Water Supply

Water Supply Vol 24 No 10, 3540 doi: 10.2166/ws.2024.207

Investigating the influence of Australia Day and Christmas Day on water demand in the Greater Sydney region

Safa Ghannam [a,*] and Farookh Hussain[b]

[a] School of Professional Practice and Leadership, Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia
[b] Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia
*Corresponding author. E-mail: safa.ghannam@uts.edu.au

SG, 0000-0001-9626-0883

ABSTRACT

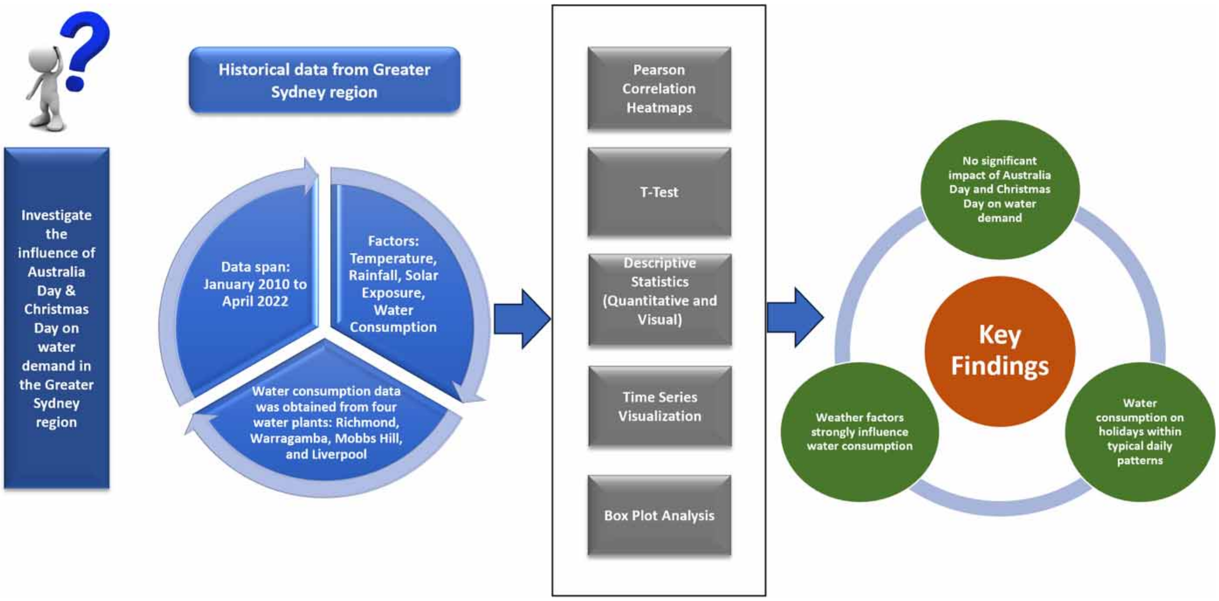
The objective of this research is to investigate whether particular occasions, such as Australia Day and Christmas Day, have a notable impact on water demand in the Greater Sydney region. By examining water demand during these events, the study aims to enhance understanding of water consumption patterns and contribute to the development of effective demand management strategies. Multivariate time series data from several water plants in the Greater Sydney region were analyzed using three methods: correlation heatmaps, *t*-tests, and descriptive statistics. The findings indicate that neither Australia Day nor Christmas Day has a significant impact on water demand at different water plants in the Greater Sydney region. These results suggest that public holidays may not need to be a critical factor in short-term water demand forecasting models for this area.

Key words: correlation, descriptive statistics, heatmap, *t*-test, water consumption, water demand

HIGHLIGHTS

- To investigate the impact of holidays, specifically Australia Day and Christmas Day, on water demand in the Greater Sydney region.
- To analyze variations in water consumption patterns during these specific occasions compared to typical days.
- To assess the statistical significance of any observed differences in water demand.
- To enhance understanding of water demand trends during holidays.
- To contribute insights for developing effective water demand management strategies in the Greater Sydney region.

GRAPHICAL ABSTRACT



1. INTRODUCTION

It is necessary to provide assistance to a system that operates in real time, including water utilities, to ensure they have the capability to meet water demand (Romano & Kapelan 2014). This study uses several methods, namely correlation heatmaps, *t*-tests, and descriptive statistics, to study the impact of Australia Day and Christmas Day on water demand in the Greater Sydney region. The descriptive statistics adopt both quantitative and visual approaches. We use a quantitative approach and comparative analysis to examine whether there are any variations in water consumption on these occasions. Additionally, we employ data visualization to investigate significant changes in water consumption patterns on public holidays. A correlation heatmap is utilized to analyze the correlation between water consumption and specific events, providing a comprehensive understanding of the relationship between water demand and these holidays.

Many researchers in the field of water demand forecasting have emphasized the critical importance of input selection for models, noting its significant impact on their performance. Choosing the right inputs is crucial for enhancing the accuracy and effectiveness of prediction processes. Zhou *et al.* (2000) developed a time series model for Melbourne, integrating seasonal, climatic, and autocorrelation factors, demonstrating robust predictive performance despite overestimation during public holidays.

This research is part of a broader collaboration with Sydney Water aimed at ensuring adequate water supply for specific suburbs. Building on previous work (Ghannam & Hussain 2023), our current research explores the influence of holidays such as Australia Day and Christmas Day on water demand. Beal & Stewart (2014) analyzed peak water demand patterns in South East Queensland, noting climate impacts and holiday effects.

Previous studies have shown that seasonal variations often have a more significant impact on water usage patterns than holidays. Guo *et al.* (2024) conducted an extensive analysis in Shanghai. The authors showed minimal holiday effects on water usage patterns compared to pronounced seasonal variations. Winter and spring patterns were consistent, whereas there were notable shifts in summer patterns during evening peaks, influenced by agricultural needs and high temperatures. Furthermore, Myka-Raduj *et al.* (2023) examined water consumption dynamics in a household setting over a year, highlighting significant variations during holidays and seasonal periods. Kavya *et al.* (2023) carried out a correlation heatmap analysis of the parameters influencing water consumption in Hubli city. The authors found that temperature and hour of the day have the maximum impact, showing a positive correlation with water consumption. However, holidays were found to have a Pearson correlation coefficient of 0.016 with water consumption.

Capt *et al.* (2021) developed an urban water demand model for El Paso, Texas, incorporating a specific factor related to holidays that affect urban water demand. The model considered long-term trends, seasonal variations, and daily meteorological conditions, showing high accuracy with an R^2 of 0.95 and a Nash–Sutcliffe efficiency of 0.94 during validation (2011–2015). The study revealed that holidays significantly impact water demand, emphasizing the need to integrate these cultural factors into water demand models for more reliable and systematic predictions. In addition to cultural factors such as holidays, weather conditions also play a pivotal role in influencing water demand dynamics. Cole & Stewart (2013) found that in Hervey Bay, Queensland, Australia, despite higher summer rainfall, temperature emerges as the predominant driver of seasonal peaks in outdoor water consumption in non-drought conditions typical of subtropical climates. This finding underscores the importance of incorporating weather variability into water management strategies to better understand and address residential water usage patterns.

Our study aims to provide a comprehensive understanding of water consumption dynamics and contribute to the development of tailored demand management strategies for the Greater Sydney region. The rest of this research is structured as follows: Section 2 reviews previous related studies in the literature. Section 3 describes the data used in this study and the studied area. Section 4 gives an overview of the approaches employed to investigate how Australia Day and Christmas Day affect water demand. Section 5 presents the findings and discussion. Lastly, Section 6 provides the conclusion and recommendations for future work.

2. RELATED WORK

The primary goal of this section is to explore previous research in the field that examines the methods used to identify the factors that affect the modeling of water demand to examine the effect between the demand for water in the Greater Sydney area and specific occasions, specifically Australia Day and Christmas Day.

A *t*-test is a statistical test used to compare the means of two groups of data. The *t*-test is utilized in certain studies to both select inputs (Quilty *et al.* 2016) and identify the factors that impact water consumption (Makki *et al.* 2015). A *t*-test produces a *t*-value and a *p*-value as its output. When the *p*-value is below the significance level (usually 0.05), this indicates a statistically significant difference between the means of the two groups, and we can reject the null hypothesis of no difference between the means (Field 2013). In this research, a *t*-test can be used to compare the mean water demand on Australia Day or Christmas Day to the mean water demand on other days of the month in multiple years. The test can help determine if there is a significant difference in water demand on these holidays, which may be useful for planning and managing water resources. The results of the test indicate either 'significant' or 'not significant', depending on whether the *p*-value of the test is less than 0.05 or not. Additionally, some studies utilize basic statistical tools to investigate the correlation between water consumption and various factors, including weather conditions and temporal features (Domene & Saurı 2006). One of the common methods for examining and comprehending a dataset is the utilization of descriptive statistics (Domene & Saurı 2006). Several other investigations evaluate the connection between diverse influencing factors and water demand by gauging the intensity of the correlation (Hussien *et al.* 2016; Zubaidi *et al.* 2018).

The use of the Pearson correlation allows for a robust analysis of the strength of the relationship between water demand and explanatory variables, providing a clear picture of the data (Hussien *et al.* 2016; Zubaidi *et al.* 2018). The addition of Christmas Day to the analysis provides a broader understanding of the impact of public holidays on water consumption, allowing for a more complete evaluation of the data.

3. STUDIED AREA AND THE DATA USED IN THIS RESEARCH

The project begins by obtaining the historical daily data on water consumption for various suburbs in the Greater Sydney region from Sydney Water. The relevant meteorological variables are sourced from the Australian Bureau of Statistics website, which is located at http://www.bom.gov.au/climate/data-services/station-data.shtml. The information that will undergo examination was obtained from four water plants located in the Greater Sydney region, specifically Richmond, Warragamba, Mobbs Hill, and Liverpool. The data spans from January 2010 to April 2022 and includes various factors such as maximum and minimum temperatures (measured in °C), the quantity of rainfall (in mm), daily global solar exposure (in MJ/m^2), and the average daily water consumption (m*3/capita.day).

3.1. Applying data preprocessing techniques

To ensure accurate results, it is important to clean and standardize data using data preprocessing techniques before conducting further analysis. Common approaches in the machine learning literature include removing duplicates, inaccuracies, outliers, noise, and missing data, as well as imputing, flattening, and normalizing the data. These techniques maintain data integrity and reduce the impact of noise (Zubaidi *et al.* 2018; Coelho & Andrade-Campos 2019). To simplify matters, the column name is altered in this investigation. Any gaps in the data will be filled using linear interpolation. To detect and eliminate anomalies, the interquartile range (IQR) visualization method is applied to each dataset by employing the technique of finding and eliminating outliers. Previous studies (Vijai & Sivakumar 2018; Zubaidi *et al.* 2018; Coelho & Andrade-Campos 2019) have described this technique, which involves using box and whisker plots. *t*-Tests, descriptive analysis, and the heatmap correlation method using the Pearson correlation will be applied on the same pre-processed dataset. However, to conduct the heatmap analysis, two additional columns were added to the dataset: 'Australia Day' and 'Christmas Day'. These columns were filled with values of 0 and 1. A value of 0 indicates that the date is not Australia Day (which falls on January 26 each year), while a value of 1 indicates that it is Australia Day. The same technique was used to fill the 'Christmas Day' column, with a value of 1 indicating that the date is December 25.

4. OVERVIEW OF THE METHODS AND THE APPROACHES

In this work, various methods are utilized to investigate the effects of Australia Day and Christmas Day water demand in the Greater Sydney region, using data obtained from four water plants, namely Richmond, Warragamba, Mobbs Hill, and Liverpool. These methods include *t*-tests, descriptive statistics, and correlation heatmaps. The descriptive statistics approach uses both quantitative and visual methods to examine the data, summarizing statistical measures such as mean, median, maximum, minimum, and standard deviation. We conduct a comparative analysis to examine whether there are any variations in water consumption on Australia Day compared to other days in January. Additionally, we use data visualization to plot the time series of water consumption in January over 12 years. Similar to the analysis for Australia Day, we conduct a comparative analysis to investigate if there are any changes in water consumption on Christmas Day compared to other days in December to identify any significant changes in consumption that occur on these two holidays. The Pearson correlation heatmap is used to analyze the relationship between water consumption and these two holidays using data from the four water plants, providing a thorough analysis of the correlation between water demand and these events. This approach will provide valuable insights into the impact of these holidays on water consumption in different locations in the Greater Sydney region. By using the Pearson correlation, the strength of the relationship between these variables can be robustly analyzed, resulting in a clear understanding of the water consumption in the Greater Sydney region. The addition of Christmas Day to the analysis provides a broader understanding of the impact of public holidays on water consumption, allowing for a more complete evaluation of the data. Overall, these heatmaps serve as a useful tool for investigating the relationship between water consumption and two public holidays using data from four water plants across the Greater Sydney region. Additionally, in this work, a *t*-test is used to compare the mean water demand on Australia Day and Christmas Day to the mean water demand on other days of the month in multiple years. The *t*-test can help determine if there is a significant difference in water demand on these holidays, which may be useful for planning and managing water resources. The water demand data are extracted for January 26 and December 25 and for the rest of January and the rest of December, for each year in the dataset. A two-sample *t*-test is then performed for each year using a *t*-test function from scipy.stats. The test compares the mean of the water demand on January 26 to the mean of the water demand for the rest of January for that year and a similar approach is used for December. The results of the test are either classified as 'significant' or 'not significant', depending on whether the *p*-value of the test is less than 0.05 or not.

5. RESULTS OF THE EVENT ANALYSIS AND DISCUSSION

In this section, we present and thoroughly analyze the results obtained by applying various methods to study the relationship between water consumption and Australia Day and Christmas Day.

5.1. The influence of Australia Day and Christmas Day on water demand using descriptive analysis

In this section, we present the findings of our quantitative analysis which explores potential variations in water consumption on Christmas Day and Australia Day, in contrast to other days in December and January, respectively. Additionally, we employ data visualization methods to plot the time series of water consumption for December and January for the given datasets, covering both Christmas Day and Australia Day, to detect any significant fluctuations in water consumption on these two holidays.

