water e-journal
Online Journal of the Australian Water Association
ISSN 2206-1991
Volume 1 No 1 2016
http://dx.doi.org/10.21139/wej.2016.002

AUTOMATED ANALYSIS OF METERED WATER CONSUMPTION

A REVIEW OF THE AUTOMATED ANALYSIS PROCESS FOR METERED CONSUMPTION DATA DEVELOPED BY SA WATER

S Zhong, N Turich, P Hayde

INTRODUCTION

The acceleration in the scope, scale and economic impact of technology has advanced at an unprecedented pace. Studies suggest that up to 90% of data worldwide has been generated within the past few years. The effective use and analysis of the ever-growing datasets in a timely manner will support utilities in making strategic business decisions in a competitive environment.

The South Australian Water Corporation (SA Water) is responsible for supplying drinking water through more than 24,000km of water mains to more than 670,000 connections across the state. SA Water collects and stores large amounts of customer meter data on a proprietary customer-billing platform. Historically, the problem with analysing meter consumption data was the volume of data and general accessibility. These obstacles made data processing and analysis both labour intensive and time consuming. In the current regulatory environment this approach was no longer acceptable, and a more timely and cost-effective approach was required that could provide whole-of-system outcomes.

To increase business agility and resilience, SA Water has developed an automated analysis process of metered water consumption. The process uses code and tools across Microsoft and ArcGIS platforms. The new process creates the ability to carry out annual detailed analysis and publication of information for metropolitan water treatment plant (WTP) areas (6), metropolitan suburbs (413), regional districts (31), regional towns (252) and groundwater supply areas (40). The analytical tool aims to systematically:

- Provide a rapid method for analysis (2–3 days versus several weeks);
- Analyse SA Water customer metered consumption data both temporally and spatially;
- Project future consumption and connection growth;
- Overlay metered consumption data with census and socio-economic data to further understand differences between each suburb's individual consumption characteristics;
- Build information management capacities to inform sustainable long-term planning.

The development of these new tools allows a deeper understanding of customer consumption across the entire network, at different spatial scales. This provides valuable information to inform infrastructure planning processes and business decisions. This paper provides details of the automated analysis process developed by SA Water.

METHODOLOGY

The automated analysis can be broken down into four major steps: Base Data and Spatial Join; Data Extraction; Consumption Analytics; and Report Publication. The four basic steps are illustrated in Figure 1.

Core SA Water Data The core data that is analysed as part of this process is composed of customer meter consumption data and customer water meter location data. These two data

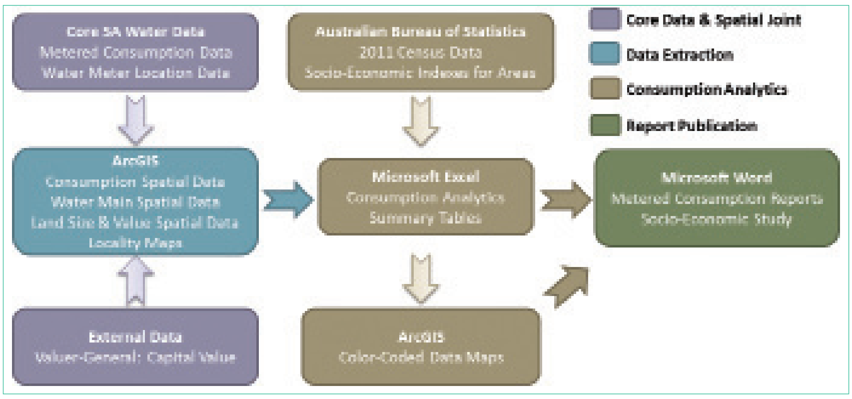


Figure 1. Process flowchart of the automated analysis of metered water consumption.

sets normally reside on two separate platforms. The customer meter consumption data is native to the corporate billing platform, while the customer water meter location data is native to the corporate ArcGIS environment. The metered consumption data is compiled and has recently become accessible in the ArcGIS environment following a spatial join based on the water meter Service Point Number (SPN) location. All personal customer identifiers that normally reside in the billing dataset are removed.

