Water Supply

water e-journal
Online Journal of the Australian Water Association

ISSN 2206-1991
Volume 3 No 4 2018
https://doi.org/10.21139/wej.2018.041

Enhancements and Validation of a Demand Forecast Tool for South Australian Water Corporation

Managing complexities in operations decision making

A Duncker

ABSTRACT

SA Water has undertaken a project to expand and enhance a set of decision support tools for its Operations Control Group (OCG). The Demand Forecast Tool (DFT) is a critical component of the tools and is used daily as part of 'business as usual'. This paper details the improvements made to the DFT, including a new process for developing and calibrating the DFT regression model, geographic expansion, region specific input data and new data sources (customer and network flow) to validate and improve forecast accuracy and inputs into other decision support tools, such as a live hydraulic model for operations (Network Operations Model – NOM).

Key Words: Demand forecasting; Machine Learning; household water use; smart meters; live hydraulic modelling

INTRODUCTION

The South Australian Water Corporation (SA Water) has been using a set of Decision Support Tools developed for its Operations Control Group (OCG) to help manage and optimise the operations of the water network and manage complexities in operations decision making, production planning and in-network transfers. The Demand Forecast Tool (DFT) is a critical component and is used on a daily basis as part of 'business as usual' for the OCG.

The DFT was developed using a multi-variable non-linear regression model to forecast demand based on climate variables such as rainfall, maximum temperature and soil moisture index (MWH, 2011). This model was built in Bentley's Amulet analytics platform and automated to generate a forecast each day. Using the tool over the last four years, the business has identified opportunities for improvements that would enable greater forecast accuracy and flexibility for further improvements or expansion.

This paper outlines a recent project undertaken to implement a number of improvements to the DFT, taking advantage of additional network monitoring and customer water use information now available as well as advances in software, including 'off the shelf' software packages for machine learning applications.

The project delivered the following:

- Flexibility and improved user interaction for demand forecast model build, calibration and forecasting through Microsoft Azure Machine Learning
- Model validation through long-term customer smart meter data from the University of Adelaide and network flow monitoring from SA Water network and water treatment plant flow meters
- Improved demand forecast accuracy, including improved geographic and customer sector demand breakdown and expansion to the Southern Myponga service area

METHODOLOGY

Over the last four years SA Water has used the DFT as part of 'business as usual' to help plan short term network operations, water treatment plant production schedules and as an input to a live hydraulic model of the metropolitan water network.

The project scope included a geographic expansion of the DFT as well as a number of enhancements to the DFT and other components of the suite of Decision Support Tools (DSTs) used by the OCG. The requirements identified by SA Water were critical for the continued use and performance accuracy of the DFT.

As this was a technology and IT project, the project team needed to work through the steps below to understand the changes required and, how to integrate the required updates with the existing tools.

1. Review of Business Requirements
2. Develop high-level scope of work statements to meet the requirements
3. Work with the business to develop detailed function requirements
4. Develop concept design and detailed design for enhancements
5. Delivery and testing phase
6. User Acceptance Testing
7. Go-Live

Business requirements critical to the DFT include:

- Improvements during DFT model calibration phase, including additional model validation steps, accuracy parameters and easier user access to setup of calibration models and data input
- Review and improvement of diurnal curve data sets and customer sector assumptions for internal and external water use (i.e. climate affected demand)
- Review of major customer types and available monitoring for customer specific diurnal curves and benefits of configuring additional customer sectors in the DFT for major customers
- Validation of demand forecast, particularly spatial breakdown of demand forecast against network flow field monitoring meter data and 24-hour use patterns
- Expansion of the demand forecast tool to Southern Myponga service area (capturing unique trends associated with these customers (i.e. low occupancy, holiday and weekend peaks).

Figure 1 gives an overview of the DFT process and enhancements added through this project.

