very unlikely owing to the fact that the raw transaction data has only two numerical features. Therefore, a more reasonable approach to take is to treat the data, particularly its portion regarding the average prices as time series data and employ relevant techniques for the intended estimation. Importantly, the fact that time series estimation techniques require reforming the data shape indeed attributes to the effectiveness of estimation due primarily to the fact that a single-point estimate is made based on more than one parameter, which usually depends on the chosen window size. As a matter of fact, this approach itself ensures that single-point estimates are performed by the model using more than one numerical feature, even though it is of the same type. For the case under consideration, prior to creating a deep neural network, a function is created in order to appropriately reform the shape of the price dataset according to the requirements of time series estimation models. A window size of three is chosen, implying that the model will predict volume weighted average allocation price for a next month based on the price data from the previous three months. Having reformed the price data shape, a deep neural network is created with two Dense hidden layers. Both hidden layers employ the 'relu' activation function, and the model is then compiled using the 'Adam' optimizer with a learning rate of 0.0001. To ensure effective training of the model, a batch size of 1 is chosen, which could certainly be improved upon further analysis. After training the model, predictions are performed for both training and test datasets to evaluate the model performance. It could be concluded that even though there is room for improvement for the model's training performance, it is apparent that the model provides satisfactory results to a certain degree provided the fact that the transaction data used is not quite reliable. However, the model obviously does not perform very well on the testing data, which could be attributed once again to the high variational distribution of the data points. The potential existence of outliers in the raw transaction dataset and their random distribution during the preparation of training and test datasets is very likely to result in this performance of the model, which once again puts a strong emphasis on the significance of reliable data. As a step to increase the model's performance on the test dataset, a seasonal price trend is extracted from the first dataset. After dividing the new dataset into training and testing sub-sets and training the same model, the following graphs are

obtained for the model’s performance on training and testing datasets. As can be seen, the model’s prediction performance on the training dataset is comparatively better, while that on the testing dataset noticeably worsened, which could possibly be attributed to the phenomenon of over-fitting.

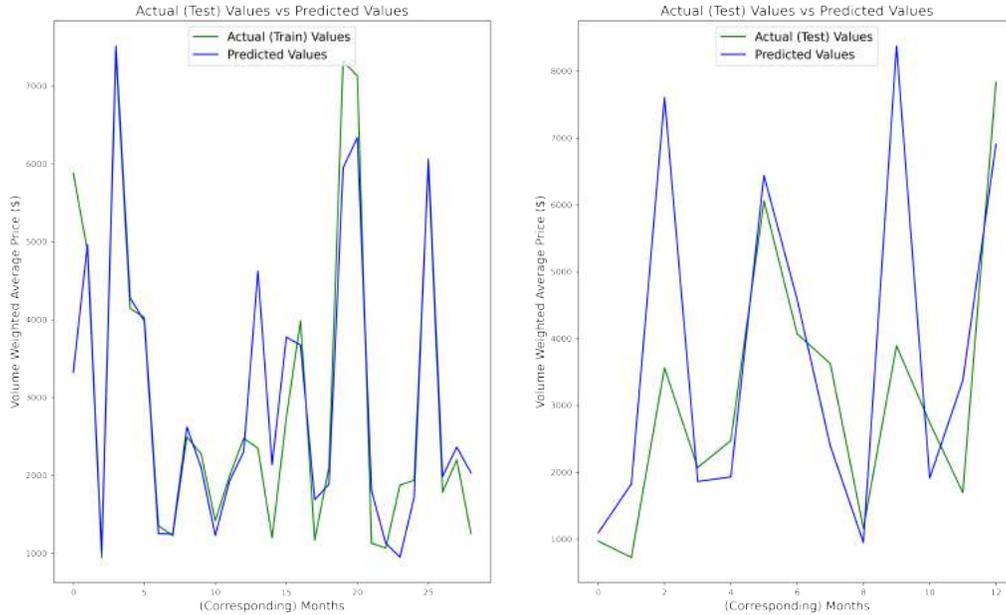


Figure 7: Water Allocation Price Prediction by a DNN model

3.8 Time Series Estimation of Volume Weighted Average Prices via an LSTM Network

Having performed time series estimation using a deep neural network with two Dense hidden layers, an LSTM network, a special type of recurrent neural networks, is created next as these networks due primarily to having feedback systems are comparatively more suited and much better for time series estimation than standard deep neural networks. As was previously performed, one of two transaction datasets for Murray is used for this estimation technique. A function previously constructed is called upon to appropriately reform the transaction data for its use in the network. The model for an LSTM network is created using a single LSTM layer and a single Dense hidden layer, both employing the ‘relu’ activation function. The model is compiled using the ‘Adam’ optimizer, and the model is trained using a batch size of 1. After the training process, predictions are made using the model for both training and testing datasets to evaluate the model performance. As can be seen from the figures, the model’s performance for both training and testing datasets could further be improved. However, it is clearly seen from the model’s performance on the training dataset, there are certain periods when the model’s predictions exactly match the training dataset. Nevertheless, it is also clearly seen that the presence of significant variance in both training and testing datasets makes it difficult for the model to learn the underlying patterns effectively. As was previously done, a seasonal price data is extracted from the first data set and used with the same module to see if further improvements could be achieved. After the model is trained, and predictions are made, it becomes clear that with seasonal data, the model’s training performance is noticeably much better, which could be attributed to the smaller training dataset. The testing performance of the model, however, needs further improvement through mainly

increasing the quality and thus reliability of the transaction data. The following figure represents predictions made by the trained LSTM model on both training and testing datasets.

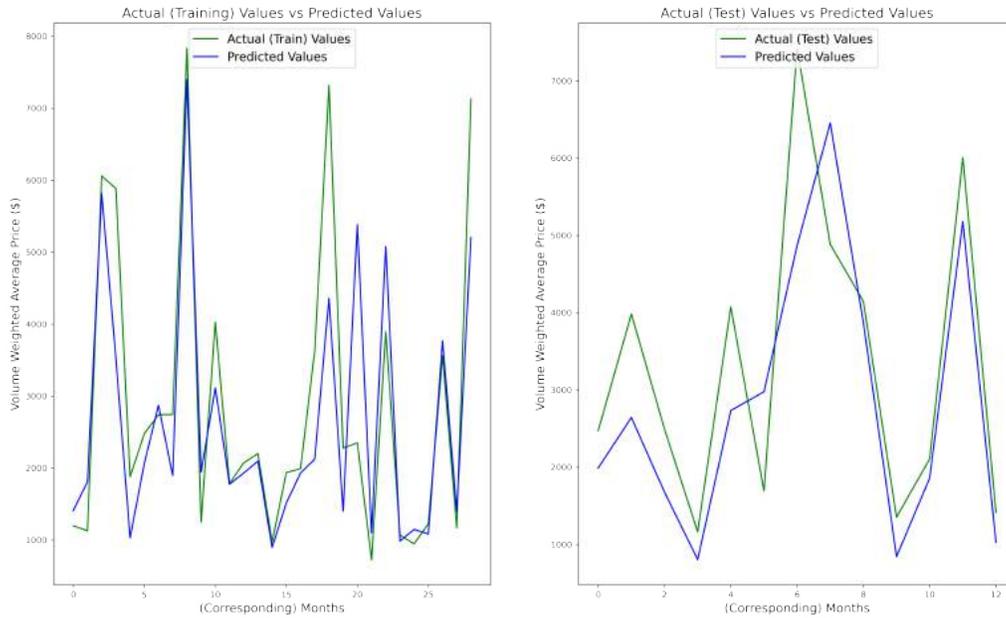


Figure 8: Water Allocation Price Prediction by an LSTM model

3.9 Time Series Estimation of Volume Weighted Average Prices via a Bayesian Neural Network

Time series estimation techniques employed so far make only single-point estimates owing to the fact that weights of the models using those techniques are single-point values. However, in a Bayesian neural network, in comparison to previously defined time series estimation models, weights are probability distributions, providing us with a valuable opportunity to decrease the likelihood of over-fitting during the training process as well as to quantify uncertainty of the model, which can be categorized as epistemic uncertainty and aleatoric uncertainty. Epistemic uncertainty is attributed to the fact that the model may not have enough information to use during its training process. However, this information can be obtained from providing new data to the model or increasing the representation capacity of the model by increasing its complexity. Importantly, this type of uncertainty can potentially be addressed and reduced. For the case under consideration, although the raw water market data contains a huge number of transactions, the prediction model is designed to predict volume weighted average allocation prices, owing to the fact that it is a usual practice and would not make much sense trying to estimate price of each potential transaction in the future. However, it shall be noted that the approach of calculating average prices for each month of each year in the data decreases the number of data points significantly, which in turn is very likely to result in an increase in the model's epistemic uncertainty. On the other hand, aleatoric uncertainty stems from the inherent nature of the data-generating process, simply meaning that for certain processes, out of a number of parameters, only a subset of these is observable and can be used in its description. Therefore, theoretically speaking, a potential way to effectively measure all these parameters would be an enabler to reproduce an event exactly without aleatoric uncertainty, which is not the case at all in most real scenarios.