External Data External datasets used include both Australian Bureau of Statistics (ABS) 2011 Census data and land valuation data from the State Valuation Office. Using the ABS TableBuilder Basic tool, available via a free download, SA Water sourced the 2011 Census population data at mesh block scale. Household income information and socio-index information was sourced at the suburb level. In the consumption versus income analysis, we used the Census 2011 – Employment, Income and Unpaid Work database to extract the dwelling counts in different Total Personal Income (Income) ranges for all suburbs in South Australia.

Similarly, there are four types of Socio-Economic Indexes assigned to each state suburb:

- State Suburb (SSC) Index of Relative Socio-economic Advantage and Disadvantage, 2011;
- State Suburb (SSC) Index of Relative Socio-economic Disadvantage, 2011;
- State Suburb (SSC) Index of Economic Resources, 2011;
- State Suburb (SSC) Index of Education and Occupation, 2011.

These indexes are the combinations of a number of factors (e.g. income, education, occupation etc) with different weighting. The detailed description of each index can be found in the *Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA)* report.

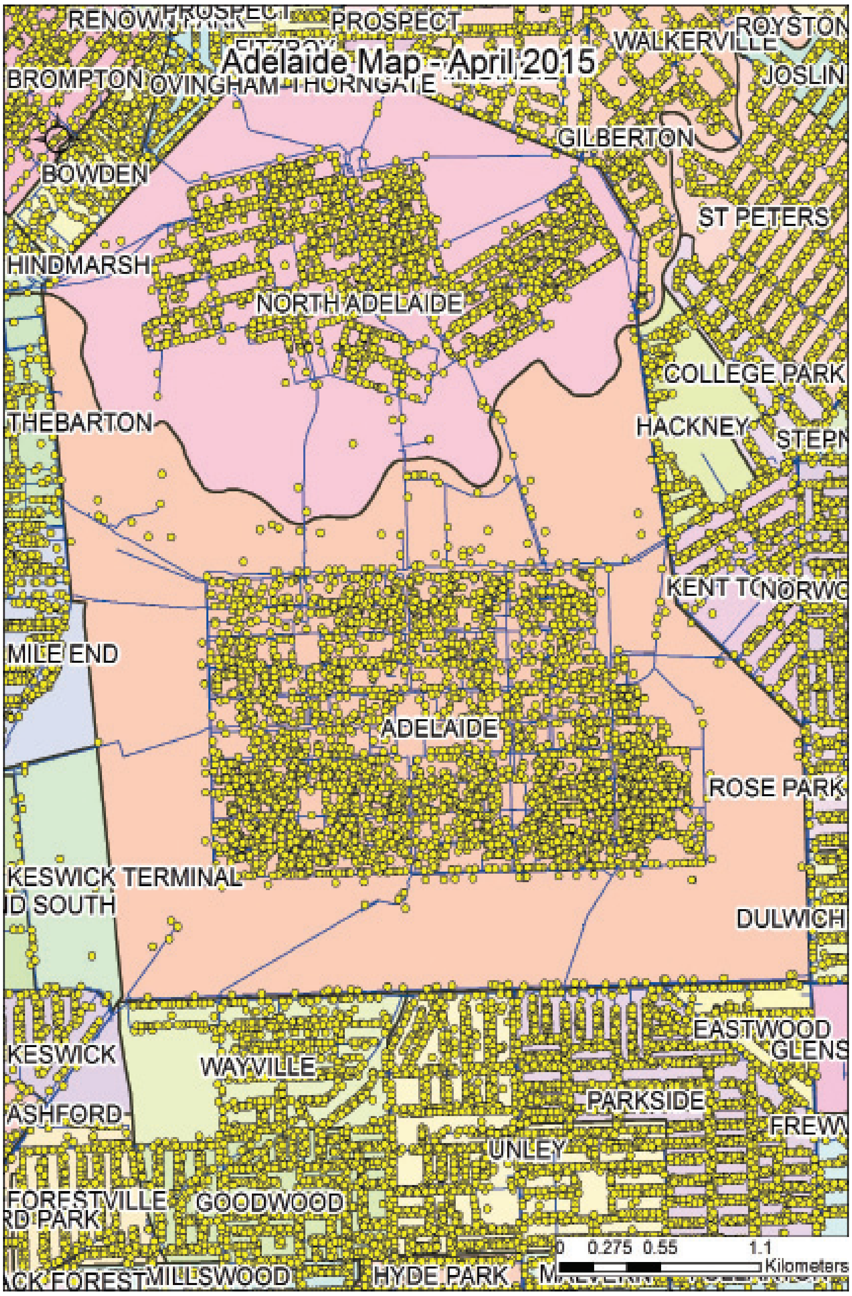


Figure 2. An example output map of the Adelaide CBD exported from ArcGIS.