5.1.1. Australia Day

Here we present the results after conducting the descriptive analysis on the datasets from the four water plants in the Greater Sydney region to investigate if Australia Day water consumption differs from the other days in January for the years from 2010 to 2022. The data were filtered to include only January from 2010 to 2022 and then the data were split into two groups: Australia Day and the rest of January. Then, we calculated the descriptive statistics (mean, standard deviation, minimum, maximum, etc.) for each group to offer further insights into the disparities in water demand between holiday and non-holiday periods, particularly the Australia Day holiday.

The data provided in Figure S1 (Supplementary Material) show that the mean water demand for Australia Day from 2010 to 2022 is 18.37, with a standard deviation of 3.66. The minimum water demand for Australia Day during this time period is 13.1, and the maximum is 22.7. The 25th percentile is 15.5, the 50th percentile is 17.7, and the 75th percentile is 22.7. In contrast, the mean water demand for the rest of January from 2010 to 2022 is 18.1, with a standard deviation of 3.6. The minimum water demand during this time period is 10, and the maximum is 24 which is slightly higher than on Australia Day. The 25th percentile is 14.9, the 50th percentile is 18.4, and the 75th percentile is 21.5. Overall, the mean water demand for Australia Day is slightly higher than the mean water demand for the rest of January, but the difference is relatively small. These data from the Richmond Water Plant provide some evidence that Australia Day may have little impact on water demand, but we cannot definitively conclude this based on the statistics alone. Further analysis, such as hypothesis testing, is necessary to determine the extent and significance of this impact on water demand before drawing any final conclusions. It is also important to note that other factors such as weather patterns may influence water demand during this time period, and these factors need to be accounted for in any analysis. One type of plot that may be useful to investigate the influence of Australia Day on water demand is a box plot. A box plot can show the distribution of water consumption on Australia Day compared to the rest of January, for each year. Figure 1 displays a box plot that represents the daily water consumption for the month of January and Australia Day from 2010 to 2022, the lower and upper boundaries calculated using the mean minus two times the standard deviation formula.

From the box plot, it is evident that there is noticeable variability in daily water consumption across different years. The median water consumption (indicated by the orange lines within the boxes) shows the fluctuations over the years, reflecting changes in water use patterns. The IQR varies year by year, indicating differing levels of variability in daily water consumption. When comparing water consumption on Australia Day to the rest of January, the data do not show a consistent pattern

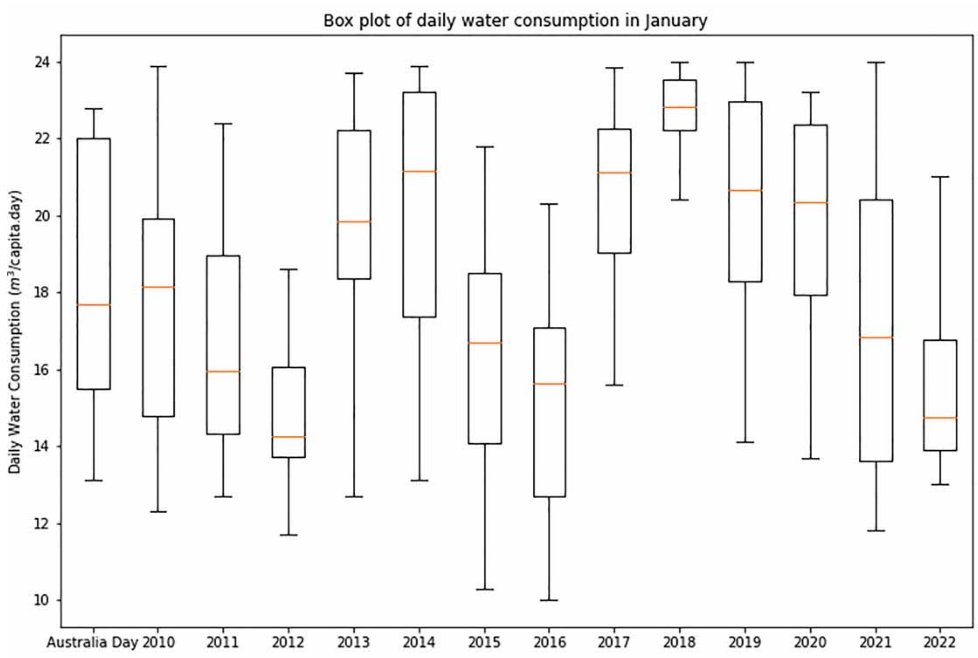


Figure 1 | The box plot for daily water consumption for the month of January and Australia Day from 2010 to 2022 – Richmond Water Plant.

of higher or lower consumption on Australia Day. In some years (e.g., 2010 and 2015), the median consumption on Australia Day appears close to the median for the rest of January, suggesting no significant difference. In other years (e.g., 2011, 2013, and 2014), slight differences are observed, but they are not consistently higher or lower.

Figures 2 and 3 show the daily water consumption for the month of January from 2010 to 2022. In this plot, Australia Day is marked with a red dot, while the rest of January is represented by the other data points.

A detailed analysis based on the daily water consumption trends, with a specific focus on the year 2022 as an example, follows. The plot for the year 2022 shows significant variability in daily water consumption throughout January 2022. Early January has relatively higher water consumption values, followed by a noticeable dip around the middle of the month. On Australia Day (marked by the red dot), there is no significant deviation in water consumption compared to the general trend observed for the rest of January. The water consumption on Australia Day is slightly above the preceding days but does not stand out as an outlier nor is there a significant spike. The daily fluctuations in water consumption are evident, with several peaks and troughs throughout the month. This variability suggests that other factors, such as weather conditions or specific events, might influence daily water consumption more than the holiday itself. The lack of a pronounced increase or decrease in water consumption on Australia Day compared to other days in January indicates that the holiday might not have a substantial impact on water demand. Similar patterns are observed across other years from 2010 to 2022. While there are annual variations in water consumption trends, the data generally do not show a significant, consistent increase on Australia Day compared to other January days. In some years, Australia Day consumption is slightly higher or lower, but these variations are within the normal range of daily fluctuations for the month.

The data presented in Figures 2 and 3 suggest that Australia Day does not significantly alter water demand patterns. While there are daily fluctuations in water consumption, these appear to be driven by factors other than the holiday itself. Further research and analysis such as a *t*-test and a correlation heatmap may be necessary to gain a more comprehensive understanding of the factors influencing daily water consumption.

The descriptive statistics for Australia Day and the rest of January, based on data from the Warragamba Water Plant for the period from 2010 to 2022, are presented in Figure S2 (Supplementary Material). The data indicate a slight increase in water demand on Australia Day compared to the rest of January. The mean water consumption on Australia Day is 3.61, which is slightly higher than the mean consumption for the rest of January (3.42). However, the maximum water consumption on Australia Day (4.94) is slightly lower than the maximum consumption for the rest of January (5.10). Moreover, it should be noted that the disparity between the means may not be statistically significant. Furthermore, the fact that the number of observations for Australia Day is much lower than the number of observations for the rest of January (13 compared to 390) suggests that the findings may not be entirely representative. It might be necessary to conduct further investigation and analysis to obtain a

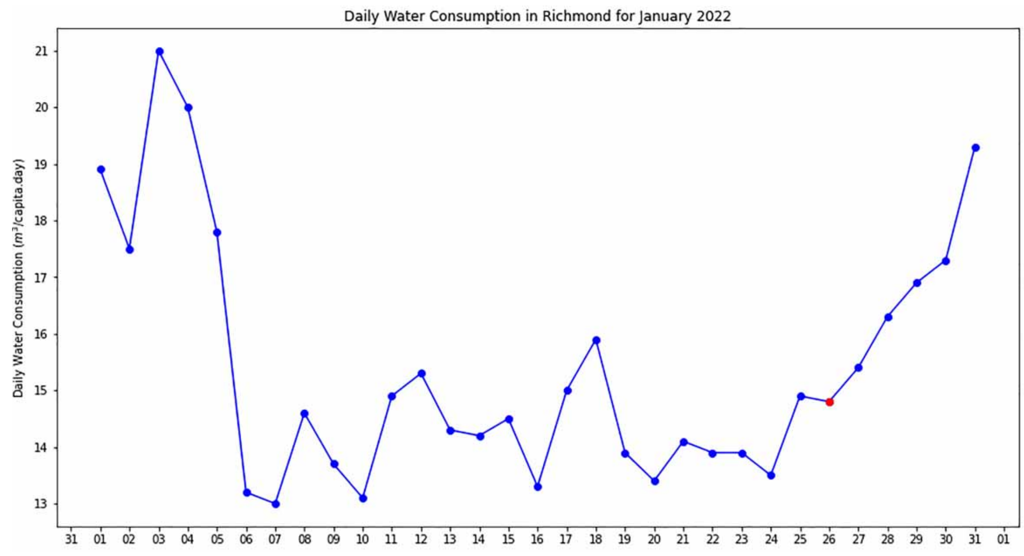


Figure 2 | Daily water consumption for the month of January 2022 – Richmond Water Plant.

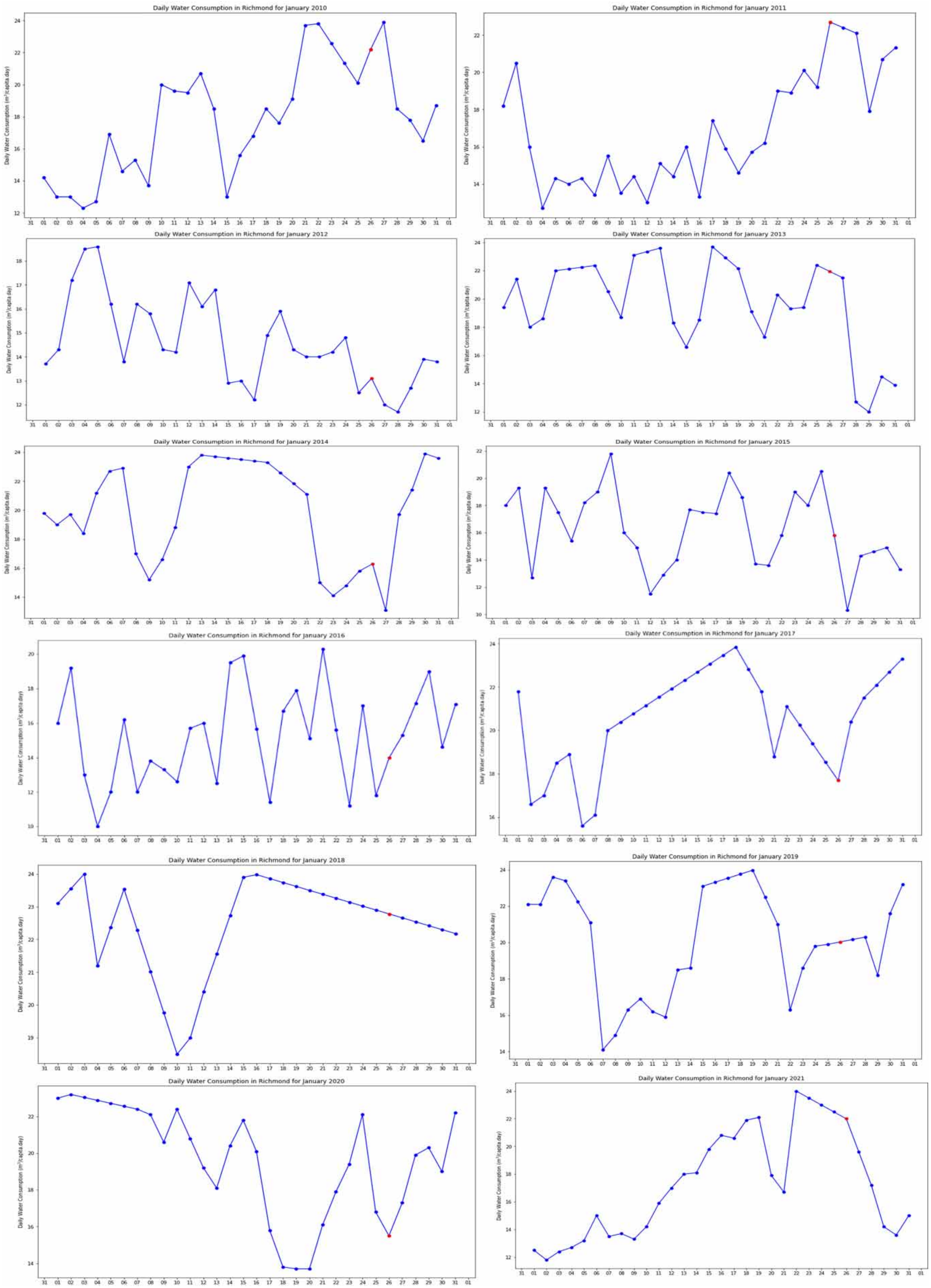


Figure 3 | Daily water consumption for January 2010 to 2021 – Richmond Water Plant.

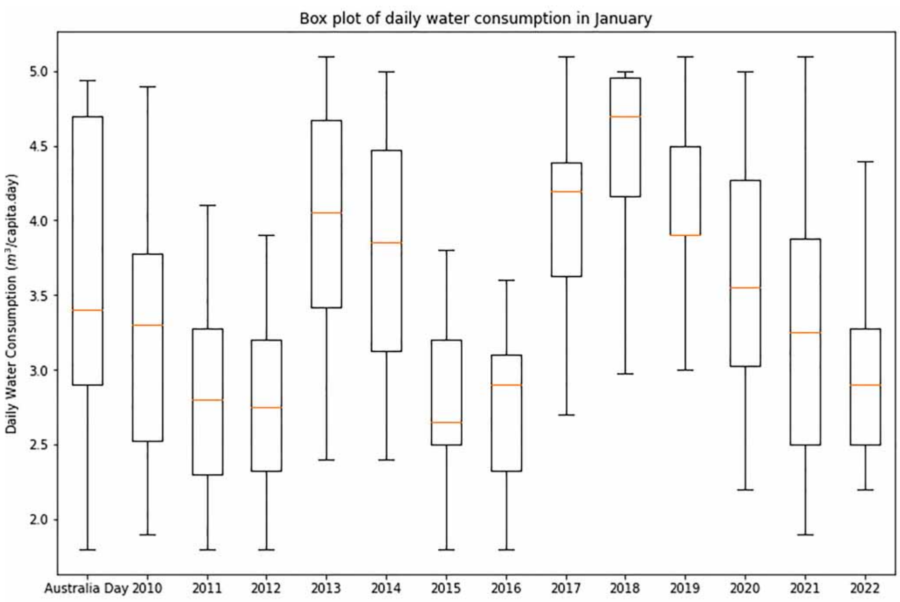


Figure 4 | The box plot for the daily water consumption for the month of January and Australia Day from 2010 to 2022 – Warragamba Water Plant.

more precise understanding of the significance of the variation. All in all, these descriptive statistics offer some preliminary understanding of the water consumption trends on Australia Day compared to the rest of January.