A Bayesian neural network for the current case is created by replacing Dense hidden layers of a deep neural network employed for time series estimation with two Dense Variational layers of the same unit size. Furthermore, the output layer is designed as a distribution layer, connected to a Dense Variational layer with a single unit, which will yield the mean of the distribution estimated by the model. In addition to this, in order to employ Dense Variational layers in the model, prior and the posterior distributions of the model weights shall also be defined beforehand. The model is then complied with the ‘Adam’ optimizer and the loss function of negative likelihood, which is a usual loss function employed in the Bayesian networks with the intention of computing the likelihood to see the true data from the estimated model distribution. As can be observed from the first figure below, while the results obtained from the use of whole transaction data reveal that the model could further be improved mainly by increasing the reliability of the data, the results for the seasonal data show really satisfactory results for both training and testing datasets. It is worth noting that compared to previous models, predictions made by the Bayesian neural network are different each time the model is asked to make them, due to the above-mentioned fact that weights of the model are not single-point estimates but rather probability distributions. It could also be easily observed from the second figure. Owing to the fact that predictions made by the model are different each time, 100 predictions are made, and each one is plotted on the figure. It is worth noting that this behavior by the probabilistic Bayesian model simply displays that the model itself is not certain in its predictions, which provides us with a great opportunity to quantify its uncertainty.

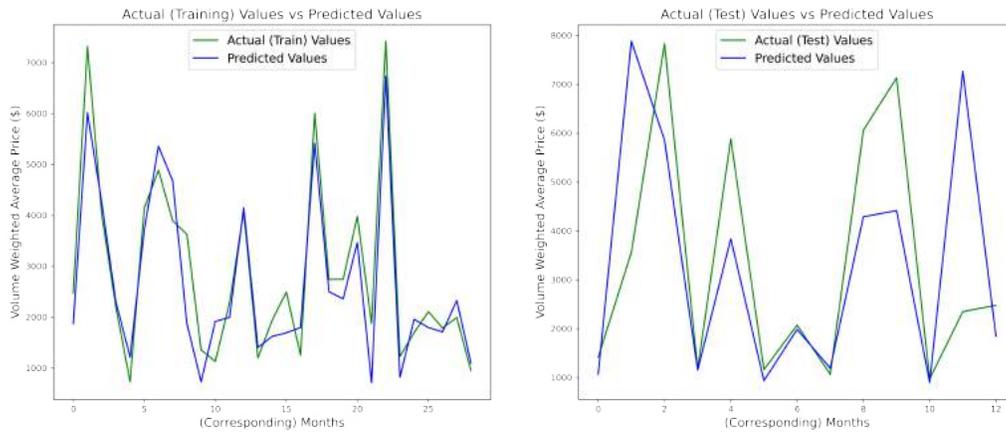


Figure 9: Water Allocation Price Prediction by a BNN model

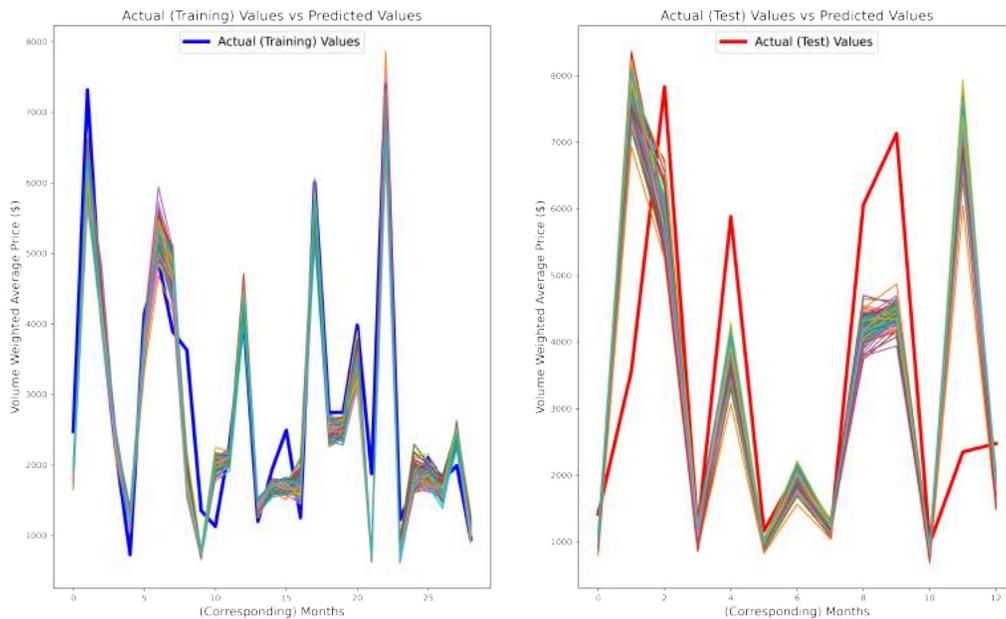


Figure 10: Water Allocation Price Prediction by a BNN model

3.10 Time Series Estimation of Volume Weighted Average Prices via a TensorFlow Structural Time Series Modelling

This forecast model developed employs TensorFlow Probability's Structural Time Series (STS) framework to forecast volume weighted average transaction prices. What makes structural time series models different to the previously employed models is the fact that these models are usually expressed as a sum of components, such as linear trend, cycles, seasonal patterns, and residuals.

As just mentioned, to create a structural time series model, some of the above-mentioned components shall be observed from the time series prediction problem of interest in order for the model to provide satisfactory estimates. Therefore, before building an STS model, it is essential to decompose the water allocation price dataset into its internal parts, which is achieved by using the 'statsmodels' library.

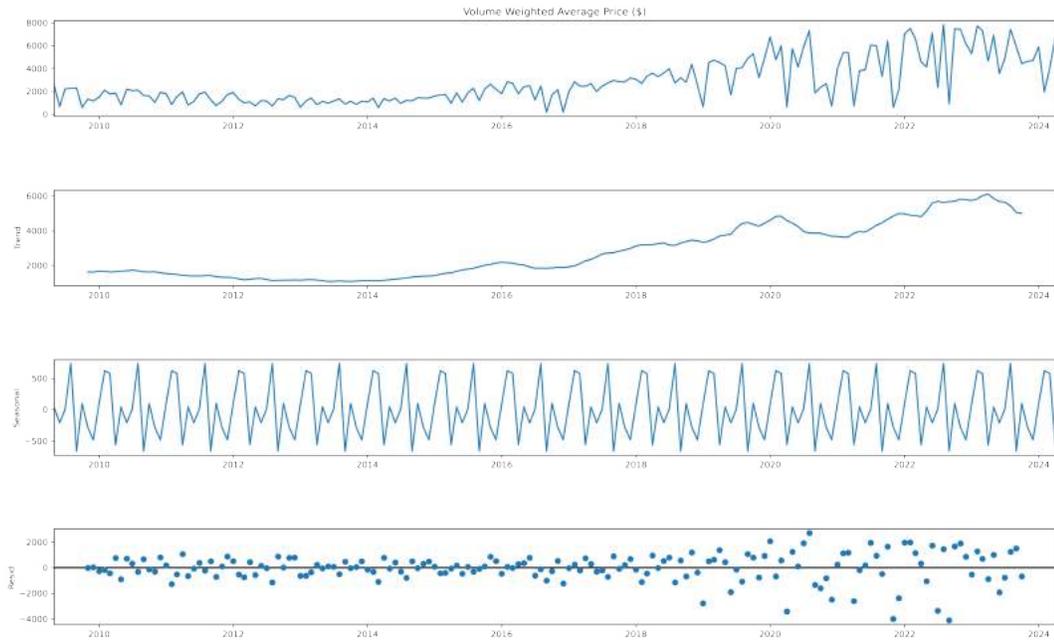


Figure 11: Trend Decomposition of Price Time Series Data

From the constituent parts of the price data, a linear pattern can be identified despite the fact that the relationship is not very linear. Further inspection also reveals the existence of a seasonal pattern. However, the residuals from the decomposed time series reveals that the variance is not constant and keeps increasing along the x axis, which is not a favorable condition for the desired prediction model.

So, based on the fact that linear and seasonal patterns from the price dataset are easily observable, these two parts will be used to construct an STS model. Once the model is created and the price dataset is divided into training and test sub-datasets, the model is fitted to the observed time series, which is the training dataset.

The following figure represents predictions made by the model using the testing dataset. It is apparent that the results are not satisfactory. It could be attributed to the fact that the 'residuals' component of the price time series data with changing variance was not used while creating a model structure.