The State Valuation Office, under the direction of the Valuer-General, is part of the Land Services Group within the Lands and Services SA division of the Department of Planning, Transport and Infrastructure. The State Valuation Office re-values all properties in the state annually, including ‘site value’ and ‘capital value’. SA Water is a ratings agency and uses the capital value to calculate customer sewerage rates. The capital value information was internally available to the T&NP team. The capital value was joined to the metered consumption data as part of a two-step process in ArcGIS.

Data Extraction The availability of metered consumption data within an ArcGIS environment has allowed for the use of innovative spatial queries and subsequent automation. The data extraction process was programmed using the ArcGIS ModelBuilder tool, providing a step-by-step flow chart tool that allowed for easy customisation of geo-processing workflows. Spatial areas of interest were identified at various scales from the WTP supply areas in the metropolitan area, regional district areas, metropolitan area suburbs and regional towns at a larger scale. Each of these areas was assigned either an existing polygon from the GIS data warehouse or a newly created polygon. The extraction process enables the selection of all the SPNs and associated customer metered data within each spatial boundary, as separate text files (CSV files). This same process may be adapted to any related dataset available within the ArcGIS platform.

As an example, to output customer information at a suburb scale, each SPN point, illustrated as a yellow dot in Figure 2, has been assigned within the suburb in which it is found. The program then selects all the SPNs with the same suburb name, compiles the data and outputs the meter information into a single CSV file. Similarly, the water main asset data within each area is also extracted using a modified version of the same process. This extraction process is automatically repeated so that all requested data for each suburb is outputted sequentially as part of a single batch run.

A locality map is published for each spatial area, centred on the spatial area of the extract, as part of the automated process. The locality map assists to visually verify that the extraction process has been undertaken correctly. An example of the locality map is given in Figure 2, highlighting the current water mains, SPN locations and boundary of the specific area being investigated.

Consumption Analytics The detailed analysis of customer-metered consumption is undertaken using Microsoft Excel. A sophisticated program has been developed to read and process the data files generated from the ArcGIS geo-processing. Once the raw customer data is imported into Excel, an automated data cleaning process is undertaken to remove customer meters that contain erroneous information that can potentially halt the analytic process.

Due to the distinctive nature of residential customers compared to non-residential customers, data from these two types of customer are segmented and analysed separately. Several charts are automatically generated to provide insights to the customer consumption within each area of interest:

- Annual consumption records of residential and non-residential customers;
- Total number of meters and active meters (consumption > 0 kL/annum);
- Average consumption per residential and non-residential customer;
- Consumption distribution within the spatial area boundary;
- Cumulative frequency of customer consumption in the current financial year;
- Ten largest customers as a percentage of total consumption;
- Ten-year annual demand projection, based on rolling three-year average;
- Consumption per unit length of water main;
- Number of active meters per unit length of water main.

While generating multiple charts, a summary table is created simultaneously. The table allows each area, at different spatial scales, to be ranked according to selected parameters. Once all the analytic and statistical results have been tabulated into a summary table, it is then exported back into ArcGIS and joined to each relevant area. Various colour-coded data maps are then generated for selected parameters to communicate visually the variation between different areas, both temporally and spatially.

To further enhance the understanding of customers, additional ABS data is incorporated into the analysis of the Adelaide metropolitan suburbs. Household income and various socio-economic indexes are used to compare against residential consumption to understand how the socio-economic level can influence water consumption.

The consumption results specific to metropolitan suburb No.155 are presented as eight charts shown in Figure 3. Each chart offers a different insight into the metered consumption characteristics of this area. The first two charts show a continuous decrease in total metered consumption from the customers located within the suburb boundary, despite the fact that the number of active water meters has been gradually increasing. This means that the average consumption per customer meter within the suburb has decreased, due to a combination of water restrictions, conservative consumption behaviour and the installation of more water-efficient appliances and fittings. The 25-year projections of consumption and connection growth are calculated using the average annual growth over the last three years. The growth in water meter connections has been driven by residential customers, who represent 95% of the total consumption in the demand distribution pie chart.