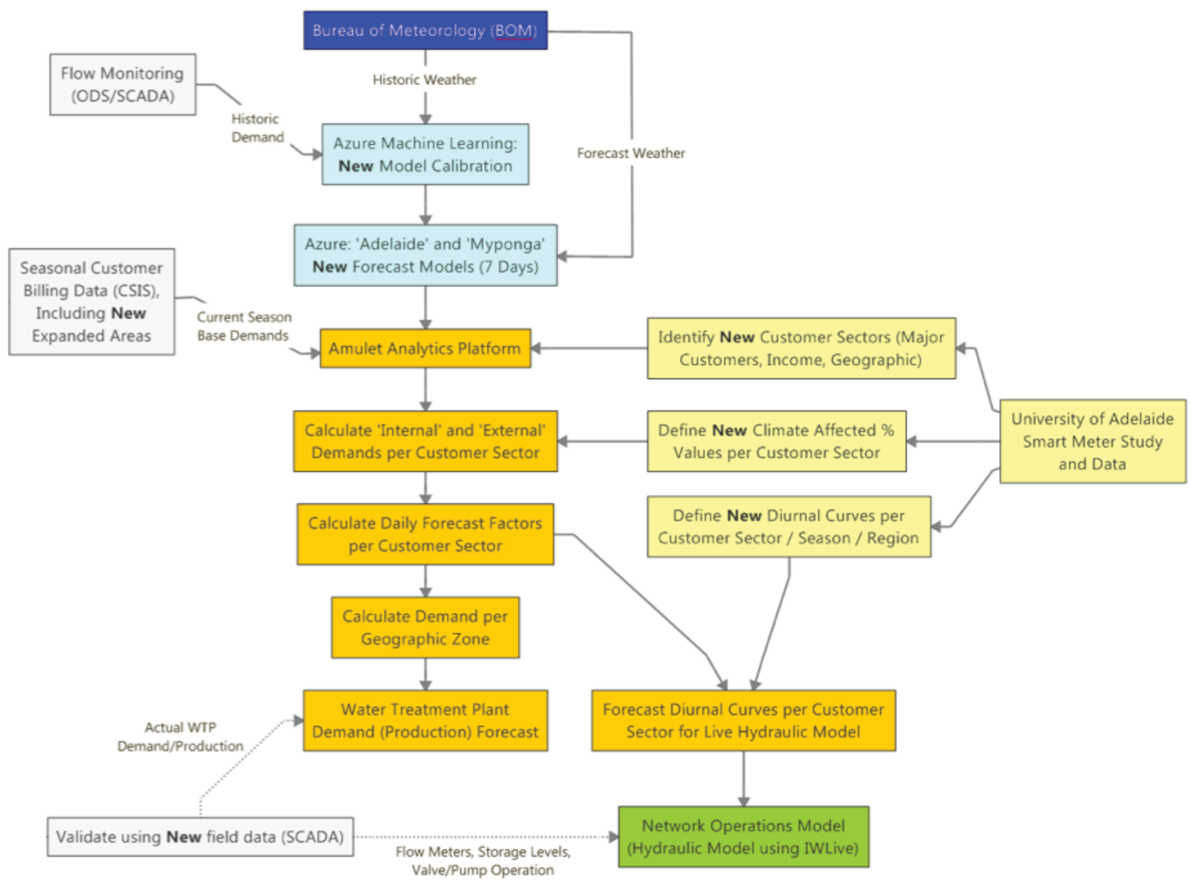

Figure 1: Demand Forecast Tool process and enhancements

Expansion of the geographic area of the DFT

The project was tasked with delivering two separate demand forecast models to cover the expanded geographic area; one model for the Adelaide Metropolitan area and one for the Myponga demand area (Figure 2). The Myponga service area covers the major townships in the Fleurieu Peninsula which is a tourism region with a signification proportion of holiday homes. A separate demand model was required to account for the different customer types, occupancy and weather compared with the metropolitan area.

The DFT has three main calculation steps:

1. Regional Forecast:

- Adelaide
- Myponga

2. Demand Area Forecast:

- 130 Demand Areas
- 13 Adelaide Customer Categories
- 10 Myponga Customer Categories
- 50 Major Customer Categories

3. Water Treatment Supply Forecast:

- Mixed supply calculations
- 6 Water Treatment Plants (WTP)

Figure 2: Adelaide and Myponga Demand Forecast model areas and water treatment plant supply areas

As illustrated in Figure 1, the regional forecast develops a 7-Day demand forecast for the region using forecast climate data. The demand area forecast uses the regional forecast per day, seasonal billing data per customer category and per demand area as well as external water use percentage per category to break down the forecast to geographic regions.