Figure 12: Water Allocation Price Prediction by TensorFlow Structural Modelling

3.11 Pruning a DNN Network using a Multi-Armed Bandit Algorithm

A DNN model created above for time series estimation had two Dense layers. In order to employ a multi-arm bandit algorithm to prune the network by removing unnecessary weights, the DNN model was selected for the sake of simplicity. The algorithm employed for this pruning technique have the following primary assumptions.

The first assumption is that each weight of the trained model is representative of a single arm, and setting the value of a weight to zero is considered to be identical to choosing an arm. The second assumption is that a reward is determined by the difference between performance scores of the network before and after the application of the pruning technique (removing a weight).

In order to employ this technique, the DNN model with the same previous structure is created and trained. Afterwards, predictions are made by the model using both training and testing datasets for the future comparison of the model performance. Then, the technique is employed in order to prune 10 percent of the model weights. Having performed this step, the model with new weights are saved and then used for making predictions using again both training and testing datasets. The following two figures depict the results of the technique. It is clearly seen that after pruning, the model's performance on the training dataset noticeably decreased, which as expected affected its performance on the testing data as well. So, with a higher level of pruning (25 percent, 50 percent, etc.), the model's performance would significantly decrease, emphasizing the significance of each weight for the model. It could be attributed to the fact that high variations in the dataset, as stated above, make it very challenging for the model to learn the relationship of interest. Therefore, a satisfactory performance by the model trained using such datasets is very prone to any removal of its weights.

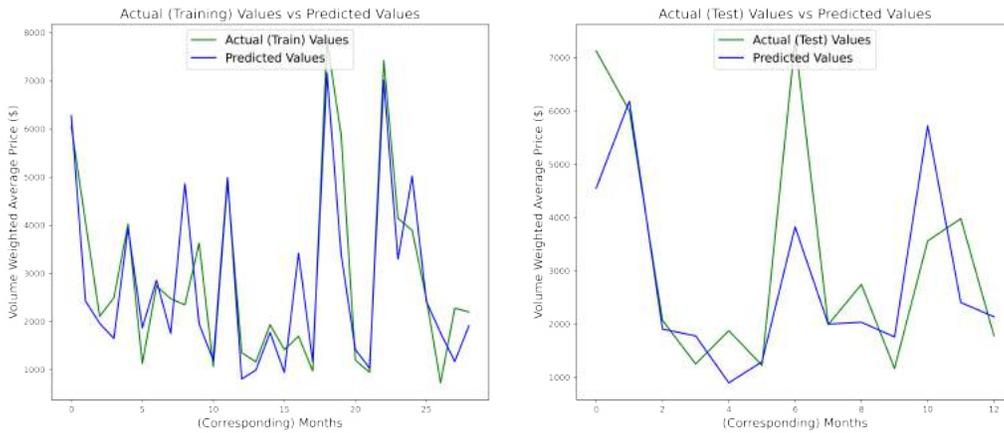


Figure 13: DNN Model Performance before Pruning

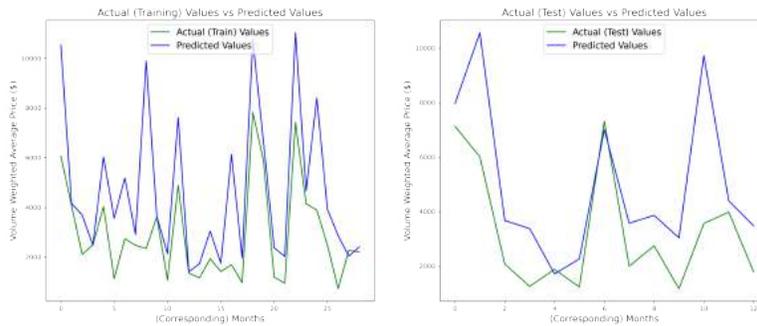


Figure 14: DNN Model Performance after Pruning

3.12 ARIMA Model Parameter Selection

We use the Bayesian information criterion when selecting ARIMA parameters (p,d,q) . The heatmap below shows the matrix of p and q that has the lowest BIC and AIC values respectively. As expected, orders that are higher generally produce better results at the expense of increased computational resources. We opt to use order $(10,2,10)$ in most of our models, as this proved most effective on tests.

It's important to note that the data is not very stationary, according to the ADF test. However, we perform differencing of the data for the basic ARMA model to enhance stationarity

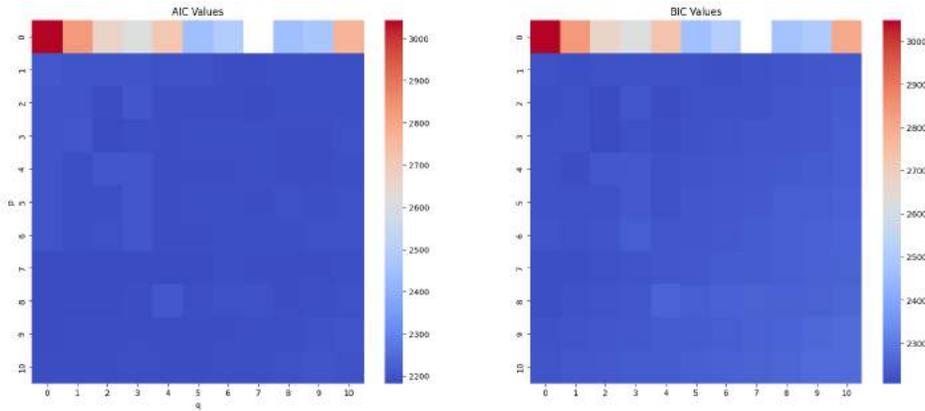


Figure 15: AIC and BIC Parameters for orders 1-10

3.13 Basic ARMA

The ARMA model combines two major components: the autoregressive (AR) part, which forecasts future behavior based on past values, and the moving average (MA) part, which models the error of the prediction as a combination of past errors. This combination is designed to capture a more accurate representation of time-series data.

We attempt to analyze future trends of the Water data based on simple auto-regressive moving average (ARMA) models. Here, we utilize 90% of the dataset for training purposes and reserve 10% for testing the model's performance.

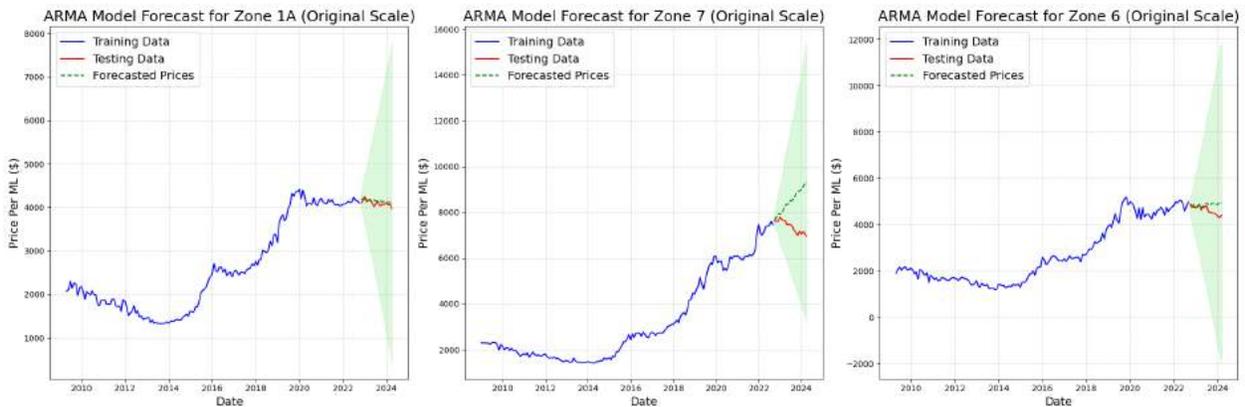


Figure 16: Differenced ARMA model with parameters AR: 10, MA: 10

Zone	MSE	MAE	R^2
Zone 1A	4134.673121	55.368150	-0.033342
Zone 7	32028.428839	161.936527	-1.975863
Zone 6	10367.589659	83.522934	0.175122

Table 2: Performance metrics for different zones

It can be seen that while the model predicts data in zone 1A with a very low R^2 , other zones do not share

the same accuracy.

3.14 Rolling ARMA and SARIMAX Models

The models explored in previous sections were adapted for later use in section 4. This was done by fitting a new model to each timestep. This simulates the process of updating a model with new data each month. The training dataset was thus the timesteps prior to the 'start' timestep and the model

For each time step

1. Train a model on the previous timesteps ('known' points)
2. Forecast the values for the remaining timesteps

The difference between the starting timestep and an individual prediction is here termed as an 'offset' or 'steps from start.' For example, let there be a model trained on the data up until $t=30$. That model's prediction for $t=46$ would have an offset of 16. As the offset of a prediction increases, naturally so does the uncertainty and large errors are produced. Managing and accounting for these errors is explored further in section 4.