The historic consumption chart reveals that there has been a significant drop in average non-residential consumption, from 1,180 kL/annum to less than 600 kL/annum. Meanwhile, average residential customer consumption reduced from 370 kL/annum to 200 kL/annum. The lower left-hand chart highlights that the majority of customers consume less than 350 kL/annum. The lower right-hand chart shows that the top 10 consumers within the suburb have used around 6% of the total consumption over the past 20 years.

Report Publication Because the batch-run analysis process produces thousands of charts and tables, a Macro program has been developed in Microsoft Word to automatically read, copy and format charts from the previous step and publish the results in a suite of annual reports. This has proven to be a significant *(continued after Figure 3)*

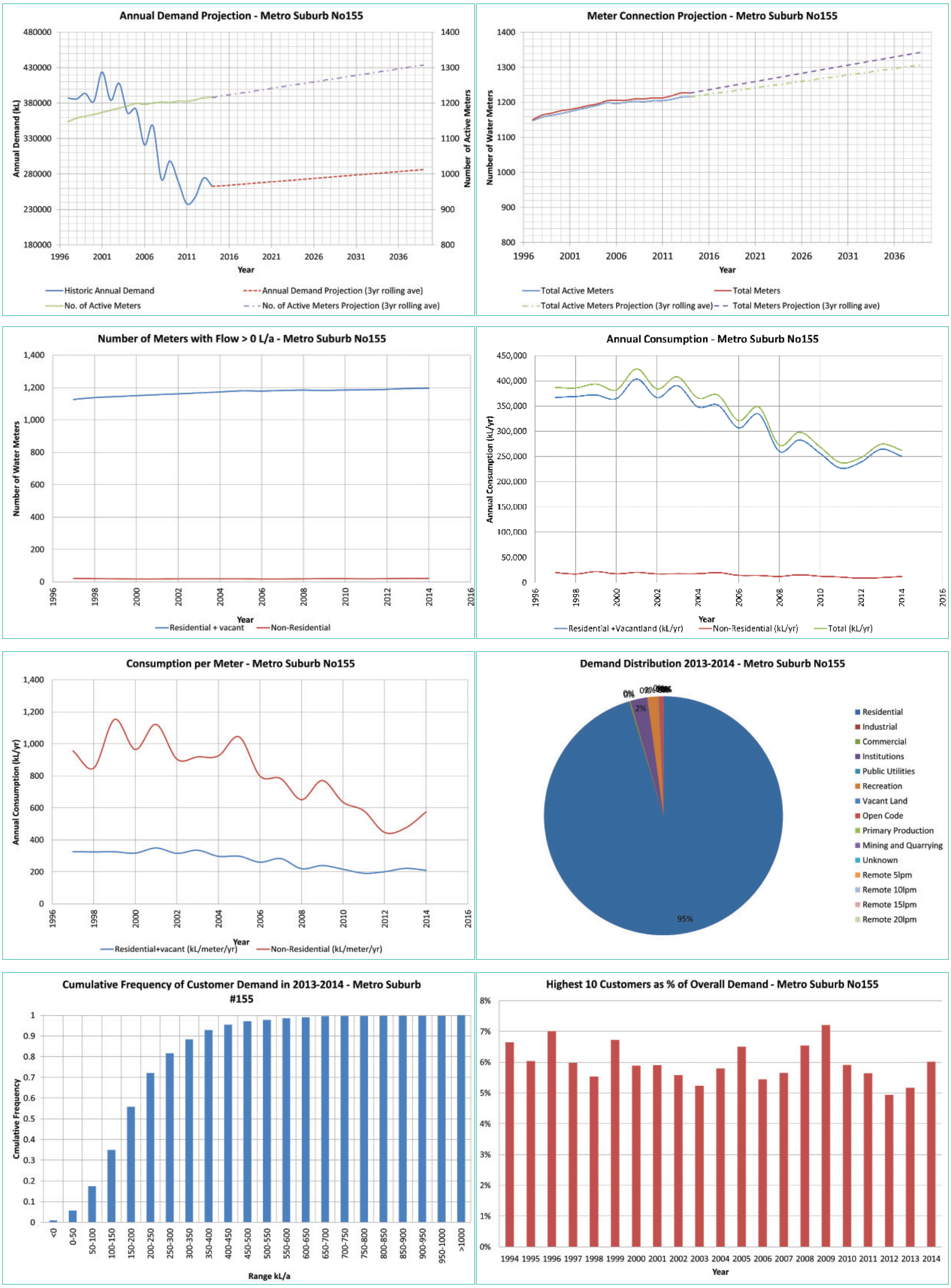


Figure 3. Automated analysis of metered water consumption results for metropolitan suburb #155.

time-saving advancement for what would previously have been a time-consuming task. The code used for publishing reports can be easily adapted to include any additional analytic results produced in future and incorporate more advanced analytical process outputs.