The WTP forecast then uses inputs from the operators on planned operating modes of the network for the next 7 days (i.e. closed boundaries, transfer modes, etc.) to identify the source of supply for each demand area (including mixed supply percentages) and calculate total WTP forecast per treatment plant for the next 7 days, see Figure 1.

Microsoft Azure Machine Learning (ML)

The Regional Forecast component of the DFT has been improved in this project by using the Microsoft Azure Machine Learning 'off the shelf' software platform. Rebuilding the regression model component of the tool with Azure ML has allowed for greater flexibility and user access in calibrating the forecast models, selecting and testing input parameters and selecting training data sets.

Similar to the original demand forecast model build in Bentley's Amulet analytics platform, the predictive model built using Azure ML produces forecast and hindcast demand at a regional level. The forecast demand is used to support daily planning of WTP demand for the next 7 days and input to a live operational hydraulic model. Hindcasts are also calculated to provide an indication of demand forecast performance when uncertainty in weather forecast data is removed.

Predictive models for the Adelaide and Myponga regions have been developed, trained and evaluated in Azure ML to determine the preferred machine learning algorithm to apply in order to achieve the best prediction of water demand in each region.

As illustrated in Figure 3, numerous input parameters are used in the training of the predictive models including:

- Historical demand
- Climate (historical and forecast rainfall and maximum temperature, including time lagged values)
- Calculated soil moisture index (SMI, included to capture changes in demand response due to past rainfall events). The calculation of SMI is based on fixed (but editable) equation coefficients that have been taken from the Amulet demand forecasting tool and stored in the SA Water Enterprise Data Warehouse (EDW).
- Calendar (days of the week, holiday periods, etc.)
- Population

The ongoing prediction of demand for each region is handled by an Extract Transform Load (ETL) process. This process collates data to be provided to Azure ML, calls a deployed web service to run the predictive model, and publishes the resulting predictions to the SA Water's Enterprise Data Warehouse.

The calculated forecast demand and hindcast demand data published to EDW is subsequently picked up by Amulet, where there are additional calculations to break the demand down to the customer category level and assist with the determination of demand distribution throughout the network at the demand area level.

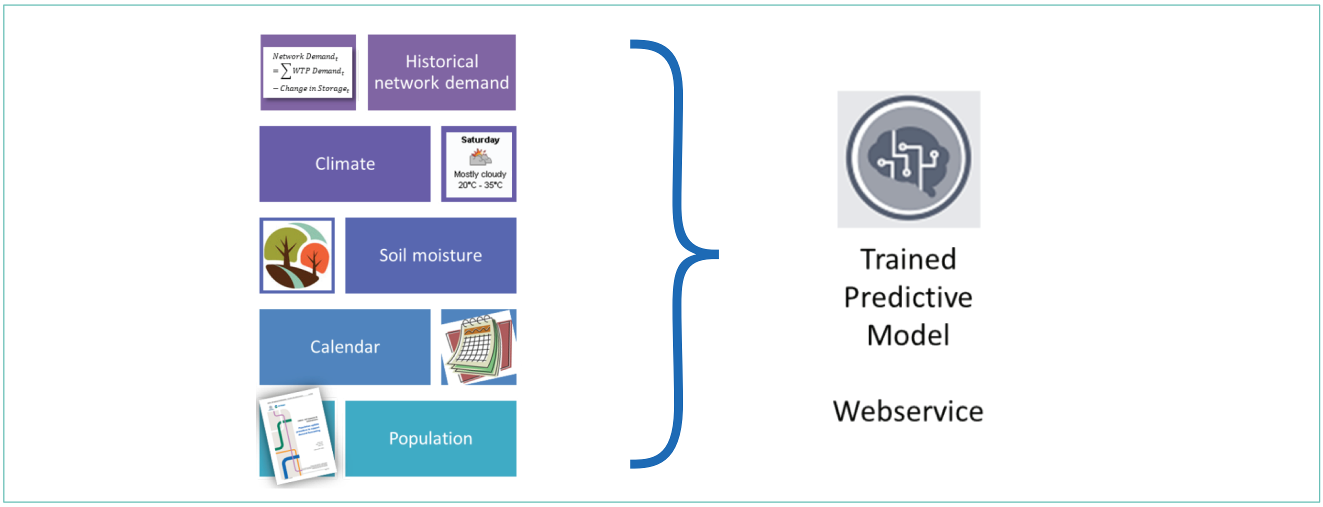

Figure 3: Azure ML demand forecast model

Azure ML allows users to access best-in class algorithms that can be used through simple drag and drop functionality to test, build and deploy analytics and models. The web-service functionality of the tool also allows the deployed service to be called from any device or location and can use any data source. By updating the demand forecast tool calibration to use the Azure ML platform, the team has also been able to integrate the new forecast outputs with the existing decision support tools and processes and limit changes to the existing IT architecture of the tools.