We thus contrast the performance of the ARMA (Using price data alone) and SARIMAX (using exogenous variables as mentioned). We thus plot the predictions, with the error as the timescale increases:

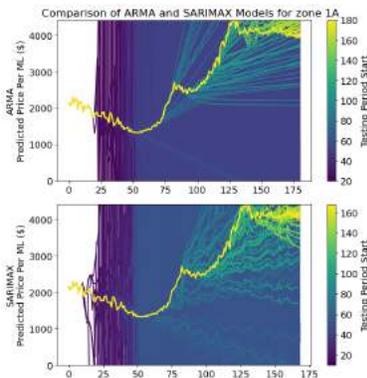


Figure 17: Zone 1A Rolling Comparison

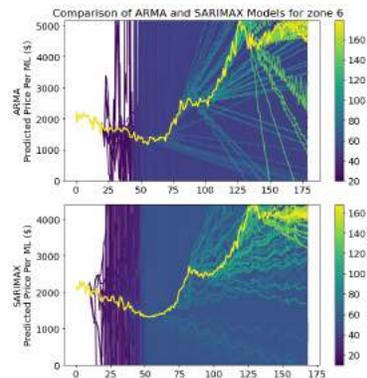


Figure 18: Zone 6 Rolling Comparison

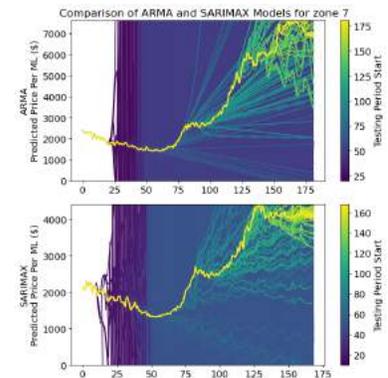


Figure 19: Zone 7 Rolling Comparison

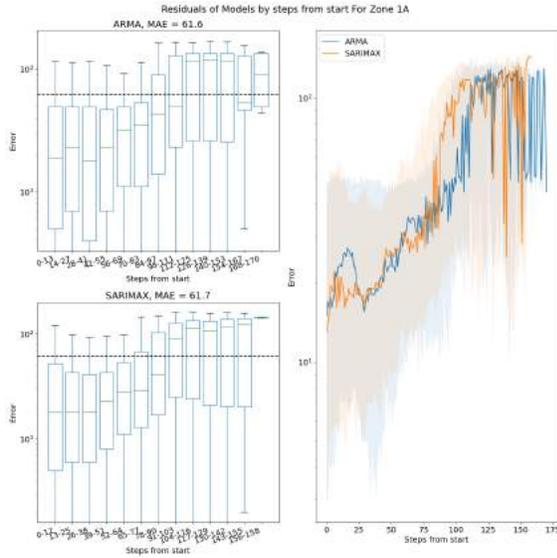


Figure 20: Zone 1A Error vs steps from start (months predicted)

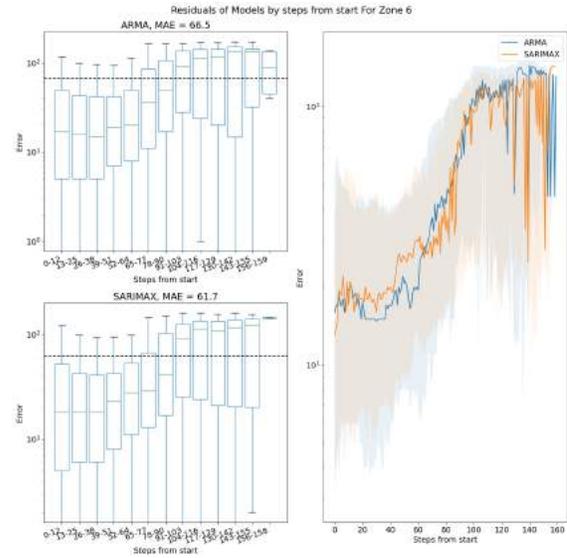


Figure 21: Zone 6 Error vs steps from start (months predicted)

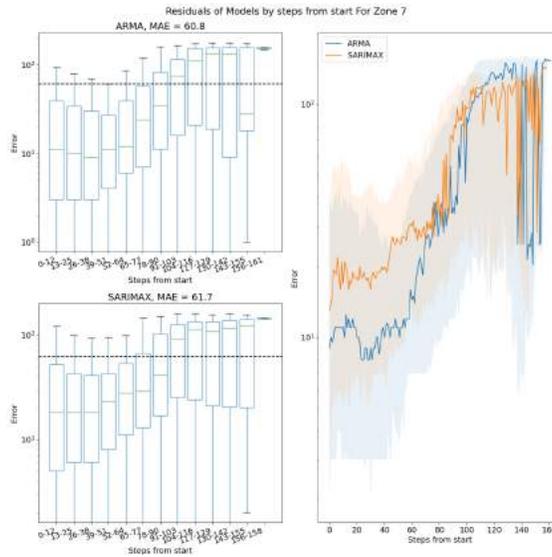


Figure 22: Zone 7 Error vs steps from start (months predicted)

We can see that in general, the MAE decreases, the less months are needed to be predicted. This lines up with expectations, as we expect the model to be more accurate, the shorter timescale we predict. However, unexpectedly, the ARMA configuration performs similarly, or even better than SARIMAX. This shows the exogenous data may not be such a good predictor of price data, in its current form.

3.15 MOMENT Model Prediction

MOMENT (Multi-Objective Multivariate Evolutionary Neural Network) is a newly developed pre-trained time series prediction model that utilizes neural networks to forecast future values in multivariate time series

data [8]. Zone 6 is excluded due to insufficient data points. The following values were obtained:

Metric	Zone 1A Value	Zone 7 Value
MSE	0.9681	1.3784
MAE	0.7498	0.7883
R^2	0.4010	0.3430

Table 3: Comparison of Metrics for Zones 1A and 7

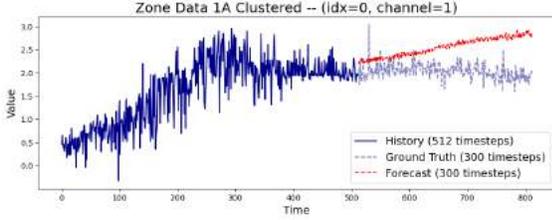


Figure 23: Model for last 512 points for zone data 1A

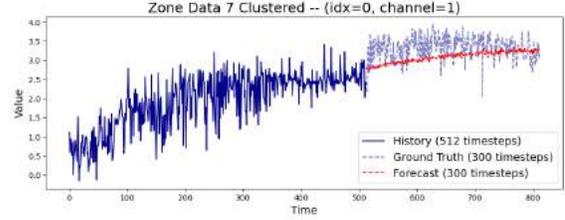


Figure 24: Model for last 512 points for zone data 1A

4 Optimisation - Timing of water purchases

The objective of this section is to develop an approach to decide the volume of entitlement to be purchased at a point time given a goal volume.

4.1 Simple Formulation

The objective function is shown in equation 4

$$\min_{V \in \mathbb{R}^+} \text{Cost}(V, x) = \sum_{t=0}^n V_t \times \text{Price}(t, x) \quad (4)$$

subject to

$$\begin{aligned} V_t &\geq 0 \forall t \in n \\ V_t &\leq \text{Volume}(t, x) \forall t \in n \\ \sum_{t=0}^n V_t &= V_{total} \end{aligned} \quad (5)$$

This can be solved as follows:

1. Find the time with the lowest price

$$t_{i^*} \leftarrow \arg \min_t \text{Price}(t, x)$$

2. Purchase as much as possible or needed

$$V_{t_{i^*}} \leftarrow \min(\text{Volume}(t_{i^*}, x), V_{total} - \sum V)$$

3. Repeat until sufficient volume acquired

$$\sum V = V_{total}$$

However this does not consider the uncertainty of price and volume predictions.

4.2 Rolling Optimisation

The price and volume predictions will decrease in accuracy at further timesteps. They can be more accurately described as

$$\text{Cost}(V, x, d)$$

$$\text{Price}(t, x, d)$$

Where d represents the difference between the time and prediction start time d .

In a rolling optimisation approach, the optimisation is conducted continually to make a decision.

1. Optimise equation 4 to find V_T
2. Update the volume needed and repeat until volume found for all timesteps $T \in n$

4.2.1 Uncertainty Correction

As a model predicts a value in a time-series further out from the training period, the more uncertain the prediction is. Thus, the previous predictions for a time-step should become more accurate as the starting time approaches. It may then be possible to account for uncertainty through regression.

This can be implemented by performing linear regression relating previous predictions for a given time to the starting time of the prediction, shown in equation 6.

$$\text{Price}(t, x, d) \approx m_t d + c_t \tag{6}$$

Since the true value is to occur when the starting time is after after the time of prediction, this relation can be used to model a more accurate prediction given by $m_t(t + 1) + c_t$.