Figure 4 displays a box plot representing the daily water consumption for the month of January and Australia Day from 2010 to 2022 using data from the Warragamba Water Plant.

From Figure 4, it appears that water consumption on Australia Day does not significantly differ from the rest of January. Following is a detailed interpretation of the conclusion based on the plot:

- Median comparison: The median water consumption on Australia Day is similar to the median for the other days of January. There is no noticeable spike or drop in the median value for Australia Day, indicating that the central tendency of water consumption remains consistent.
- IQR: The IQR for Australia Day is comparable to those for the other days. This suggests that the variability in water consumption is not significantly different on Australia Day compared to the rest of the month.
- Visual inspection: Visually, the box for Australia Day blends with the rest of the boxes, reinforcing the conclusion that water consumption patterns are stable and not notably affected by the holiday.

Figures 5 and 6 show the daily water consumption for the month of January from 2010 to 2022 using data from the Warragamba Water Plant. In this plot, Australia Day is marked with a red dot, while the rest of January is represented by the remaining data points.

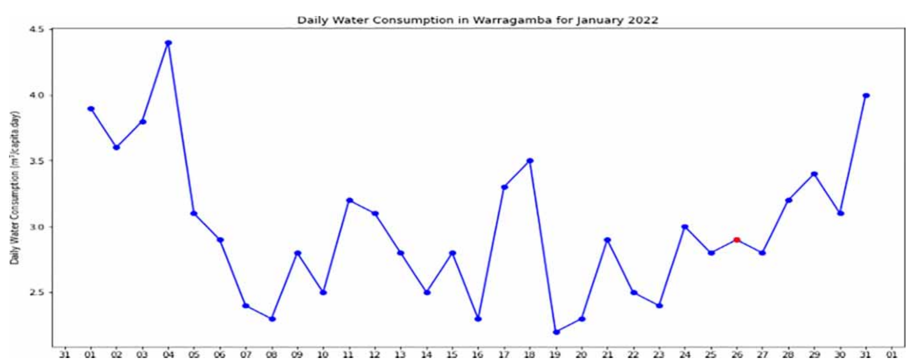


Figure 5 | Daily water consumption for the month of January 2022 –Warragamba Water Plant.

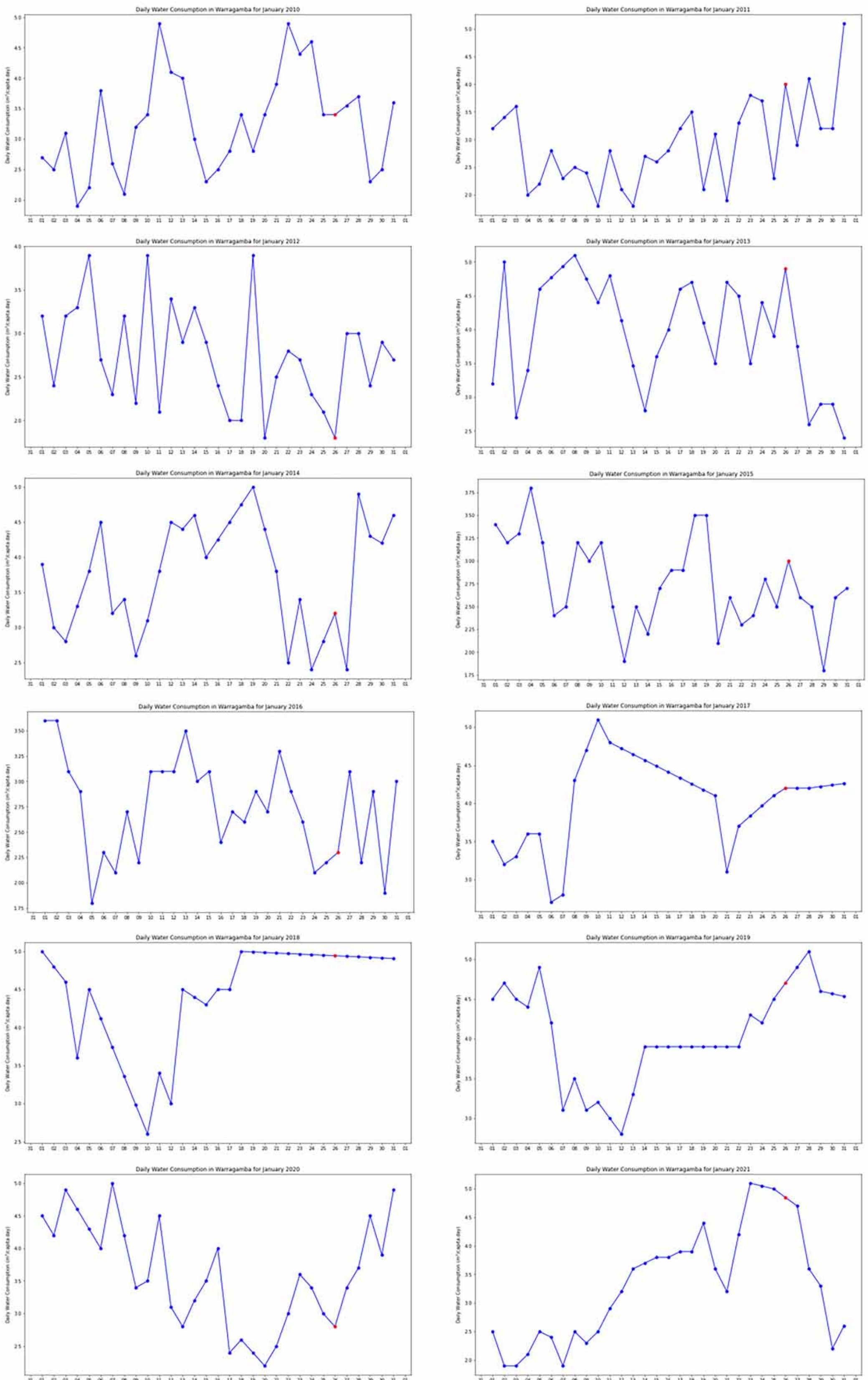


Figure 6 | Daily water consumption for January from 2010 to 2021 – Warragamba Water Plant.

Based on the data presented in Figures 5 and 6, it seems that Australia Day has no effect on water demand at the Warragamba Water Plant. The plot does not show any noticeable difference in water demand on Australia Day when compared to the other days in January. However, to obtain a more complete understanding of the relationship between Australia Day and water demand, further research is necessary. Overall, the presented data suggest that Australia Day does not have a significant impact on water demand at the Warragamba Water Plant.

According to the data provided in Figure S3 (Supplementary Material), which presents the descriptive statistics for Australia Day and the rest of January at the Mobbs Hill Water Plant for the period from 2010 to 2022, it seems that water consumption on Australia Day may have slightly increased. The average water consumption on this day is 31.70, which is a bit higher than the average for the rest of January, which is 31.26. However, the disparity between the means is not very large, so it might not be statistically significant. Moreover, the sample size for Australia Day is much smaller than that of the rest of January, with only 13 data points as opposed to 390 data points. This could affect the accuracy of the results, and it should be considered when interpreting the findings. In summary, while Australia Day could potentially impact water consumption, more research and statistical analysis is needed to confirm this.

Figure 7 displays a box plot representing the daily water consumption for the month of January and Australia Day from 2010 to 2022 at the Mobbs Hill Water Plant.

The box plot in Figure 7 shows that daily water consumption varies significantly across different years. The median water consumption, shown by the orange lines within the boxes, fluctuates annually, indicating changes in water consumption patterns. The IQR also varies each year, reflecting differing levels of variability in daily water consumption.

When comparing water consumption on Australia Day to the rest of January, the data from the Mobbs Hill Water Plant do not exhibit a consistent pattern of higher or lower consumption on Australia Day. In some years (e.g., 2013), the median consumption on Australia Day is very close to the median for the rest of January, suggesting no significant difference. In other years (e.g., 2016, 2018, 2019, and 2022), slight differences are observed, but they are not consistently higher or lower.

Figures 8 and 9 show the daily water consumption for the month of January from 2010 to 2022 at the Mobbs Hill Water Plant. In this plot, Australia Day is marked with a red dot, while the rest of January is represented by the remaining data points.

Based on the results from Figures 8 and 9, Australia Day has no discernible impact on water demand at the Mobbs Hill Water Plant. The plot indicates that water demand on Australia Day is not noticeably different from that of other days in January.

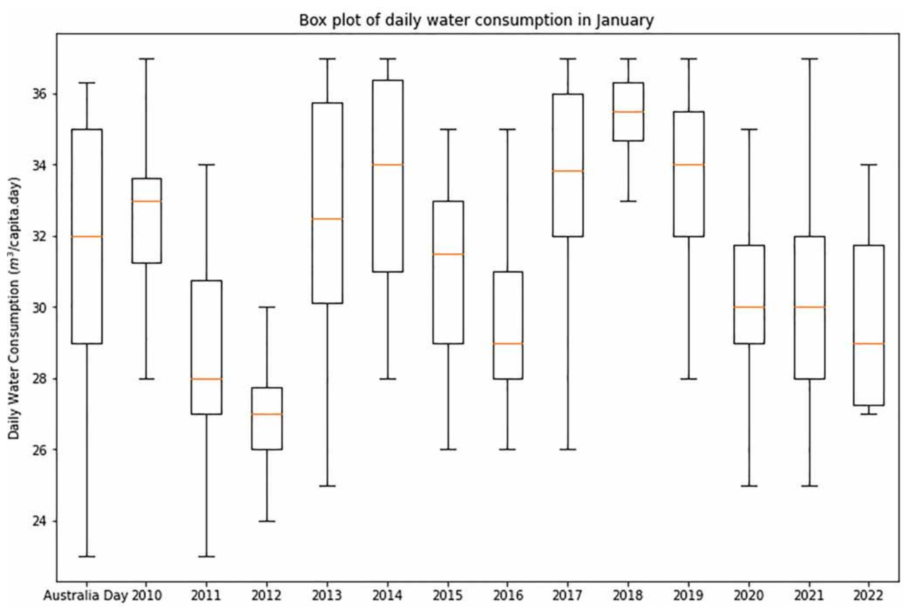


Figure 7 | The box plot for daily water consumption for the month of January and Australia Day from 2010 to 2022 – Mobbs Hill Water Plant.

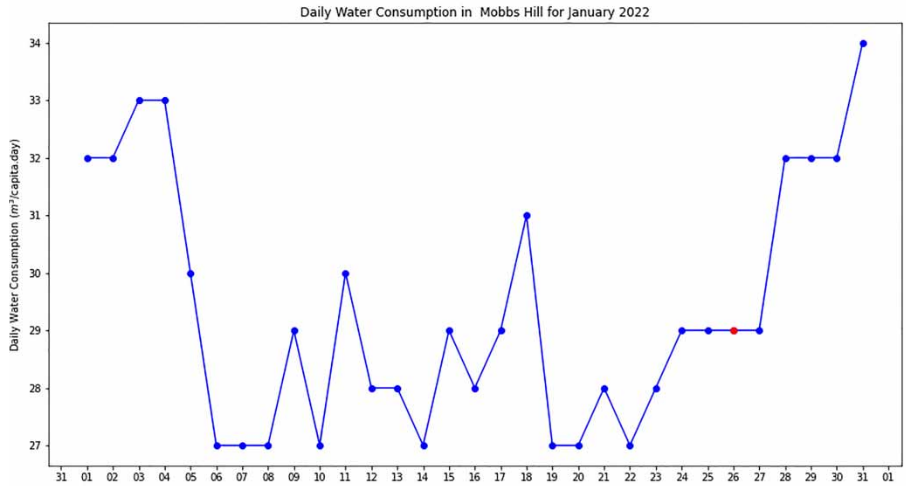


Figure 8 | Daily water consumption for the month of January 2022 – Mobbs Hill Water Plant.

Figure S4 (Supplementary Material) shows the descriptive statistics for Australia Day and the rest of January, based on data from the Liverpool Water Plant for the period from 2010 to 2022. These statistics indicate that the mean water consumption on Australia Day (64.70 units) is very close to that of the rest of January (64.72 units). The standard deviations (Australia Day: 8.18 units, rest of January: 8.26 units) suggest comparable variability in water consumption between the two periods. The range of consumption values (from minimum to maximum) is slightly narrower on Australia Day (50–78 units) compared to the rest of January (46–79 units), but this difference is not significant. Overall, these findings suggest that Australia Day does not have a substantial impact on water consumption patterns at the Liverpool Water Plant. Figure 10 displays a box plot representing the daily water consumption for the month of January and Australia Day from 2010 to 2022 based on data from the Liverpool Water Plant.

As shown in the box plot in Figure 10, when analyzing water consumption on Australia Day compared to the rest of January, data from the Liverpool Water Plant spanning 2010–2022 does not indicate consistently higher or lower consumption patterns on Australia Day.

Figures 11 and 12 show the daily water consumption for the month of January from 2010 to 2022 based on data from the Liverpool Water Plant. In this plot, Australia Day is marked with a red dot, while the rest of January is represented by the other data points.