Figure 3 shows how the new Azure ML demand forecast model steps have been integrated into the overall DFT design and the end-to-end DFT process. Some of the advantages of changing the Regional Forecast calculations from a component within the Amulet platform to the Azure ML platform include:

- Flexibility to apply different regression methods
- Better tools for validation (training vs validation data sets) of the forecast models
- Providing the better access for the business to machine learning tools
- Ability to test forecast model scenarios and alternative methods/data without making a commitment to the outcomes
- Drag and drop interaction within the user interface for model building (data, parameters)
- Upgrade to new technology (machine learning) not available when the initial project was developed

Household water use study data

The University of Adelaide (UoA) were engaged on this project to provide data, insights and updates on water use patterns used in the original DFT. UoA has been involved in a water use study as part of the Goyder Optimal Water Resource Mix for Metropolitan Adelaide (OWRM) project.

The OWRM project included continuous sub-daily monitoring of household demand and detailed surveying of the household demographics and characteristics. Additionally, smart meter data was collected by UoA for three years (March 2013 to July 2016) from 120 of the 150 households involved in the OWRM project. Together this provides a unique data set of South Australian water use and drivers which incorporates the variability between seasons and households.

The OWRM study households were chosen to represent metropolitan Adelaide based on the demographics including income and age. The study households represent 60-65 percent of the households in metropolitan Adelaide with regard to demographics and dwelling structure (owner occupied established detached households) (Arbon et.al, 2014). Due to study participation restraints under-represented households included low income, single parent family and non-family households. Units, townhouses and renters were excluded. Figure 4 shows the approximate household locations and percentage of study households within the demand areas defined in the DFT.

The key outcomes from the OWRM project (Arbon et.al, 2014, Thyer 2015) relevant to the enhancement of the DFT included:

Detailed information and insights on Adelaide water use at the end-use and household scale. This includes information on water use patterns such as diurnal curves, internal/external water use proportions and the drivers of changes in water use patterns.

Distinct household usage types (classified in terms of income/age/attitude) were identified with different usage patterns due to differences in appliance type and behaviour. Areas with different proportions of these household types will have different water use and water use patterns.

Evaluation of the Behavioural Stochastic End-use (BESS) framework which predicted end-uses by explicitly considering appliance type and behavioural drivers of household water use. BESS was found to provide reliable predictions (<5% errors) of indoor end-use totals and proportions. BESS was used for predictions of past, present and future water use for metropolitan Adelaide.

The additional monitoring of the households provided outdoor water use variability between seasons and between households.

The assumed water use patterns (diurnal curves and climate affected indoor/outdoor patterns) in the DFT were compared and evaluated against the seasonal water use patterns in the UoA smart meter data. The potential effect of the variation in the water use patterns due to demographics was also evaluated at the Demand Area (as shown in Figure 4) level using data from the Australian Bureau of Statistics.