4.3 Results

The optimisation was performed using outputs of ARIMA models (developed in section ??) of order (10, 2, 10) for the monthly median of price and sum of volume of entitlement trades. The target volume was set to 80% of the available volume. This is an arbitrary target chosen for its achievability.

The rolling optimisation was not able to reproduce the theoretical optimal chain of decisions (the algorithm run with true values rather than predictions). As seen in figure 25, the rolling optimisation algorithm was not able to recognise the trend of increasing price and thus was not able to take advantage of it. However when the rolling algorithm used predictions corrected with regression as described in 4.2.1, it was able to more accurately predict the trends in price and purchase entitlements at the beginning while they were cheaper.

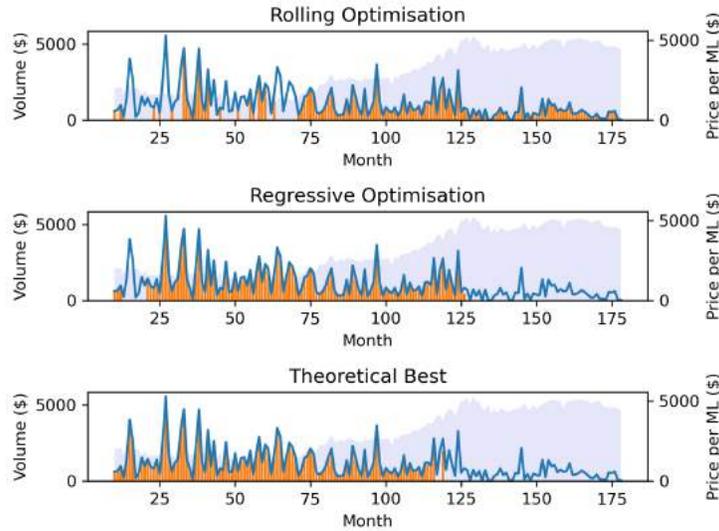


Figure 25: Comparison of volume purchase decisions by simple rolling optimisation algorithm and the theoretical optimum. The line represents the volume available at each time and the bar represents how much is to be purchased.

A random approach was used as a point of comparison for performance evaluation. If the model provides similar performance to random, then the algorithm is not useful. The rolling optimisation approach tended to perform better than random as seen in figure 26 however there is some overlap. Thus there is need to improve the model.

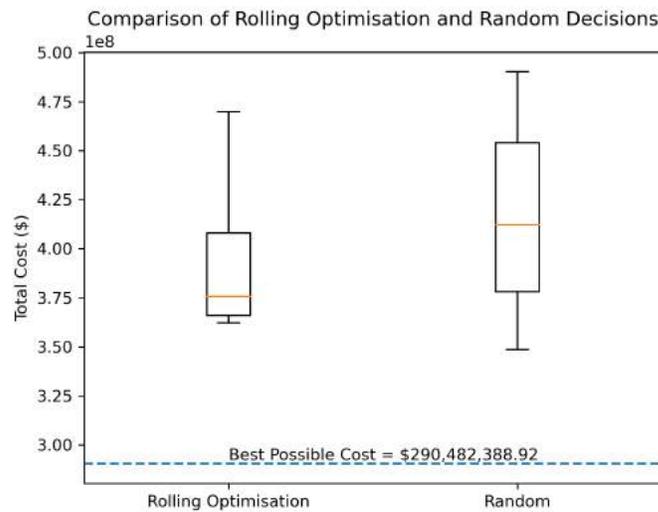


Figure 26: Comparison of distribution of total cost for 100 realisations optimised with predicted price or random values. For these values, the rolling optimisation used price and volume predictions sampled from a random distribution based on the confidence intervals of the prediction.

4.4 Continued Optimisation

4.4.1 Maximum Likelihood

The errors can be modelled as a normally distributed random variable which increases in spread as the predictions lie further from the starting point.

$$\epsilon \sim N(0, \sigma(l)) \quad (7)$$

The relationship between the standard deviation and offset can be estimated using maximum likelihood estimation.

The likelihood of observation x given predictions μ is given by equation 8.

$$\mathcal{L}(\sigma) = \prod_{i=0}^n Pr(x_i|\mu_i, \sigma) = \prod_{i=0}^n \frac{1}{\sigma(l_i)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i - \mu_i}{\sigma(l_i)}\right)^2} \quad (8)$$

A log-transformation eases this problem for both analytic solutions and computational. From a computational perspective, the log-transform reduces the errors introduced by precision loss.

$$\ell(\sigma) = \sum_{i=0}^n \ln Pr(x_i|\mu_i, \sigma) = -\frac{n}{2} \ln 2\pi - \sum_{i=0}^n \ln \sigma(l_i) - \frac{1}{2} \sum_{i=0}^n \left(\frac{x_i - \mu_i}{\sigma(l_i)}\right)^2 \quad (9)$$

4.4.2 Updating Volume Predictions

The standard deviation for each lag is assumed to be equal. Thus the following equation can be used to find the standard deviation for each lag which maximises likelihood of observations.

Equation 10 can be simplified when estimating the standard deviation associated with each lag l . This can then be formulated as an optimisation problem:

$$\hat{\sigma}_l = \arg \max_{\sigma_l} \ell(\sigma(l_i)) = \arg \max_{\sigma_l} -\frac{n}{2} \ln 2\pi - n \ln \sigma_l - \frac{1}{2\sigma_l^2} \sum_{i=0}^n (x_i - \mu_i)^2 \quad (10)$$

$$\frac{\partial \ell}{\partial \sigma_l} = -\frac{n}{\sigma_l} + \frac{1}{\sigma_l^3} \sum_{i=0}^n (x_i - \mu_i)^2 \quad (11)$$

For the function to be concave, the second derivative must be non-positive. However this is not the case.

$$\frac{\partial^2 \ell}{\partial \sigma_l^2} = \frac{n}{\sigma_l^2} - 3 \frac{1}{\sigma_l^4} \sum_{i=0}^n (x_i - \mu_i)^2 \quad (12)$$

$$\frac{n}{\sigma_l^2} \leq 3 \frac{1}{\sigma_l^4} \sum_{i=0}^n (x_i - \mu_i)^2$$

$$\sigma_l^2 \leq \frac{3}{n} \sum_{i=0}^n (x_i - \mu_i)^2$$

Which gives an interval of

$$0 < \sigma_l \leq \sqrt{\frac{3}{n} \sum_{i=0}^n (x_i - \mu_i)^2}$$

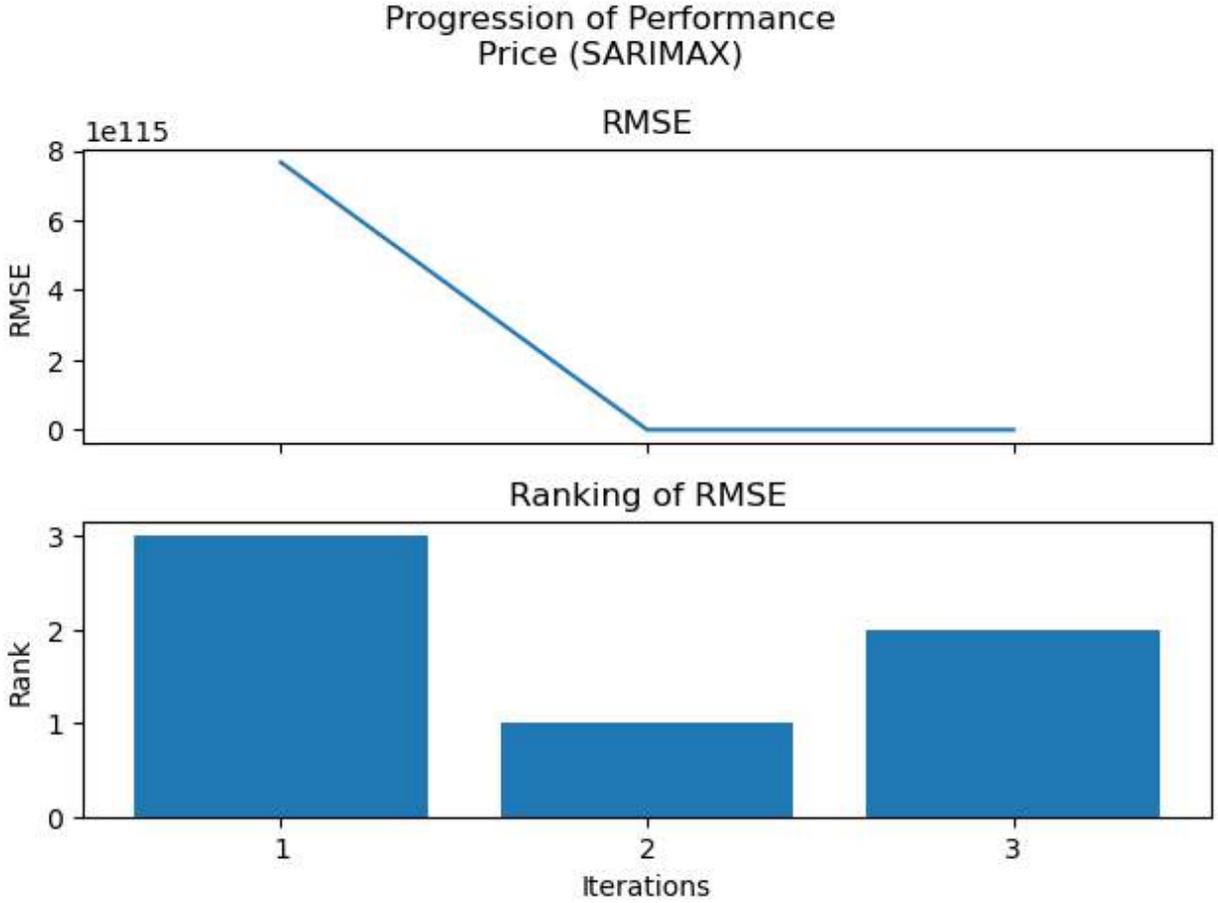


Figure 27: The change in MSE of predictions of volume after expectation maximisation