The visualization plots in Figures 11 and 12 suggest that Australia Day has no clear influence on water demand at the Liverpool Water Plant, as there is no noticeable difference in water demand between Australia Day and other days in January. However, for a comprehensive understanding of the relationship between Australia Day and water demand, further investigation using statistical tests such as the *t*-test and correlation heatmap may be necessary. Therefore, based on the available data, it appears that Australia Day does not have a significant impact on water demand at the Liverpool Water Plant.

5.1.2. Christmas Day

This subsection presents the outcomes of the descriptive analysis of the datasets from four water plants in the Greater Sydney region. The aim of this analysis is to investigate whether there is a difference in water consumption on Christmas Day compared to the other days in December for the years 2010–2021. To do this, we filtered the data to include only December from 2010 to 2021, and then divided it into two groups: Christmas Day and the remaining days in December. Descriptive statistics such as mean, standard deviation, minimum, and maximum were calculated and presented for each group. This was done to provide further insights into any differences in water demand between holiday and non-holiday periods, particularly on Christmas Day. Figure S5 (Supplementary Material) provides the descriptive statistics for Christmas Day and the rest of December, based on data from the Richmond Water Plant from 2010 to 2021. The statistics suggest that Christmas Day may have only a minimal impact on water demand at the Richmond Water Plant; however, further analysis is needed to confirm this. Hypothesis testing would be necessary to determine the significance of this impact on water demand before drawing

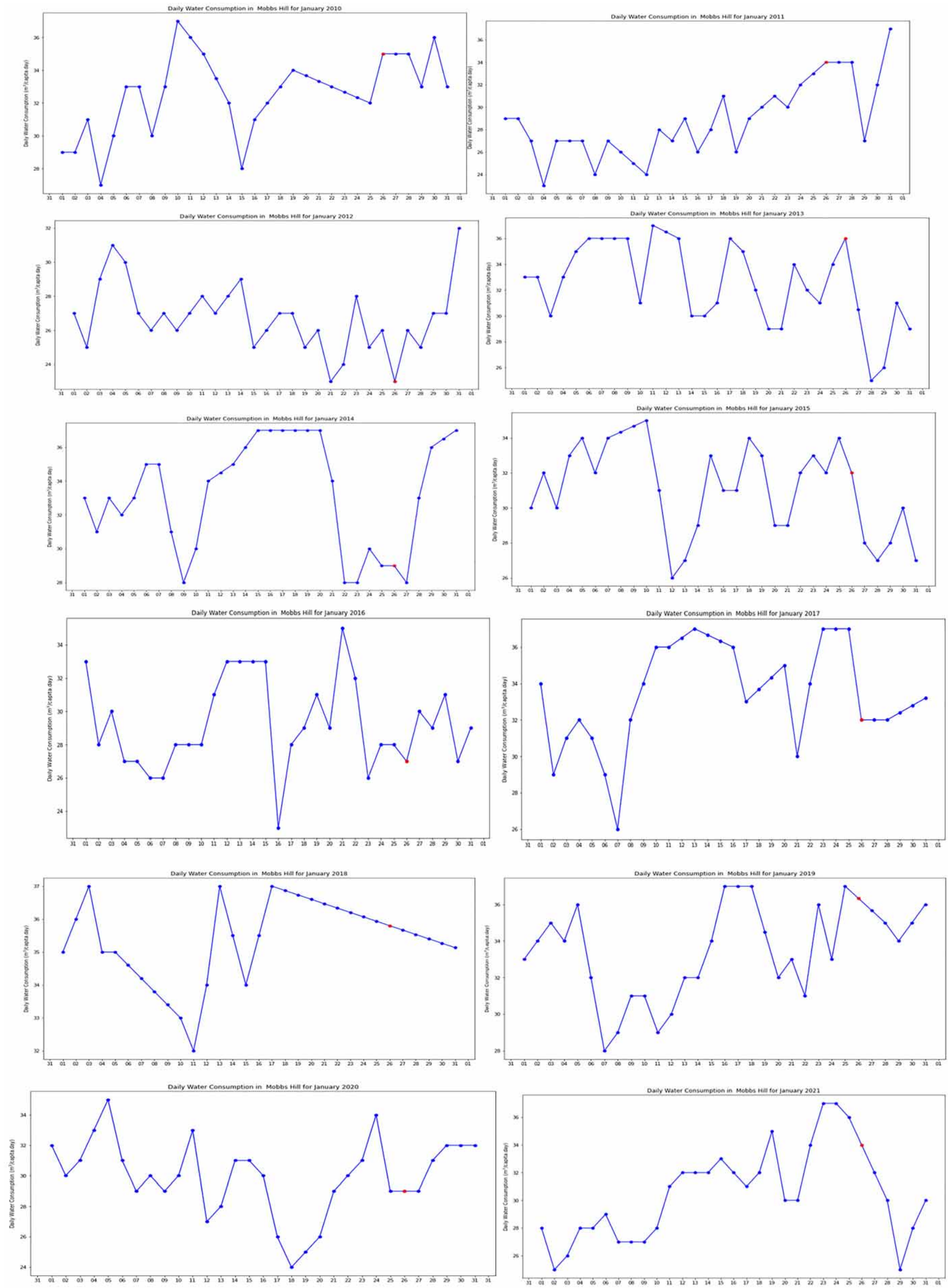


Figure 9 | Daily water consumption for the month of January 2010 to 2021 – Mobbs Hill Water Plant.

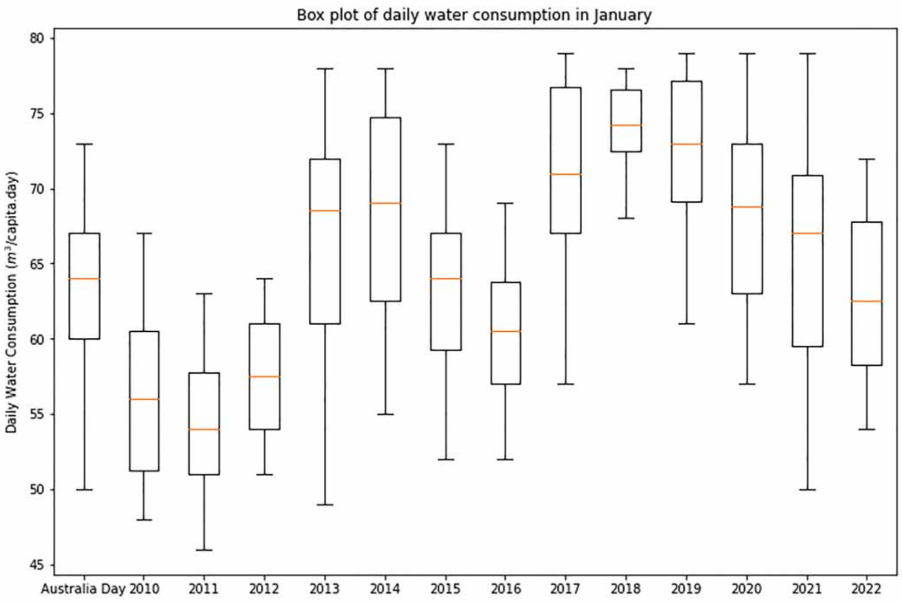


Figure 10 | The box plot for the daily water consumption for the month of January and Australia Day from 2010 to 2022 – Liverpool Water Plant.

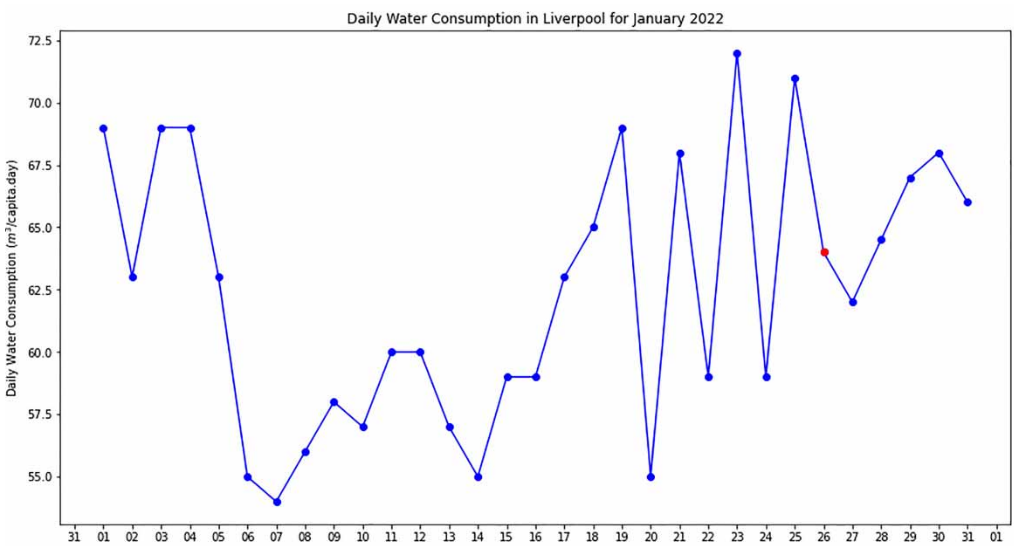


Figure 11 | Daily water consumption for the month of January 2022 – Liverpool Water Plant.

any definitive conclusions. A box plot can show the distribution of water consumption for each year on Christmas Day compared to the rest of December.

Figure 13 depicts a box plot which illustrates the daily water consumption for December and Christmas Day from 2010 to 2021, with lower and upper boundaries calculated using the mean minus two times the standard deviation formula.

Figure 14 shows the daily water consumption for the month of December from 2010 to 2021. In this plot, Christmas Day is marked with a red dot, while the rest of December is represented by the remaining data points.

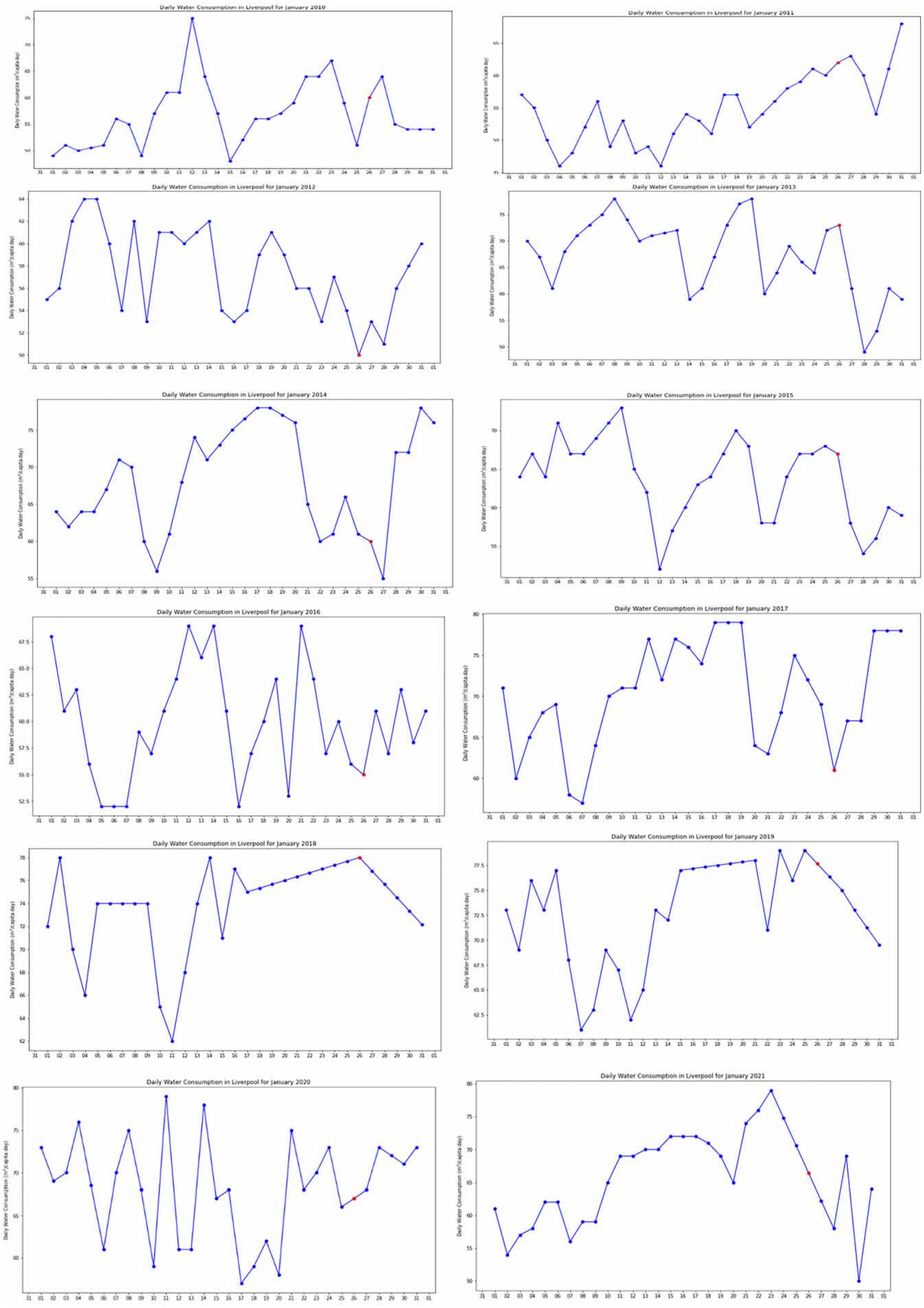


Figure 12 | Daily water consumption for the month of January 2010 to 2021 – Liverpool Water Plant.