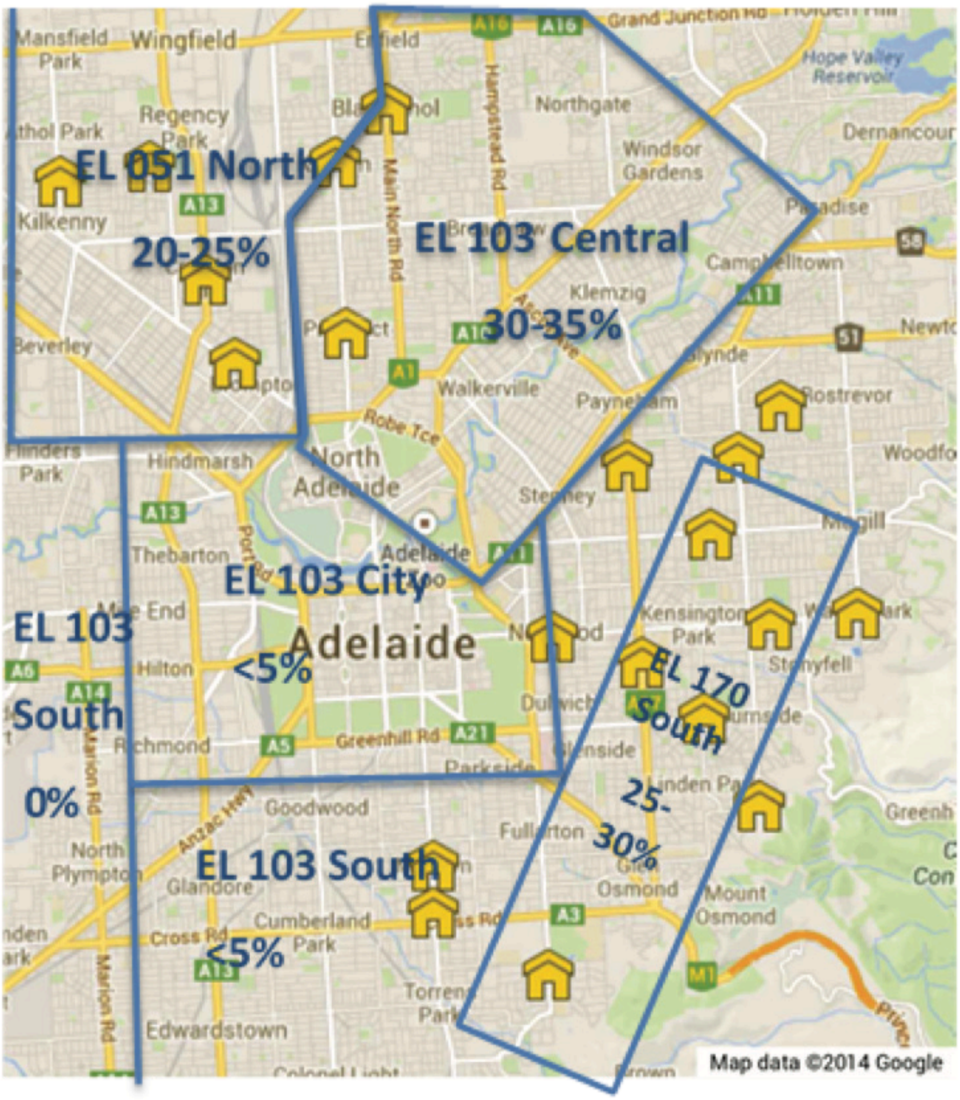

Figure 4: Smart water meter household monitoring sites across model demand areas

UoA was also engaged to help evaluate the assumptions needed for customers in the Myponga areas. As data from OWRM could not be directly applied as Myponga is outside of metropolitan Adelaide, the BESS framework was used to assess the water use drivers which have the greatest impact on the assumptions of the demand forecasting tool.

Field monitoring and billing data

As part of the North South Interconnector System project completed in 2013, 50 new flow meters were installed across the Adelaide metropolitan water network at key locations. These locations represented inlets and outlets at storages and critical flow paths (in-network transfers). This data allows continuous monitoring of network performance across almost the entire metropolitan network and has provided a valuable source of long-term data for validation of the live hydraulic model and the demand forecast tool.

In addition to flow monitoring, a number of major customers have also been set up with smart meters to record interval data on consumption. These major customers are significant water users and many have unique water use patterns that vary seasonally or between weekends and weekdays.

Using major customer flow data, consumption trends for weekday verses weekend and differences between typical commercial, industrial or recreational patterns normally used for these major customers could be understood.

The network flow meter data was used as validation data for the live hydraulic model. Comparing model performance against field data for an average day or peak day also allowed the demand assumptions and forecast to be tested and validated. Differences between total flows (mass balance to a group of zones) and 24-hour flow patterns were used to test and improve forecast accuracy (namely distribution per demand area) and diurnal patterns.

DISCUSSION AND RESULTS

Azure model training

Results of the model training for the regional forecast calculation undertaken for the DST expansion project demonstrated good performance for the Adelaide region and satisfactory performance for the Myponga region. Data available for training the Myponga model has some limitations which were identified as part of the project. In particular the measurement of volume in the large Nettle Hill storage tank. SA Water plans to retrain the Myponga model when improved data is available (through new flow monitoring and/or recalibration of flow meters).