Solving equation 11 for $\frac{\partial \ell}{\partial \sigma_i} = 0$ gives:

$$\hat{\sigma}_i = \sqrt{\frac{\sum_i^n x_i - \mu_i}{n}} \quad (13)$$

Which sits in the derived interval and is thus a local optimum.

The prediction can then be maximised with the following equation

$$\hat{\mu} = \sum_i \frac{x_i}{\sigma_i^2} \div \sum_i \frac{1}{\sigma_i^2} \quad (14)$$

These can be repeated and alternated as a form of expectation maximisation.

1. Fit a standard deviation for each lag based on predictions and observations
2. Fit a new prediction based on the new standard deviation

This process was applied to both price and volume however only volume saw a continued improvement with further iterations, shown in figures 27 and 28

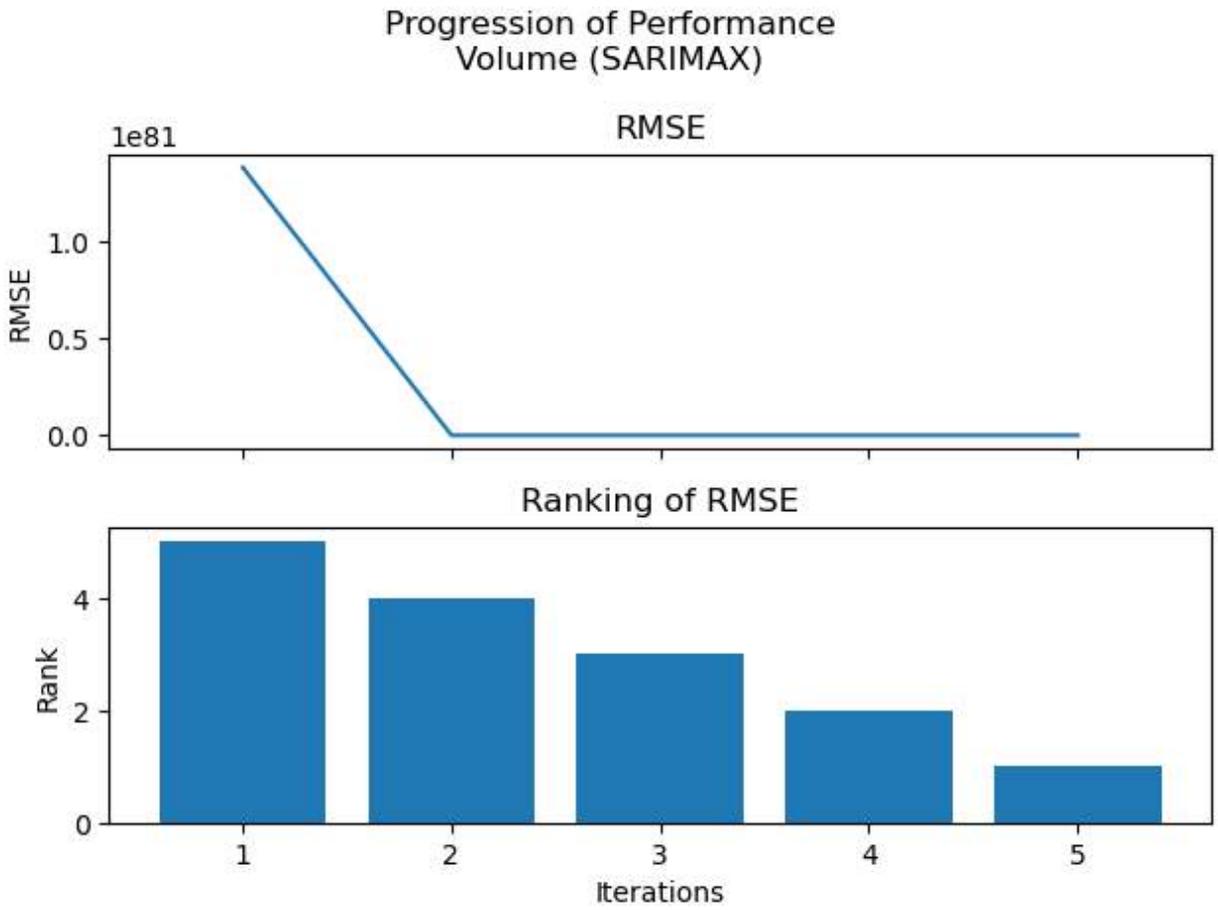


Figure 28: The change in MSE of predictions of price after expectation maximisation

	a	b	c
Differential Evolution	579.814019	3.627484e+06	-3.627234e+06
Nelder-Mead 1	579.820887	4.239408e+06	-4.239159e+06
Nelder-Mead 2	579.820887	4.239408e+06	-4.239159e+06

Table 4: Table of results for sequential optimisation algorithm runs maximising log likelihood of standard deviation - offset polynomial

4.4.3 Improving Price Predictions

The price was improved by fitting the standard deviation to a polynomial. This was achieved using a series of optimisation functions from the scipy optimisation library.

1. Differential evolution (a global optimisation algorithm)
2. Nelder-Mead (a local optimisation algorithm)

The order allowed for exploration followed by exploitation. Differential Evolution algorithms are able to quickly explore a bounded solution space and iterate through possible solutions using techniques inspired by genetics and natural selection. Nelder-Mead does not require bounds and works from one starting point and explores the space using a simplex of points however as a local algorithm it is more likely to become stuck in a local optimum. Thus, the best point found by differential evolution can be efficiently improved upon by passing it as the starting point to Nelder-Mead.

The standard deviation was modelled as a quadratic function of the offset. The coefficients are given by \vec{w}

$$\sigma(l) = \sum_{i=0}^3 w_i l^i \quad (15)$$

All of the data was included, including points in the future. This is unrealistic in how the approach may be applied but serves as an exploration of the best case.

The path of the nelder-mead optimisation run is shown in figure 29. The objective function was evaluated over a grid for the purposes of visualisation. It is important to note that the other parameter is kept at its optimum value for this grid evaluation and thus the contour does not match exactly the evaluations made by each iteration of Nelder-Mead.

The improved predictions and standard deviation estimate can be used to predict the likelihood of whether the price will be lower or higher than at the decision time.

The results from each run can be found in table 4.

From these parameters, the probability that a price in the future will be higher than the price at the start of predictions can be modelled.

The positive constraint can be modelled by finding the probability of the price being positive and using it as a divisor. Since the values are positive this is equivalent to the conditional probability.

$$Pr(x = X|X > 0) = \frac{Pr(X > 0|x = X)Pr(x = X)}{Pr(X > 0)} \quad (16)$$

4.5 Applying to Decision Making

The improved estimates of volume were used in the decision making algorithm. The probability of the higher price is used in 3 strategies.