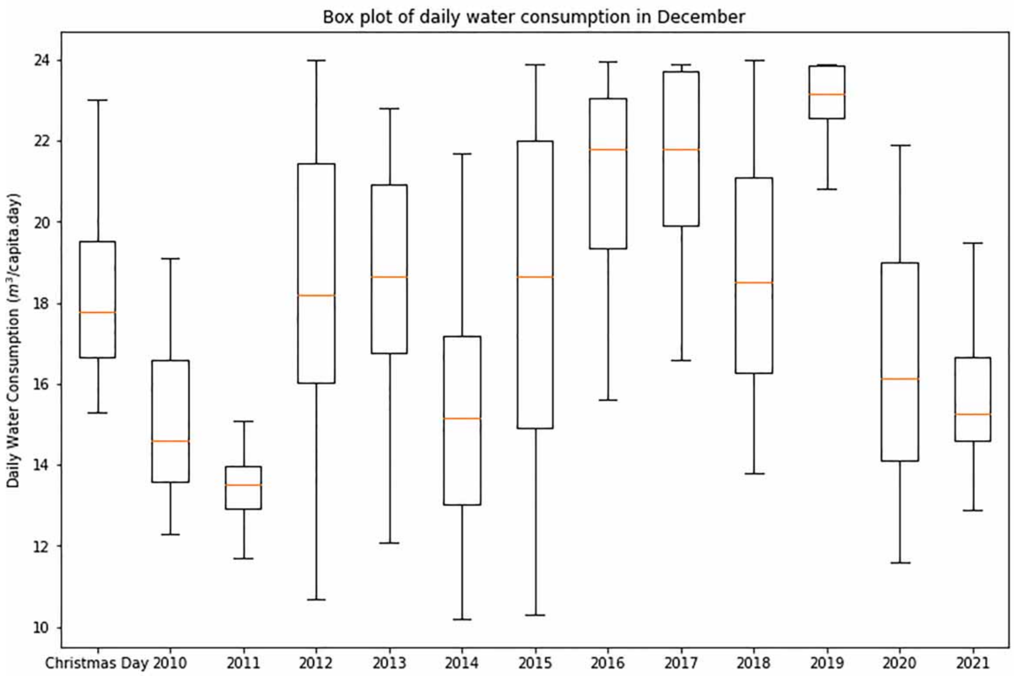


Figure 13 | The box plot for daily water consumption for the month of December and Christmas Day from 2010 to 2021 – Richmond Water Plant.

Based on the visualization plot in Figure 14, it seems that Christmas Day has no impact on water demand at the Richmond Water Plant. There is no noticeable difference in water demand on Christmas Day compared to the rest of December. However, to gain a more complete understanding of the situation, additional research and analysis, such as a *t*-test and correlation heatmap, may be required.

Figure S6 (Supplementary Material) shows the descriptive statistics for Christmas Day and the rest of December, based on data from the Warragamba Water Plant for the period from 2010 to 2021. In conclusion, the study examined the influence of Christmas Day on water demand compared to the rest of December from 2010 to 2021. The statistics in Figure S6 (Supplementary Material) shows that the mean water demand on Christmas Day is slightly lower than the mean water demand for the rest of December. Additionally, the standard deviation of water demand for the rest of December (0.98) is much higher than that for Christmas Day (0.62). This indicates that water demand on Christmas Day is more consistent compared to the rest of December. Figure 15 depicts a box plot that illustrates the daily water consumption for December and Christmas Day from 2010 to 2021, with lower and upper boundaries calculated using the mean minus two times the standard deviation formula.

Figure 16 shows the daily water consumption for the month of December from 2010 to 2021 at the Warragamba Water Plant. In this plot, Christmas Day is marked with a red dot, while the rest of December is represented by the remaining data points.

Based on the visualization plot in Figure 16, it seems that Christmas has no impact on water demand at the Warragamba Water Plant. There is no noticeable difference in water demand on Christmas Day compared to the rest of December.

Referring to Figure S7 (Supplementary Material), the statistics show that the mean water consumption was slightly lower on Christmas Day compared to the rest of December, based on data from the Mobbs Hill Water Plant for the period from 2010 to 2021. Therefore, based on the results, it can be concluded that Christmas Day does not have a significant impact on water demand. Future studies can explore other factors that may affect water demand during the holiday season, such as temperature and population density.

Figure 17 depicts a box plot that illustrates the daily water consumption for December and Christmas Day from 2010 to 2021.

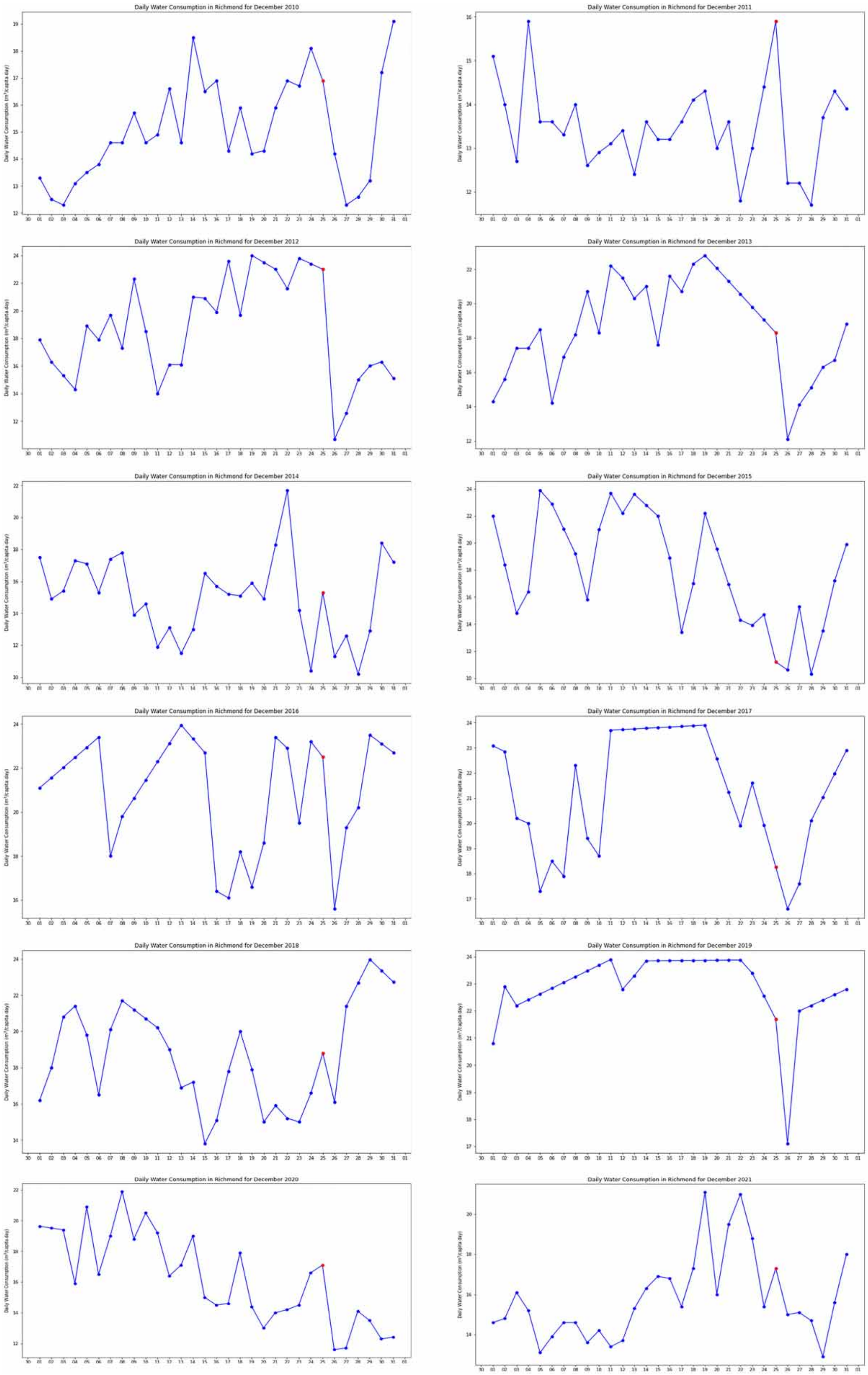


Figure 14 | Daily water consumption for December 2010 to 2021 – Richmond Water Plant.

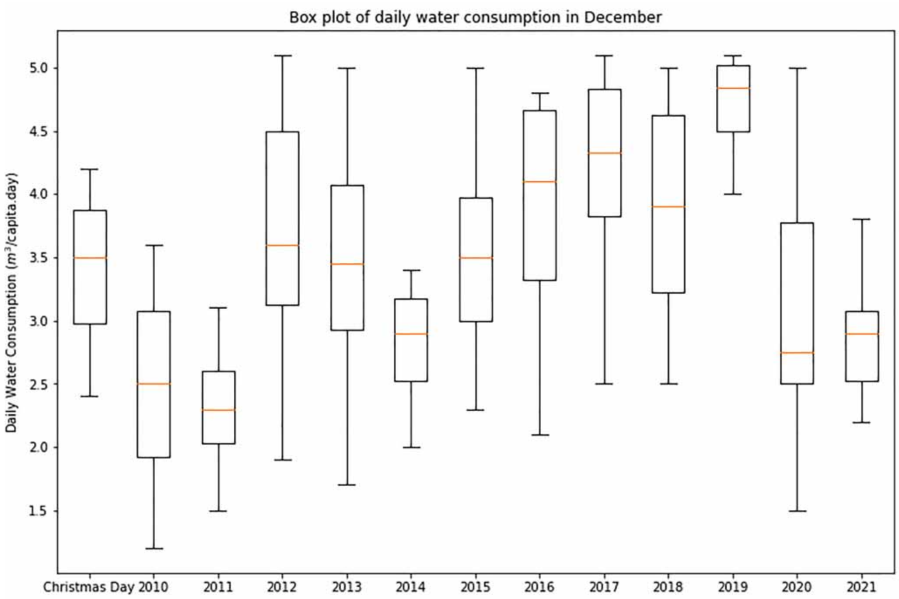


Figure 15 | The box plot for daily water consumption for the month of December and Christmas Day from 2010 to 2021 – Warragamba Water Plant.

Figure 18 shows the daily water consumption for the month of December from 2010 to 2021 based on data from the Mobbs Hill Water Plant. In this plot, Christmas Day is marked with a red dot, while the rest of December is represented by the other data points.

From the visualization plot in Figure 18, it appears that Christmas Day has no discernible impact on water demand at the Mobbs Hill Water Plant. The water demand on Christmas Day is similar to that of the rest of December. Nevertheless, a comprehensive investigation of the situation may necessitate further research and analysis.

Referring to Figure S8 (Supplementary Material), the descriptive statistics compare the water demand on Christmas Day to that of the rest of December from 2010 to 2021 at the Liverpool Water Plant. The results indicate that the mean water demand on Christmas Day is slightly higher than the rest of December. However, the difference in means is not statistically significant enough to conclude that Christmas Day has a significant impact on water demand. The descriptive statistics show variability in water demand, but no clear pattern or trend emerges. Therefore, the study suggests that Christmas Day may not significantly influence water demand compared to the rest of December. Figure 19 depicts a box plot that illustrates the daily water consumption for December and Christmas Day from 2010 to 2021.

Figure 20 shows the daily water consumption for the month of December from 2010 to 2021 using data from the Liverpool Water Plant. In this plot, Christmas Day is marked with a red dot, while the rest of December is represented by the remaining data points.

Based on the results from Figure 19 and the visualization plot in Figure 20, water demand at the Liverpool Water Plant shows consistency on Christmas Day, with no observable effect on water consumption when compared to the rest of December. These findings indicate that Christmas Day does not significantly impact water demand patterns at the Liverpool Water Plant during this period.

5.2. The influence of Australia Day and Christmas Day on water demand using *t*-test inferential statistics

To explore the impact of Australia Day and Christmas Day on water demand in the region, *t*-tests were performed on water consumption data. The *t*-tests compared the mean water demand on Christmas Day and Australia Day with the mean water demand on other days for the months of December and January, respectively, using the provided dataset.

Parameter values:

- *t*-Value: The *t*-value calculated from each *t*-test.
- *p*-Value: The *p*-value calculated from each *t*-test.
- Significance level: The significance level was set at 0.05.

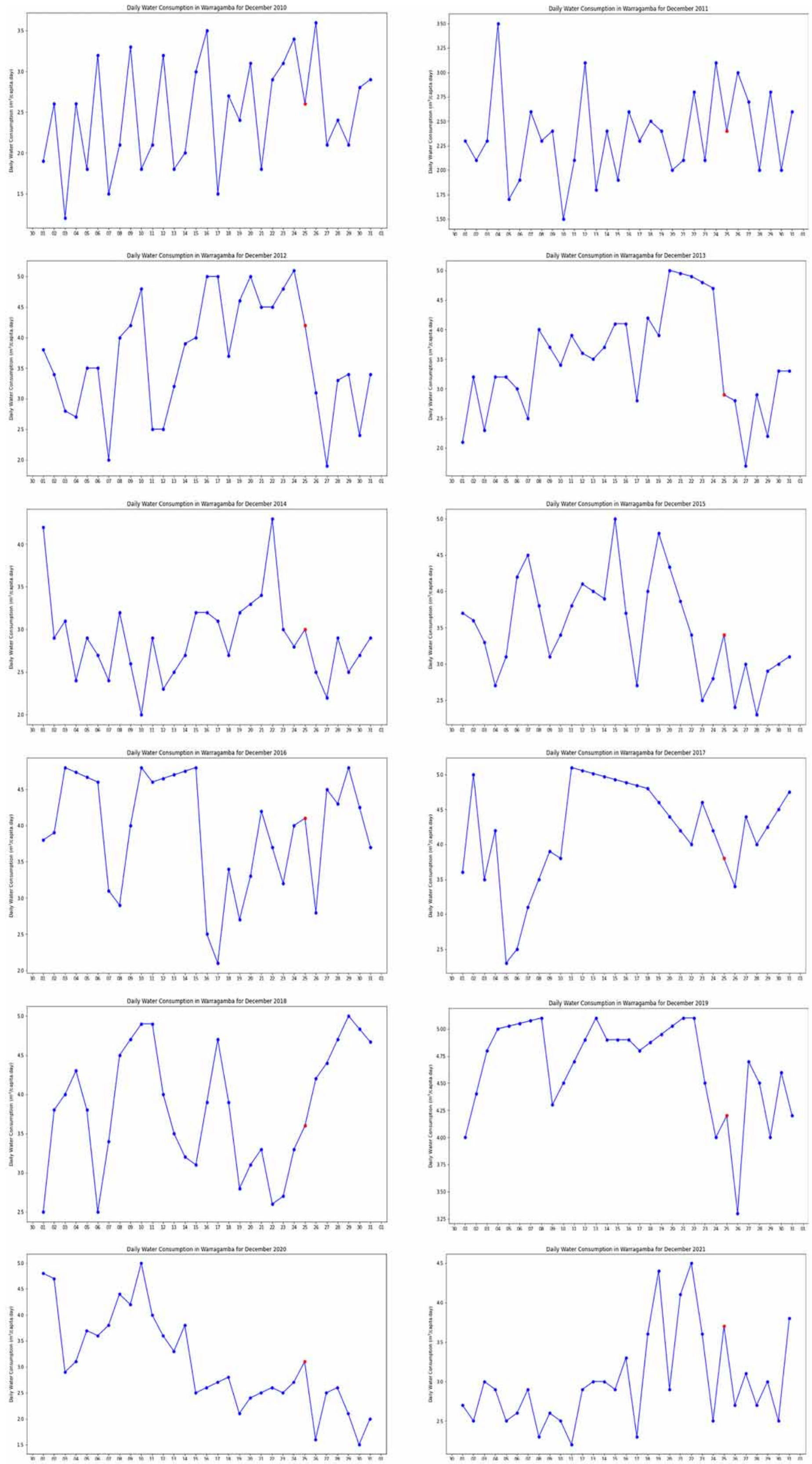


Figure 16 | Daily water consumption for December 2010 to 2021 – Warragamba Water Plant.