RESULTS AND INSIGHTS

Many insights were obtained from the information generated, which includes the ability to rapidly quantify long-held anecdotal information. Internal business units now have readily available data in a spatially relevant format to temporally monitor a suite of factors, highlight changes and update the decision-making processes accordingly.

Insights gained from the data included the following:

- Quantification of the degree of disparity between connection growth rates in metropolitan and country areas, and between country towns and country lands;
- The relative weighting of country towns' versus country lands' annual demand varied by district;
- Average annual consumption in suburbs with a higher socio-economic index is showing a higher 'bounce-back' post-restriction than those with a lower socio-economic index. Inversely, the data points to where customers may be in financial stress;
- At a suburb scale the average annual consumption was not necessarily determined by land size; however, within each suburb, the relationship between annual demand and land size was more evident;
- There was a relatively high coefficient of determination between the average annual consumption at the metro suburb scale and respective socio-economic indexes;
- Relative high consumption of top 10 customers in some areas highlighted some vulnerability to shifts in the industrial profile of the economy;
- In most areas the total annual demand has dropped, despite an increase in customer meter connections;
- By adopting a systems approach we are able to rank all areas using a range of indicators. This will assist in focusing our planning resources, our understanding of available capacity, informing timing of asset upgrades and investment, and provide insights into customer behaviour. Updated annual reporting of customer demands spatially, and reprioritisation of investments, may lend themselves to delayed capital expenditure and corporate savings.

FUTURE APPLICATIONS

The current method is an automated process for creating meaningful analytics of metered annual consumption data across the entire network at various spatial scales in a single pass. Opportunities under consideration for the future include:

- Integration of a regression model to future demand projections for each spatial area as part of the automated process.
- Extend analytics to water main asset performance and water quality information.
- Extend the reach of analysis to more external databases in the search for further insights into customer behaviours.
- Integrate further with hydraulic modelling platforms to build better predictive tools for capacity assessments.
- Annually publish reports for metered water consumption across the entire network at all spatial scales.
- The method can be easily applied to analyse consumption for recycled water meters.
- Expand the annual reports to include analysis of quarterly customer meter data.
- The method has scope to be adapted to a study of the sewerage network.

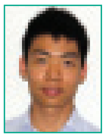
CONCLUSION

A rapid, low-cost and robust automated analysis tool was created, based on metered water consumption data, which will significantly contribute to the business intelligence of SA Water. The analytics allows SA Water to make better-informed infrastructure planning decisions within a regulatory environment. The process developed for use with the ArcGIS and MS platforms provides accurate information in a rapid manner and can be easily adapted for new applications depending on the needs of the business. The process was developed internally using available software platforms and publicly available external datasets. The automation process allows for a vastly enhanced spatial and temporal analysis to be performed rapidly in comparison to the previous manually based practices.

ACKNOWLEDGEMENTS

The automation process relies on and leverages off the metered water consumption data conversion to Service Point Number (SPN) work carried out by the Asset Information team within SA Water.

THE AUTHORS



Shane Zhong (email: shane.zhong@sawater.com.au) is a Graduate Engineer, Treatment & Network Planning, Strategy & Planning, SA Water. Shane is currently studying for a Graduate Diploma of Data Science at Monash University.

Nick Turich (email: nick.turich@sawater.com.au) is a Senior Planner, Treatment and Network Planning, Strategy and Planning, SA Water. Nick has undertaken planning and projection work for water, sewerage and recycled water over the past six years.

Patrick Hayde (email: patrick.hayde@sawater.com.au) is Manager for Water in the Treatment and Network Planning team at SA Water. Patrick has 20 years of experience in water modelling, design and infrastructure planning.