1. Using the probability against a threshold, if it is $\geq 50\%$ then it is predicted to be higher

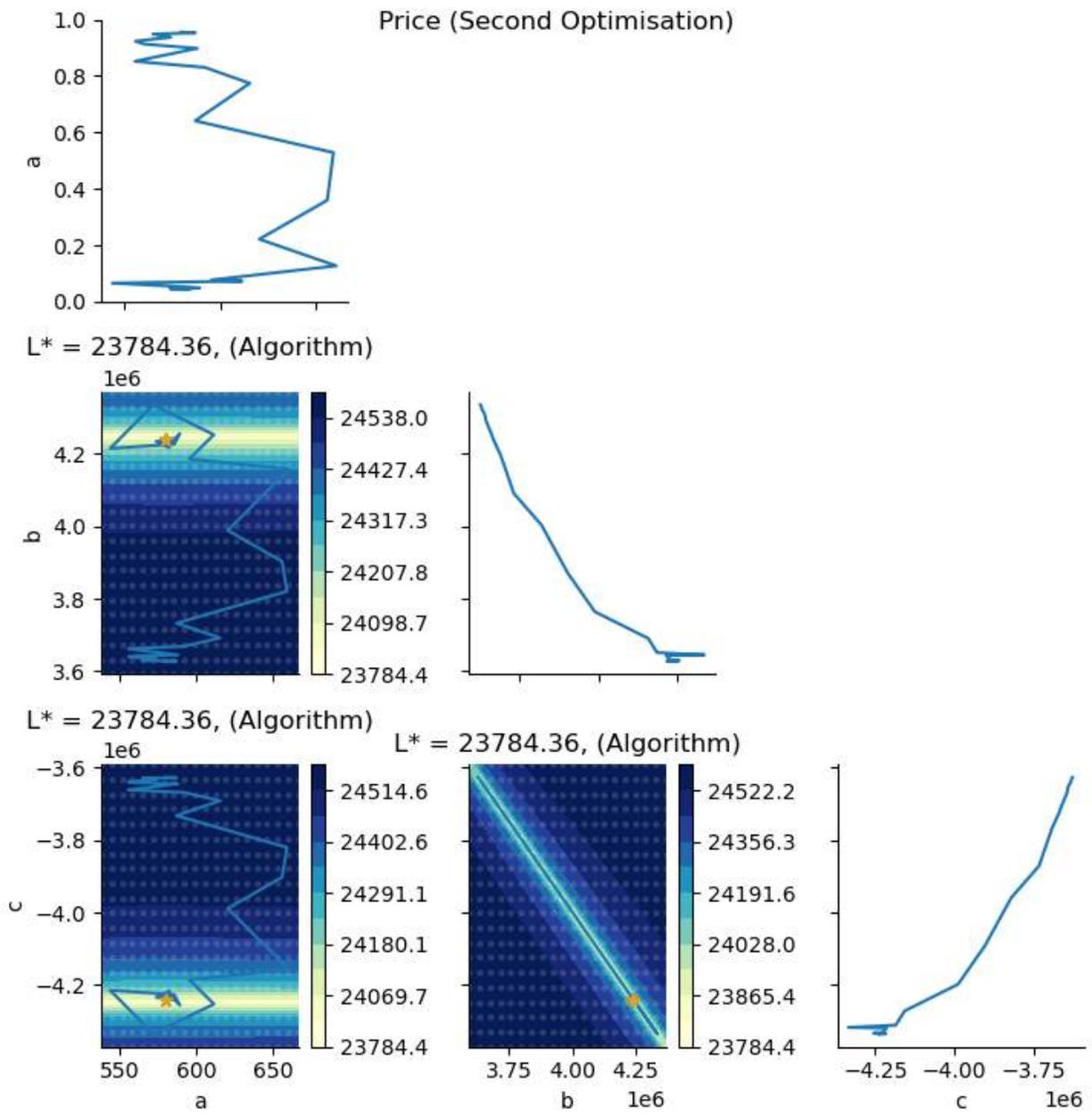


Figure 29: Graph showing the optimisation path of a Nelder-Mead simplex algorithm for fitting the weights in eq 15. 50 points were sampled in the axis area to visualise the values of the objective function (negative log likelihood) and are seen as a contour map. The diagonal plots show the objective function values associated with each parameter value.

2. Predicting the total volume available in the future
3. A combination of the above, applying the second when the threshold was not met to purchase the full volume.

The distribution of the sum of normal distributions follows the relationship

$$\sum_i N(\mu_i, \sigma_i) = N\left(\sum_i \mu_i, \sum_i \sigma_i\right) \quad (17)$$

To take into consideration the uncertainty of whether it would be cheaper in the future, the predictions μ were multiplied by the probability the price would be higher.

The results of these approaches, along with the others, can be found in fig 33.

4.5.1 Further Work

The decision making algorithm should be improved by

1. Implementing other forecasting models, potentially correcting for more granular trends and then modelling smaller scale processes.
2. Investigation of joint distribution methods to analyse the relationship between price and volume as well as between predictions.

5 Optimisation 2 - Optimising Portfolio Allocation

Often, there is a trade-off for irrigators (farmers) whether to invest in other assets such as water, stocks, or farm equipment. There is also a similar trade-off for government entities of whether to buy water allocation for environmental needs, or spend money on other public services. In economic theory, the primary way to measure risk-adjusted returns, and hence allocate assets to a portfolio is via the Sharpe ratio. This section investigates the optimal weighting of water shares for each of the trading zones for a portfolio, compared to the ASX200 based on data from time period of 2018-2023.

5.1 Problem Formulation

Objective: Maximize the Sharpe Ratio of the portfolio. By maximising this ratio, the highest possible return for risk can be quantified:

$$\max_w \frac{R_p - R_f}{\sigma_p} \quad (18)$$

where:

- R_p : Expected return of the portfolio:

$$R_p = \sum_{i=1}^N w_i R_i \quad (19)$$

with the weighted average of expected return of asset return R_i

- R_f : Risk-free rate of return (Set at 2.04% - based on Ausgrid rate of return [3])

- σ_p : Standard deviation of the portfolio's return, defined by the variance:

$$\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} \quad (20)$$

with σ_{ij} as the covariance between the returns of assets i and j .

- w : Asset weights in the portfolio

Constraints:

- The weights must sum to 1.

$$\sum_{i=1}^N w_i = 1 \quad (21)$$

- The weights must be non-negative, we assume borrowing to sell is not allowed.

$$w_i \geq 0, \quad \forall i \in \{1, 2, \dots, N\} \quad (22)$$

5.2 Primal Problem Formulation and data synthesis

Substituting the expected return, risk free rate and the covariance, we are able to formulate the Lagrangian, where λ and μ_i are the Lagrange multipliers for the equality and inequality constraint respectively.

$$\mathcal{L}(w, \lambda, \mu) = - \left(\frac{\sum_{i=1}^N w_i R_i - R_f}{\sqrt{\sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}}} \right) + \lambda \left(\sum_{i=1}^N w_i - 1 \right) + \sum_{i=1}^N \mu_i (-w_i) \quad (23)$$

where λ and μ_i are the Lagrange multipliers associated with the sum of weights constraint and non-negativity constraints, respectively.

5.2.1 Data Synthesis

Four price datasets from the year starting 2018 to the year ending 2024 (5 years) were compared. The first was the ASX [4], and the other three was the cleaned and flattened water market data for each associated zone (Trading zones 1A, 7, 6) as described in section 3. The returns and covariance matrix are as described:

Zone	Value
ASX	0.00378477
Zone_1A	0.00488902
Zone_6	0.00709633
Zone_7	0.00842891

Table 5: Returns % per month

	ASX	Zone_1A	Zone_6	Zone_7
ASX	0.001585	-0.000139	-0.000081	-0.000161
Zone_1A	-0.000139	0.001817	0.000356	0.000574
Zone_6	-0.000081	0.000356	0.004357	0.000457
Zone_7	-0.000161	0.000574	0.000457	0.001749

Table 6: Covariance Matrix

5.3 Solving using Stochastic Gradient Descent (SGD) with vanishing gradient

Stochastic Gradient Descent (SGD) is an iterative optimization algorithm that can be used to find the optimal portfolio weights. 20 iterations were completed, with a step size of 0.1, and decay ratio of 0.98 were completed. Thus portfolio allocations were as follows. This result was unexpected at face value as zone 7 returns outperformed zone 6 returns:

Optimal Allocation	
ASX	0
Trading Zone 1A	0
Trading Zone 6	0.8499
Trading Zone 7	0.1871

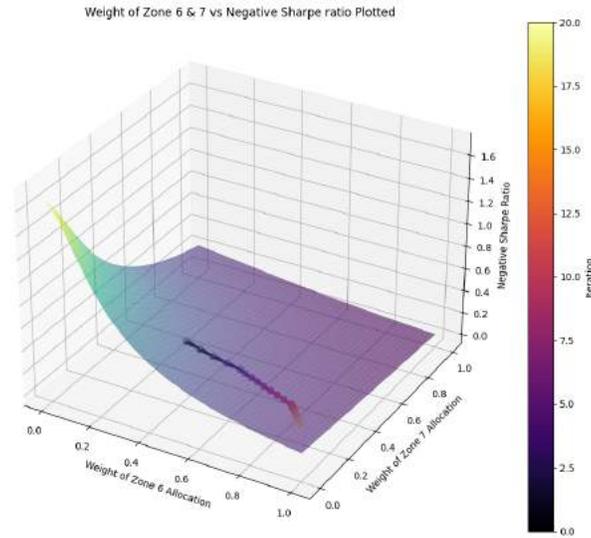


Figure 30: Sharpe Ratio vs weights of Zone 6 and 7

5.4 Solving using Newton's Method

Newton's Method uses the path of a quadratic equation to better approximate the gradient, rather than descending linearly like SGD, and uses the Hessian matrix to find the optimal solution.