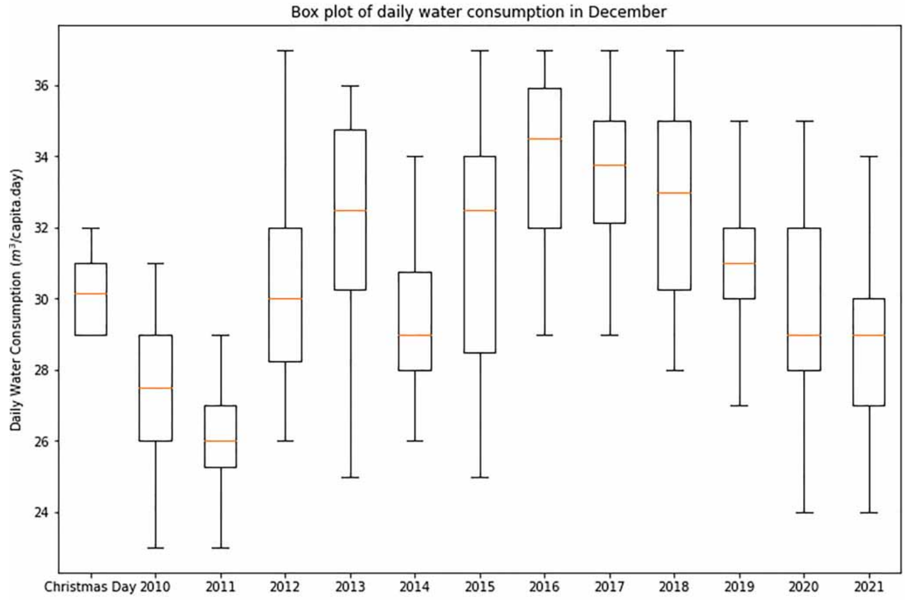


Figure 17 | The box plot for the daily water consumption for the month of December and Christmas Day from 2010 to 2021 – Mobbs Hill Water Plant.

Testing process description:

1. Data collection and preparation: Water demand data for each water plant were collected and stored in a structured format using a pandas DataFrame. The date column was converted to datetime format to facilitate temporal analysis.
2. Data segmentation: For each year in the study period, data for Christmas Day and the remaining days of December, as well as for Australia Day and the rest of January, were extracted separately for each water plant.
3. Statistical analysis: Utilizing the stats.ttest_ind function from scipy.stats, independent two-sample *t*-tests were conducted for each year to compare:

   - the mean water demand on Christmas Day against the mean demand for the rest of December at each water plant.
   - the mean water demand on Australia Day against the mean demand for the rest of January at each water plant.

Interpretation criteria

The results were interpreted based on the calculated *p*-values: If the *p*-value was less than 0.05, it was concluded that there was a statistically significant difference in water demand between the respective holiday and the rest of the month. If the *p*-value was greater than or equal to 0.05, it was concluded that there was no statistically significant difference in water demand between the respective holiday and the rest of the month.

5.2.1. Christmas Day

Statistical analysis using *t*-tests was performed on water demand data for December 25 and the remaining days of December for each year across the Richmond, Warragamba, Mobbs Hill, and Liverpool Water Plants. Based on the *t*-test results from 2010 to 2022, it can be concluded that there is no significant difference in water demand on December 25 compared to the rest of December. This suggests that water demand on December 25 does not exhibit a consistent pattern of being significantly higher or lower than the rest of December; observed differences are likely due to random variation rather than a systematic trend. These findings are valuable for water management and planning, indicating that no special measures are necessary to accommodate water demand specifically on December 25.

5.2.2. Australia Day

A statistical analysis using the *t*-test was conducted on water demand data for January 26 and the remaining days of January for each year across the four water plants. The results indicate that there is no significant difference in water demand on

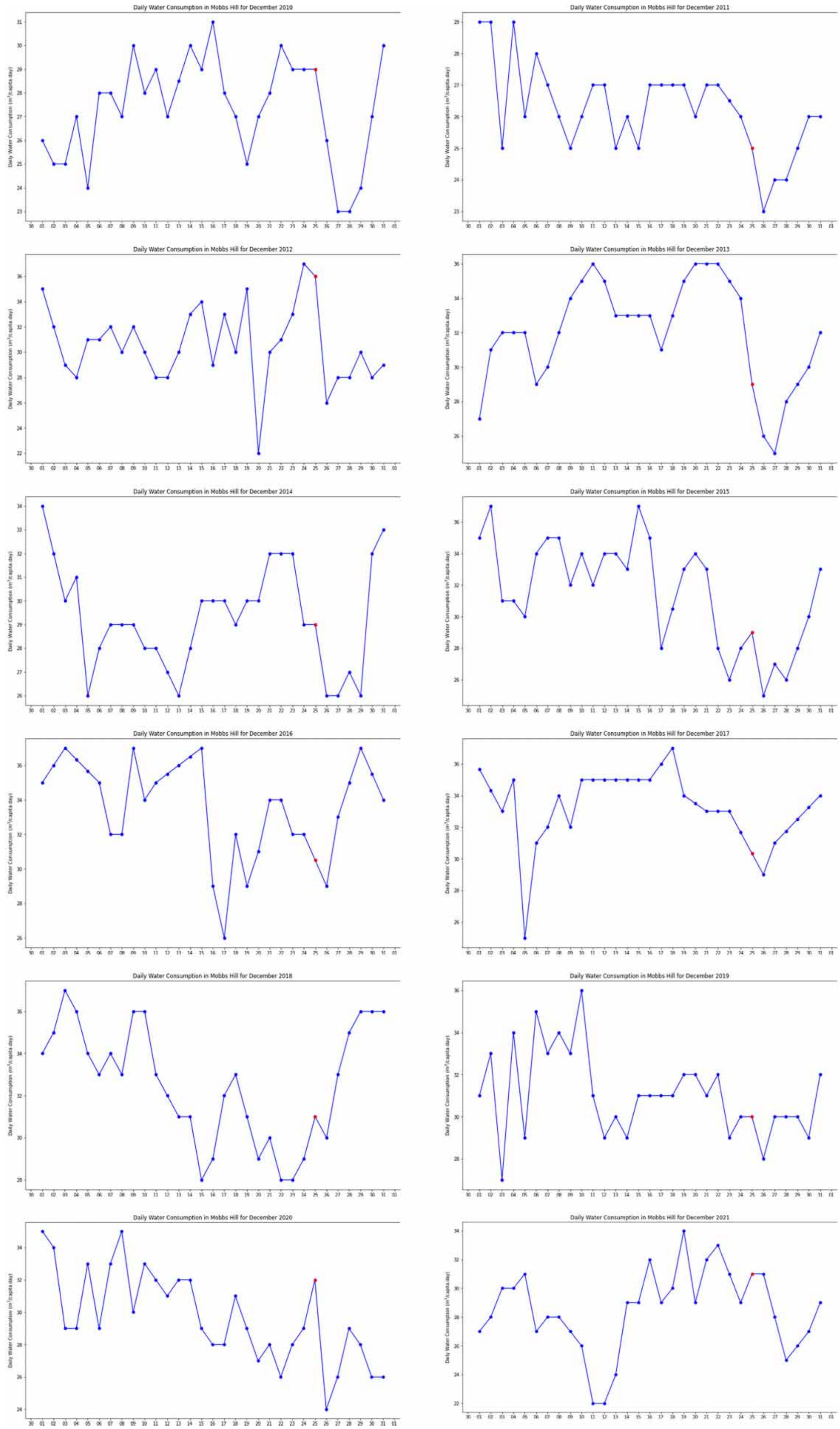


Figure 18 | Daily water consumption for December 2010 to 2021 – Mobbs Hill Water Plant.

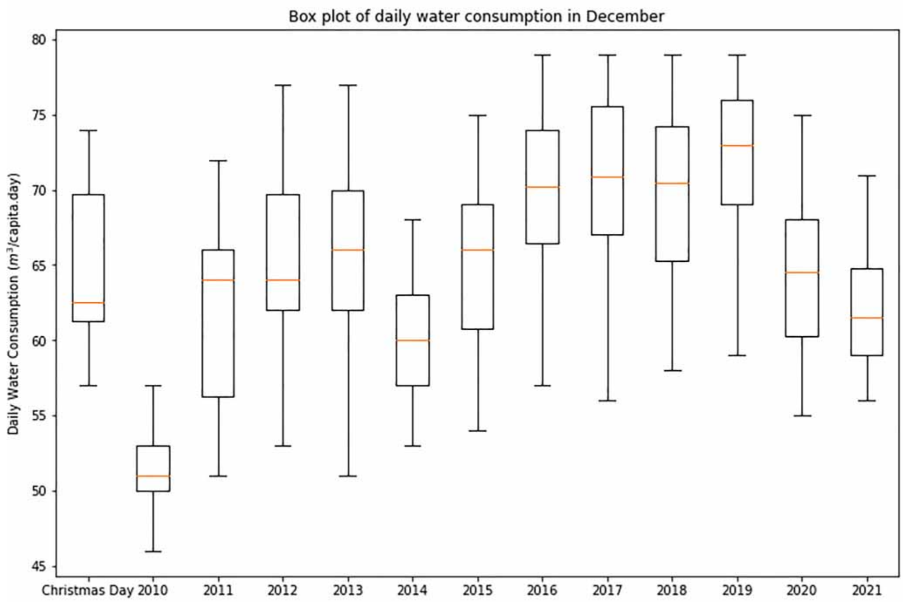


Figure 19 | The box plot for daily water consumption for the month of December and Christmas Day from 2010 to 2021 – Liverpool Water Plant.

January 26 compared to the rest of January from 2010 to 2022 across the four water plants. This finding suggests that water demand at these facilities on January 26 is not influenced by Australia Day or any specific factors that differ from the rest of January. These findings are valuable for water management and planning, as they suggest that no special measures are necessary to accommodate water demand, specifically on January 26.

The following subsection discusses the creation of a correlation heatmap to examine the relationships between water demand, Australia Day, Christmas Day, and other relevant factors such as weather variables.

5.3. The influence of Australia Day and Christmas Day on water demand using a Pearson correlation heatmap

Data visualization is a method of conveying information in a clear and effective manner. Humans have an innate instinct to interpret visual images, so presenting a visually appealing map that illustrates the connection between water consumption and certain events and other factors can facilitate the examination of their impact. Figures 21–24 present heatmaps that utilize the Pearson correlation to analyze the level of correlation between water consumption, Australia Day, Christmas Day, and the weather variables across four water plants.

The heatmap results from Figures 21–24 show a weak association between water consumption and Christmas Day and Australia Day, implying that these holidays have little or no correlation with water demand in the Greater Sydney region. However, the results suggest that weather factors have a stronger relationship with water consumption than Australia Day or Christmas Day. Higher values of weather-related variables may play a more significant role in explaining variations in water consumption in the Greater Sydney region. The analysis reveals the following insights:

- the maximum temperature and daily global solar exposure are the primary factors affecting water consumption in some suburbs in Greater Sydney, with a strong positive correlation.
- there is a strong negative correlation between water consumption and rainfall.
- the water consumption and minimum temperature have a high correlation.

Time series analysis provides valuable insights into how various factors change over time. Through data visualization, trends, correlations, and the potential impacts of these factors on water consumption can be identified. Such analysis aids in understanding the relationship between these variables and their influence on water consumption. Figure 25 presents the time series analysis of various weather factors alongside water consumption using the Richmond water plant data to examine the impact of weather variables on water consumption. The results for the other water plants are available upon request.

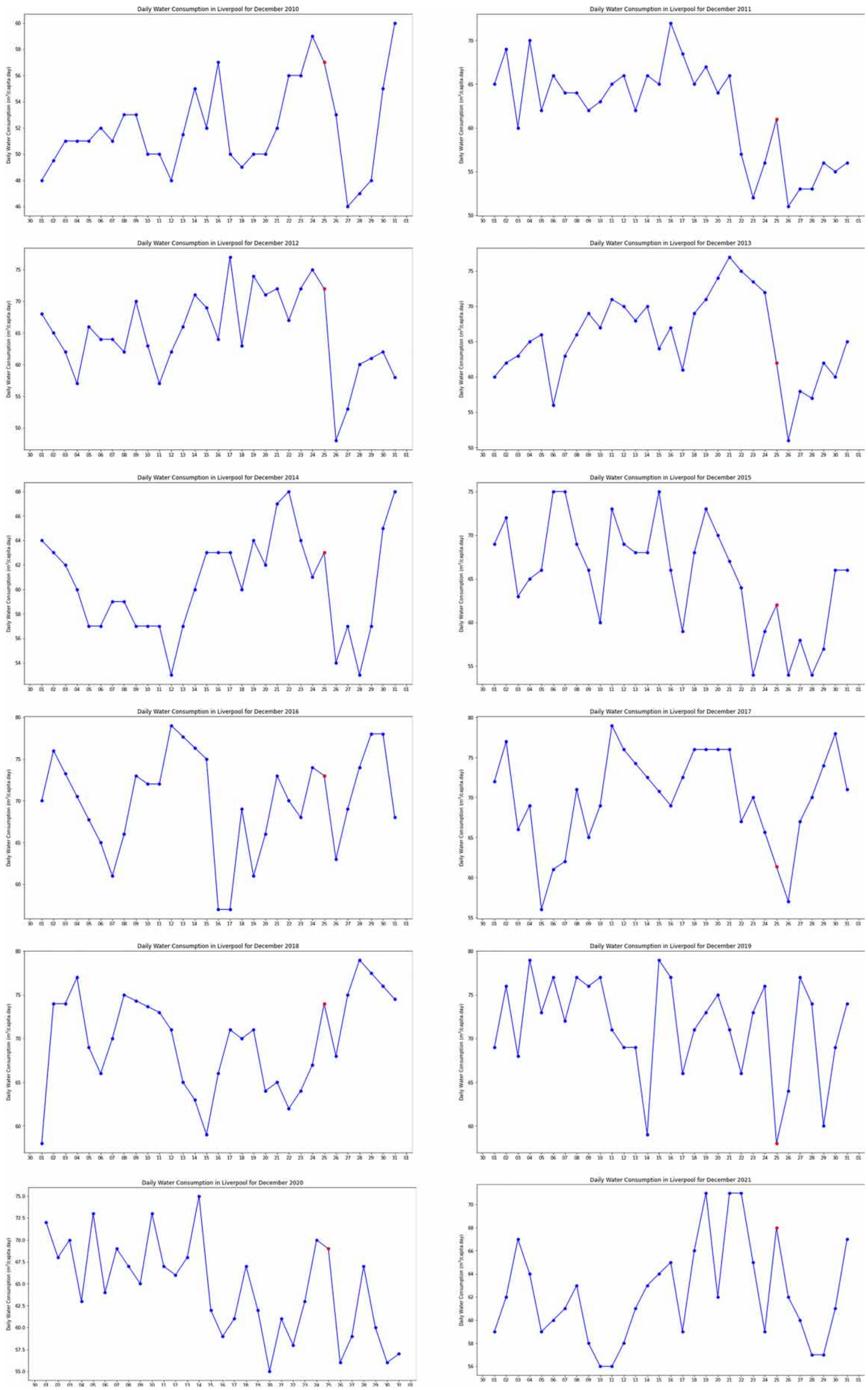


Figure 20 | Daily water consumption for December 2010 to 2021 – Liverpool Water Plant.

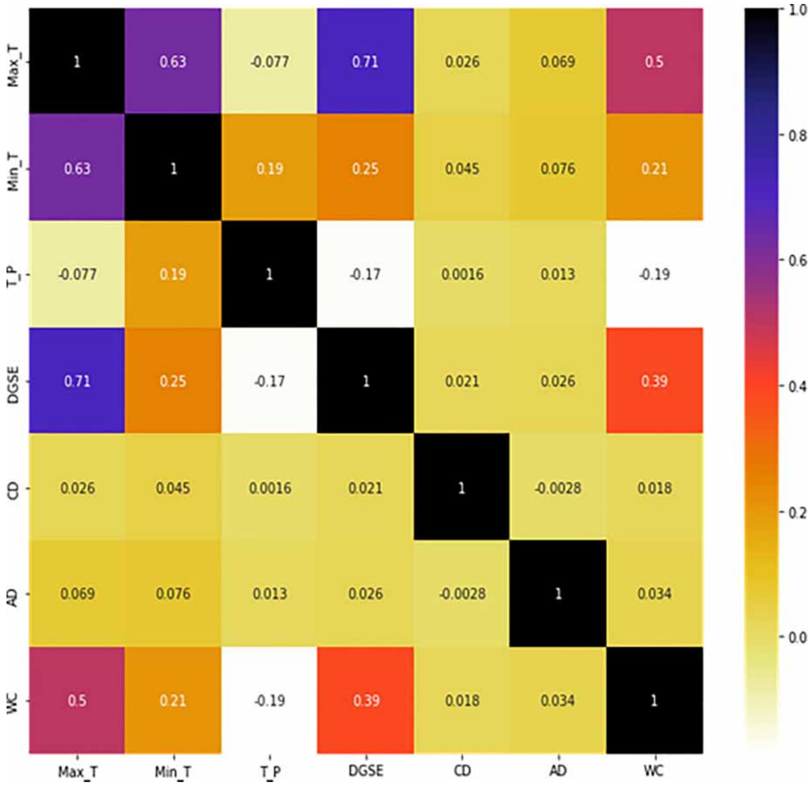


Figure 21 | Heatmap illustrating the Pearson correlation coefficients between water consumption, weather variables, Australia Day, and Christmas Day – Richmond Water Plant.

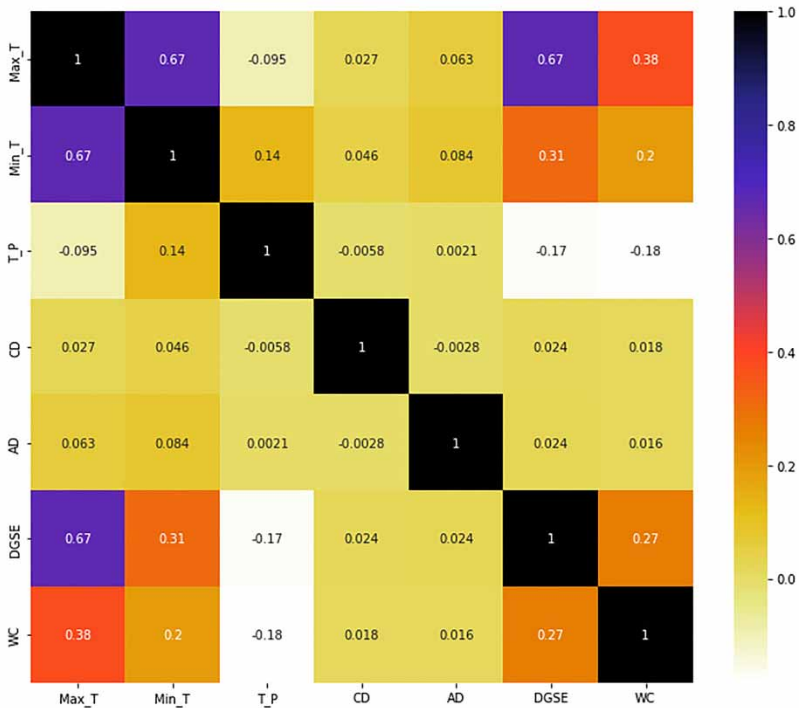


Figure 22 | Heatmap illustrating the Pearson correlation coefficients between water consumption, weather variables, Australia Day, and Christmas Day – Warragamba Water Plant.

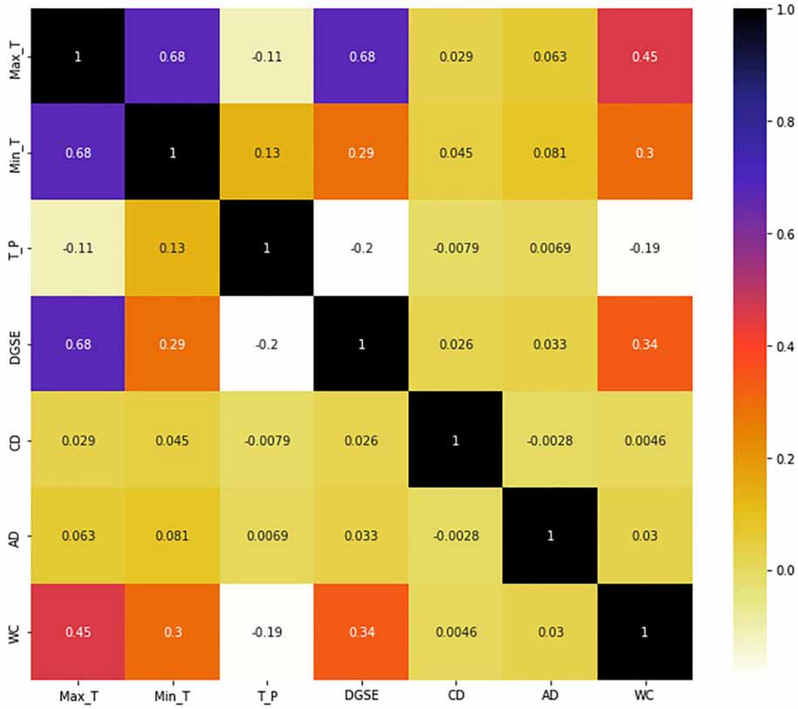


Figure 23 | Heatmap illustrating the Pearson correlation coefficients – Mobbs Hill Water Plant.

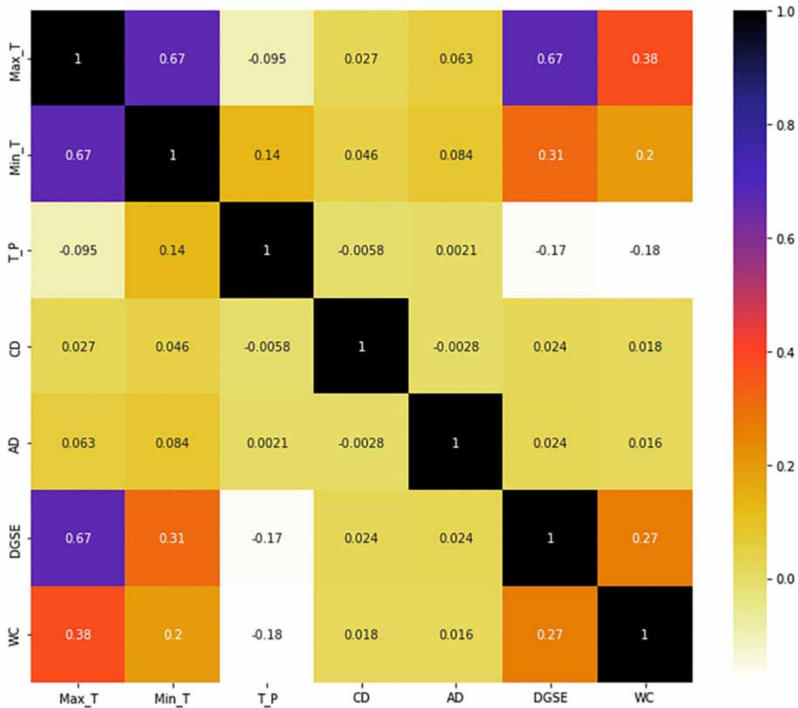


Figure 24 | Heatmap illustrating the Pearson correlation coefficients – Liverpool Water Plant.

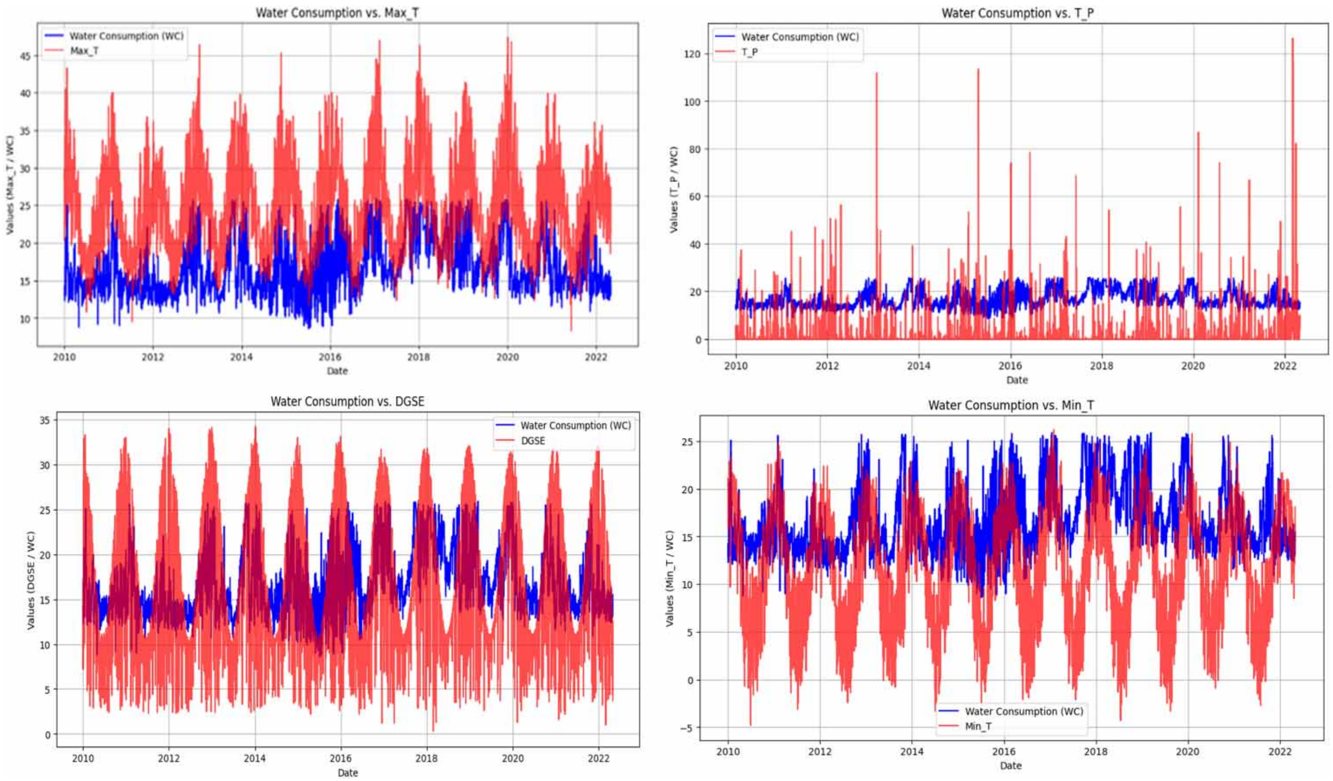


Figure 25 | Time series analysis of weather factors and water consumption – Richmond Water Plant.