We can see that the optimum allocation changes to favour zone 7, differing from SGD results. Additionally, the sharpe ratio increases after a certain point, signalling that it may have too large of a step size.

Optimal Allocation	
ASX	0
Trading Zone 1A	0
Trading Zone 6	0.0752
Trading Zone 7	0.9248

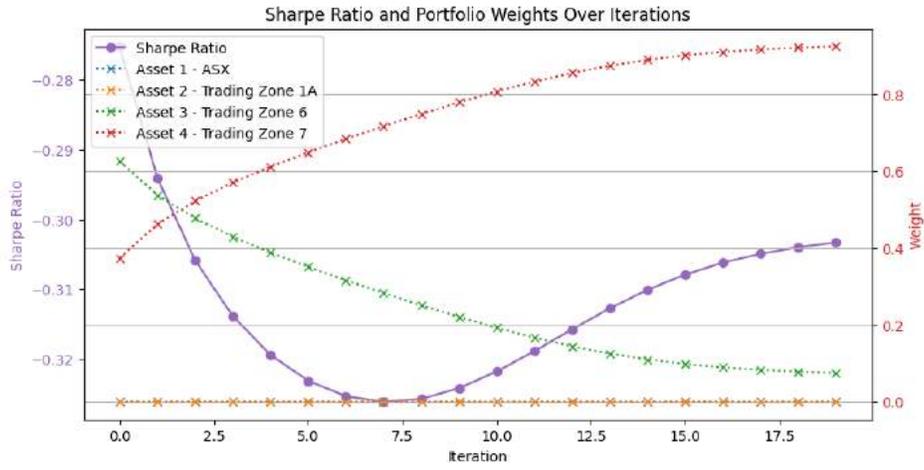


Figure 31: Newton Optimisation comparison (Minimising negative sharpe ratio)

6 Appendix A1

For exogenous variables, we source the data from every weather station online using the [A1] section script to fetch it using the water data online API



Figure 32: Weather stations data sourced from in Victoria from [5]

6.1 Data Preprocessing for daily data

The raw data from the weather stations often contains missing values, which is challenging for accurate analysis. To address this, we employ a two-step process:

Drop rows where more than half the columns have missing values. Given a data matrix X with n rows and m columns, we define a threshold $t = 0.5m$. For each row i , if the number of missing values $k_i > t$, we remove the entire row.

$$X_{\text{filtered}} = x_i \in X : k_i \leq t \quad (24)$$

Drop any columns with missing values. After the first step, we remove any remaining columns that contain even a single missing value. If x_{ij} represents the value in row i and column j , we remove column j if $\exists i : x_{ij} \text{ is NaN}$.

These steps ensure that the data is not imputed with guessed values for `pca`, even when there is less values to assess. Removing rows may result in fewer time data points, preserving spatial relationships but not trends. Removing columns preserves the time relationship but results in a small number of usable stations in Victoria.

References

- [1] Salem Ameen and Sunil Vadera. *Pruning Neural Networks Using Multi-Armed Bandits*. en. Sept. 2019. DOI: [10.1093/comjnl/bxz078](https://doi.org/10.1093/comjnl/bxz078). URL: <http://dx.doi.org/10.1093/comjnl/bxz078>.
- [2] Australian Bureau of Agricultural and Resource Economics and Sciences. *Economic effects of water recovery in the Murray–Darling Basin*. <https://www.agriculture.gov.au/abares/products/insights/economic-effects-of-water-recovery-in-murray-darling-basin>. Accessed on June 11, 2024.

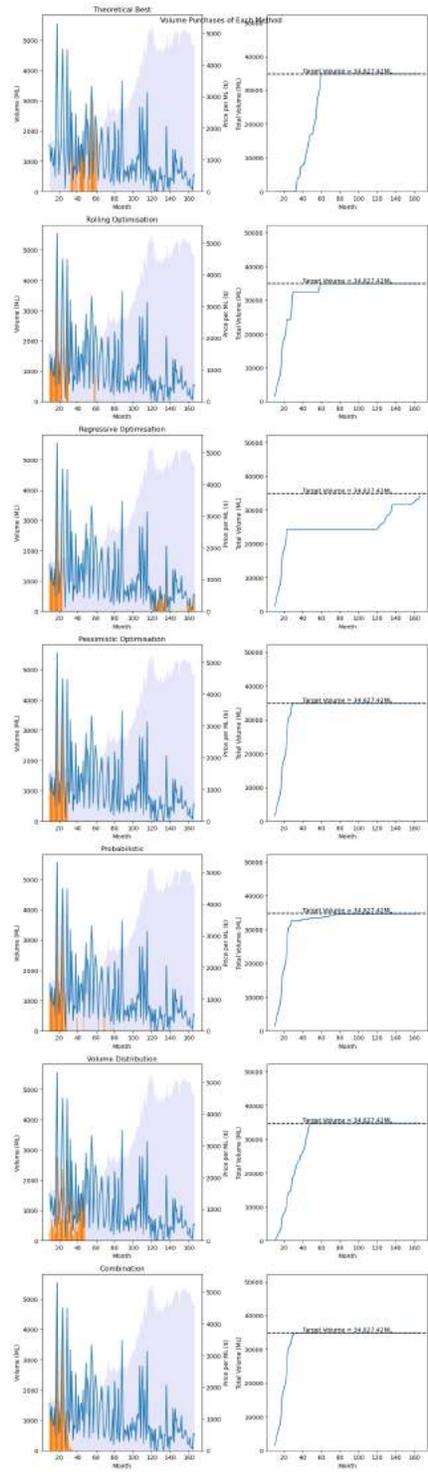


Figure 33: The results of each optimisation approach

- [3] Australian Energy Regulator. *Draft Decision Attachment 3 - Rate of return - Ausgrid - 2024-29 Distribution revenue proposal*. Tech. rep. Australian Energy Regulator, Sept. 2023. URL: <https://www.aer.gov.au/system/files/2023-10/AER%20-%20Draft%20Decision%20Attachment%203%20-%20Rate%20of%20return%20-%20Ausgrid%20-%202024-29%20Distribution%20revenue%20proposal%20-%20September%202023.pdf>.
- [4] Australian Securities Exchange. *Market Data*. 2024. URL: <https://www.asx.com.au/market-data/>.
- [5] Bureau of Meteorology. *Water Data Online*. Accessed: 2024-06-12. 2023. URL: <http://www.bom.gov.au/waterdata/>.
- [6] Tao Chen. “Round-number biases and informed trading in global markets”. In: *Journal of Business Research* (2018). URL: <https://doi.org/10.1016/j.jbusres.2018.07.027>.
- [7] Department of Agriculture, Fisheries and Forestry. *Australian Crop Report: Victoria*. Australian Government - Department of Agriculture, Fisheries and Forestry. Accessed: 2024-06-11. 2023. URL: <https://www.agriculture.gov.au/abares/research-topics/agricultural-outlook/australian-crop-report/victoria#daff-page-main>.
- [8] Mononito Goswami et al. “MOMENT: A Family of Open Time-series Foundation Models”. In: (2024). DOI: [10.48550/ARXIV.2402.03885](https://arxiv.org/abs/2402.03885). URL: <https://arxiv.org/abs/2402.03885>.
- [9] Tai Nguyen-ky et al. *Predicting water allocation trade prices using a hybrid Artificial Neural Network-Bayesian modelling approach*. en. Dec. 2018. DOI: [10.1016/j.jhydrol.2017.11.049](https://doi.org/10.1016/j.jhydrol.2017.11.049). URL: <http://dx.doi.org/10.1016/j.jhydrol.2017.11.049>.
- [10] Australian Bureau of Statistics. *Monthly Consumer Price Index Indicator*. tags: data. 2024. URL: <https://www.abs.gov.au/statistics/economy/price-indexes-and-inflation/monthly-consumer-price-index-indicator/latest-release#data-downloads>.
- [11] Victorian Water Register. *Water Entitlements: Entitlement Statistics*. Accessed: 2024-06-12. 2023. URL: <https://www.waterregister.vic.gov.au/water-entitlements/entitlement-statistics>.
- [12] Water Register Victoria. *Water Share Trading*. <https://waterregister.vic.gov.au/water-trading/water-share-trading>. Accessed on June 11, 2024.
- [13] Xiying Zhang et al. *Dry matter, harvest index, grain yield and water use efficiency as affected by water supply in winter wheat*. en. Aug. 2008. DOI: [10.1007/s00271-008-0131-2](https://doi.org/10.1007/s00271-008-0131-2). URL: <http://dx.doi.org/10.1007/s00271-008-0131-2>.