The results from Figure 25 reveal a clear and significant relationship between weather variables, including maximum temperature, minimum temperature, rainfall amount, and daily global solar exposure, and water consumption. The evidence indicates that changes in these weather factors have a significant impact on water demand, with temperature and solar exposure exhibiting strong positive correlations with water consumption. This insight is particularly valuable for water resource management and planning, as it underscores the need for adaptive strategies in response to changing climate patterns. In contrast, our analyses show that the occurrence of special days, such as Australia Day and Christmas Day, has no discernible impact on water consumption trends in the region. These events appear to have minimal influence on water demand, suggesting that daily weather patterns play a more pivotal role. These findings have practical implications for water resource managers and policymakers. Understanding the dynamic relationship between weather variables and water consumption can inform more effective demand management strategies and resource allocation, especially in the face of climate change and increasing weather variability. By prioritizing strategies that consider the impact of weather on water demand, the Greater Sydney region can enhance its resilience and sustainability in the management of this critical resource.

5.4. Summary of key findings

The findings from our comprehensive analysis of water demand on Australia Day and Christmas Day, utilizing data from four water plants in the Greater Sydney region over the years 2010–2022 provide several important insights.

Time series visualization and descriptive statistics: Through time series visualization, we examined daily water consumption patterns for January (2010–2022) and December (2010–2021). Each plot highlighted water consumption on Australia Day and Christmas Day, respectively. Based on these visualizations, the data indicate that while there are daily fluctuations in water consumption, Australia Day and Christmas Day do not consistently exhibit significant deviations from typical consumption patterns observed during these months. Additionally, descriptive statistics were calculated for these holidays and compared against the rest of January and December. The results consistently showed that the mean and median water consumption on these holidays were within the range of daily fluctuations observed throughout the rest of the month. For instance, the mean water demand from the Richmond Water Plant on Australia Day was 18.37, closely matching the mean of 18.1 for the rest of January. Similarly, Christmas Day water consumption trends paralleled those of other days in December.

Box plot analysis: Box plots further illustrated the distribution of daily water consumption values, emphasizing the IQR and median consumption levels. These visualizations confirmed that water consumption on both Australia Day and Christmas Day did not significantly deviate from consumption patterns observed on other days in their respective months. The consistency in median values and overlapping IQRs suggests that these holidays do not introduce exceptional variations in water demand. In some years, the median consumption on these holidays was similar to that of other days in January and December, indicating no significant difference. While slight differences were observed in other years, they were not consistently higher or lower.

Statistical significance and *t*-test results: To solidify our findings, we conducted hypothesis tests, specifically *t*-tests, to determine if the differences in water consumption on Australia Day and Christmas Day were statistically significant when compared to the rest of January and December. The *p*-values obtained from these tests were greater than the typical significance threshold of 0.05, indicating that any observed differences were not statistically significant. These results reinforce the conclusion that these holidays do not have a substantial impact on water demand.

Pearson correlation heatmap analysis: Heatmaps using the Pearson correlation coefficient analyzed the relationship between water consumption, Australia Day, Christmas Day, and various weather variables across four different water plants. The heatmap results indicate a weak association between water consumption and both Christmas Day and Australia Day, suggesting that these holidays have little to no impact on water demand. In contrast, weather variables showed a stronger correlation with water consumption, emphasizing the influence of weather factors over holiday events.

The following findings are noteworthy regarding water consumption patterns in the Greater Sydney region: maximum temperature and daily global solar exposure show strong positive correlations with water consumption. Conversely, there is a clear negative correlation between water consumption and rainfall. Moreover, there is a notable relationship between water consumption and minimum temperature. These insights underscore how weather conditions significantly influence water consumption trends in this region.

Time series analysis of weather factors: Time series analysis was conducted to examine the impact of weather variables on water consumption. The results revealed significant correlations between water consumption and weather variables such as maximum temperature, minimum temperature, rainfall amount, and daily global solar exposure. These factors exhibited strong correlations with water consumption, highlighting their substantial influence on water demand.

Summary research findings: The comprehensive analysis, encompassing time series visualization, descriptive statistics, box plot analysis, *t*-tests, and Pearson correlation heatmaps consistently indicates that neither Australia Day nor Christmas Day significantly impacts water demand in the Greater Sydney region. Despite minor variations in daily water consumption, these holidays do not cause notable deviations from the typical daily patterns observed throughout January and December.

Implications for water management: Our findings suggest that water consumption is influenced more by other factors, such as weather conditions and overall seasonal trends, rather than by the occurrence of these specific holidays. This conclusion is vital for water resource management and planning, as it indicates that special measures to address potential spikes in water demand on these holidays may not be necessary.

Understanding the dynamic relationship between weather variables and water consumption can inform more effective demand management strategies and resource allocation, especially in the face of climate change and increasing weather variability. By prioritizing strategies that consider the impact of weather on water demand, the Greater Sydney region can enhance its resilience and sustainability in the management of this critical resource.

6. CONCLUSION AND RECOMMENDATION

The accuracy and effectiveness of water demand forecasting hinge critically on selecting inputs that significantly impact model performance (Ghannam & Hussain 2024). In our study, we explored the influence of holidays, specifically Australia Day and Christmas Day, alongside weather variables such as maximum temperature, minimum temperature, daily global solar exposure, and rainfall amount, on water consumption in Greater Sydney.

Our analysis reveals that holidays such as Australia Day and Christmas Day do not notably alter water consumption patterns compared to typical days. However, weather variables, particularly maximum temperature and daily global solar

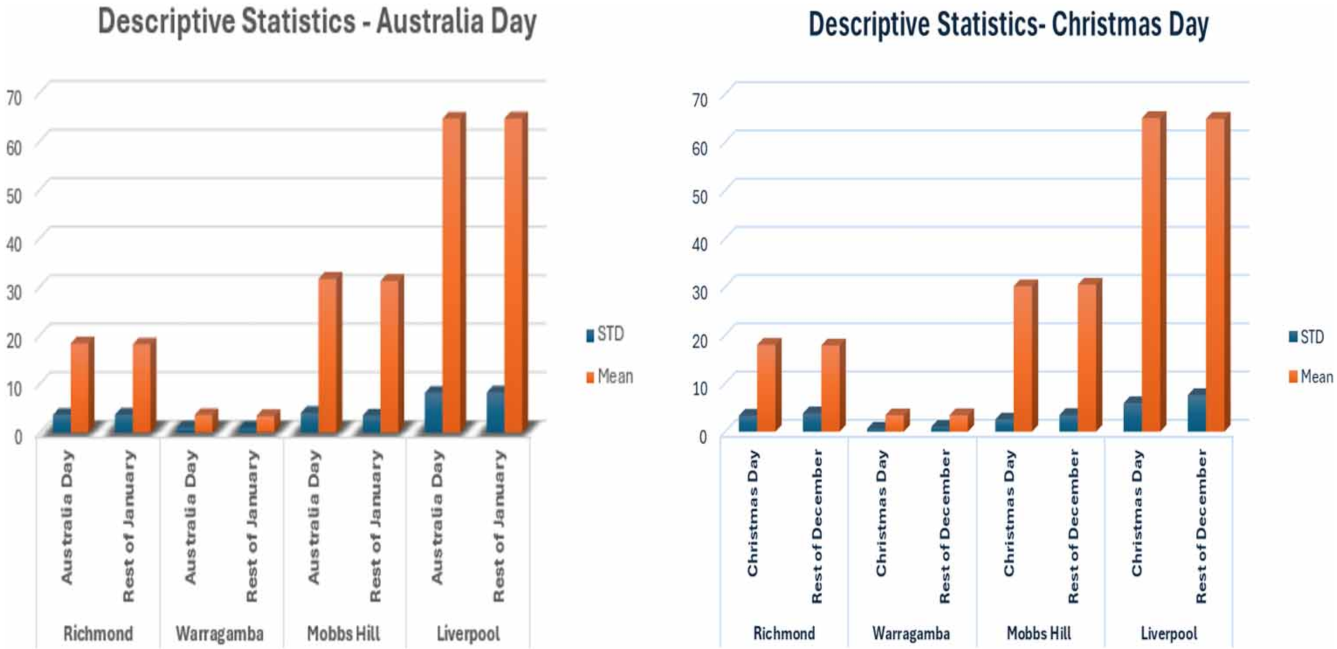


Figure 26 | Comparison of mean water demand and standard deviation between Australia Day and Christmas Day versus the rest of January and December across Four Water Plants in the Greater Sydney region.

exposure, exert a significant influence on water demand. This highlights the importance of monitoring these factors, especially during extreme weather events, to optimize water management strategies effectively.

The comparative analysis of mean water demand and standard deviation on Australia Day and Christmas Day versus the rest of January and December, as depicted in Figure 26, suggests that water demand remains relatively consistent across the Greater Sydney region. Both the mean water demand and standard deviation values for these holidays closely match those of the rest of their respective months. This indicates that major holidays do not significantly impact overall water consumption patterns in the region. Such findings imply that water consumption on these holidays is stable and comparable to other days in their respective months, suggesting that no special consideration for water planning is necessary for these occasions.

Understanding historical trends in holiday-related water demand is crucial. While our study found no significant deviations during holidays, historical data indicate evolving patterns influenced by urbanization and changing consumer behaviors. Climate change projections suggest intensifying weather extremes, which could unpredictably exacerbate water demand dynamics. Adaptive strategies must therefore anticipate these challenges, ensuring resilience in water resource management.

Although our focus is on Greater Sydney, our methodologies such as correlation heatmaps, *t*-tests, descriptive statistics, and time series analysis offer a robust framework applicable in diverse geographical contexts. This adaptability, combined with meticulous variable selection and analytical tools, positions our approach to address water management challenges beyond our study region.

In summary, our findings underscore the importance of selecting relevant inputs for accurate water demand forecasting models and emphasize the significant impact of weather variables on consumption patterns. These insights are pivotal for guiding effective water management strategies aimed at ensuring sustainable urban water use. By considering historical contexts and anticipating future climate impacts, policymakers can proactively address the challenges posed by climate change and fluctuations in holiday-related water demand, thereby promoting resilient and adaptive water management practices.

AUTHORS' CONTRIBUTIONS

S.G. contributed to conceptualization and methodology; wrote the first draft; validated the work; participated in formal analysis; wrote the original draft; and also reviewed and edited the manuscript. F.K.H. reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Beal, C. D. & Stewart, R. A. 2014 Identifying residential water end uses underpinning peak day and peak hour demand. *Journal of Water Resources Planning and Management* **140** (7), 04014008.

Capt, T., Mirchi, A., Kumar, S. & Walker, W. S. 2021 Urban water demand: Statistical optimization approach to modeling daily demand. *Journal of Water Resources Planning and Management* **147** (2), 04020105.

Coelho, B. & Andrade-Campos, A. G. 2019 Short-term forecasting of hourly water demands – A Portuguese case study. *International Journal of Water* **13** (2), 173–207.

Cole, G. & Stewart, R. A. 2013 Smart meter enabled disaggregation of urban peak water demand: Precursor to effective urban water planning. *Urban Water Journal* **10** (3), 174–194. doi: 10.1080/1573062X.2012.716446.

Domene, E. & Saurı, D. 2006 Urbanisation and water consumption: Influencing factors in the metropolitan region of Barcelona. *Urban Studies* **43** (9), 1605–1623.

Field, A. 2013 *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications, London, UK.

Ghannam, S. & Hussain, F. 2023 Comparison of deep learning approaches for forecasting urban short-term water demand a greater Sydney region case study. *Knowledge-Based Systems* **275**, 110660. doi:10.1016/j.knosys.2023.110660.

Ghannam, S. & Hussain, F. 2024 Short-term water demand forecasting: A review. *Australasian Journal of Water Resources* 1–19. doi:10.1080/13241583.2024.2350102.

Guo, H., Liu, X. & Zhang, Q. 2024 Identifying daily water consumption patterns based on K-means clustering, agglomerative hierarchical clustering, and spectral clustering algorithms. AQUA – water infrastructure. *Ecosystems and Society* **73** (5), 870–887.

Hussien, W. A., Memon, F. A. & Savic, D. A. 2016 Assessing and modelling the influence of household characteristics on per capita water consumption. *Water Resources Management* **30**, 2931–2955.

Kavya, M., Mathew, A., Shekar, P. R. & Sarwesh, P. 2023 Short term water demand forecast modelling using artificial intelligence for smart water management. *Sustainable Cities and Society* **95**, 104610.

Makki, A. A., Stewart, R. A., Beal, C. D. & Panuwatwanich, K. 2015 Novel bottom-up urban water demand forecasting model: Revealing the determinants, drivers and predictors of residential indoor end-use consumption. *Resources, Conservation and Recycling* **95**, 15–37.

Myka-Raduj, A., Jóźwiakowski, K., Siwiec, T. & Raduj, W. 2023 Changes of water consumption in a Forester's lodge in Polesie National Park (Poland)—Case study. *Water* **15** (17), 3157.

Quilty, J., Adamowski, J., Khalil, B. & Rathinasamy, M. 2016 Boot-strap rank-ordered conditional mutual information (broCMI): A nonlinear input variable selection method for water resources modeling. *Water Resources Research* **52** (3), 2299–2326.

Romano, M. & Kapelan, Z. 2014 Adaptive water demand forecasting for near real-time management of smart water distribution systems. *Environmental Modelling & Software* **60**, 265–276.

Vijai, P. & Sivakumar, P. B. 2018 Performance comparison of techniques for water demand forecasting. *Procedia Computer Science* **143**, 258–266.

Zhou, S. L., McMahon, T. A., Walton, A. & Lewis, J. 2000 Forecasting daily urban water demand: A case study of Melbourne. *Journal of Hydrology* **236** (3–4), 153–164.

Zubaidi, S. L., Gharghan, S. K., Dooley, J., Alkhaddar, R. M. & Abdellatif, M. 2018 Short-term urban water demand prediction considering weather factors. *Water Resources Management* **32** (14), 4527–4542.

First received 16 May 2024; accepted in revised form 10 August 2024. Available online 20 September 2024