Figure 5 shows some of the model calibration metrics generated when training the Adelaide demand regression model using Azure ML.

The Adelaide model achieved a Coefficient of Determination (R2) close to 0.95. Data used as part of the model development is divided into both a training and validation data set, which was one of SA Water's key business requirements in developing a robust demand forecast model for the regional forecast calculation.

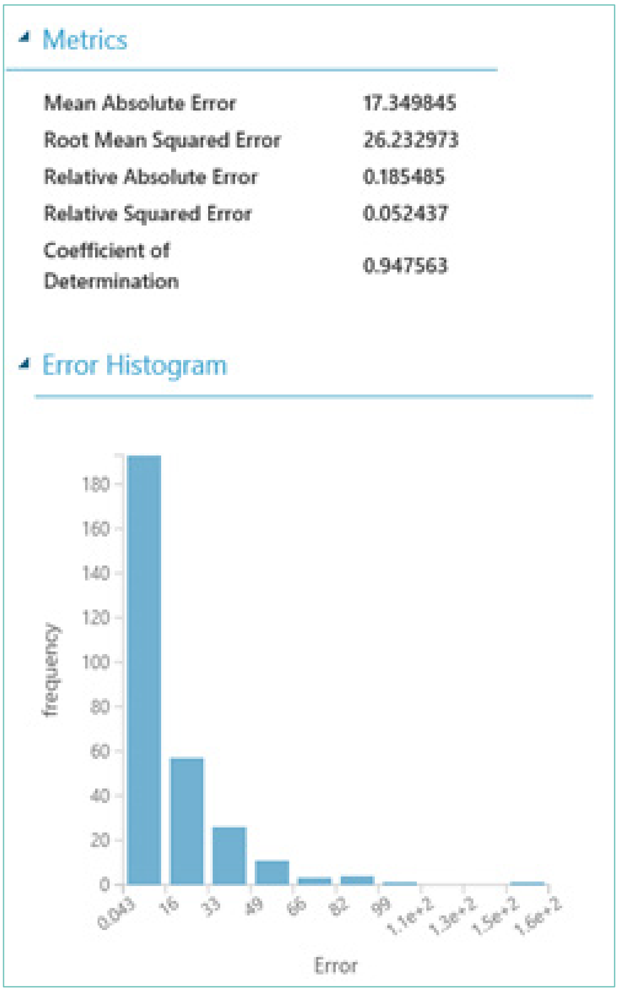

Figure 5: Azure ML DFT model training outcomes

Demand forecast tool accuracy (regional)

Once the regional demand forecast models have been trained, the predictive model is executed as a web service. The web service runs each day when the climate forecast (max. temperature, rainfall) is available from the Bureau of Meteorology (BOM). Other variables such as SMI and evaporation are also calculated from the forecast and passed to the predictive model. The model calculates a 7-day demand forecast. Table 1 shows examples of the model results for the Adelaide and Myponga forecast models, for both Day 1 and Day 7 forecasts compared to the actual demand.

Table 1: Example forecast accuracy of DFT

Adelaide	**Day 1: 24 Jan**	**Day 7: 30 Jan**
Forecast	600.5 ML/d	429 ML/d*
Hindcast	630.1 ML/d	393 ML/d
Actual	618.0 ML/d	423 ML/d
Forecast diff.	17.5 ML/d	6 ML/d
% difference	3 %	1.4 %
Myponga	**Day 1: 24 Jan**	**Day 7: 30 Jan**
Forecast	15.3 ML/d	14.4 ML/d*
Hindcast	14.9 ML/d	9.8 ML/d**
Actual	17.7 ML/d	10.1 ML/d
Forecast diff.	2.4 ML/d	4.3 ML/d
% difference	13 %	40%

*1 Day Forecast on the 30th Jan = 435 ML/d & 14.1 ML/d,

**Change in weather even on one day forecast as demonstrated by the hindcast value

Figure 6 illustrates the Adelaide metro forecast over December/January compared to actual demand with percentage differences for both the one day forecast and hindcast values against actuals. On average over this period the percentage difference is 7 percent. Average demand over this period is 564 ML/d.

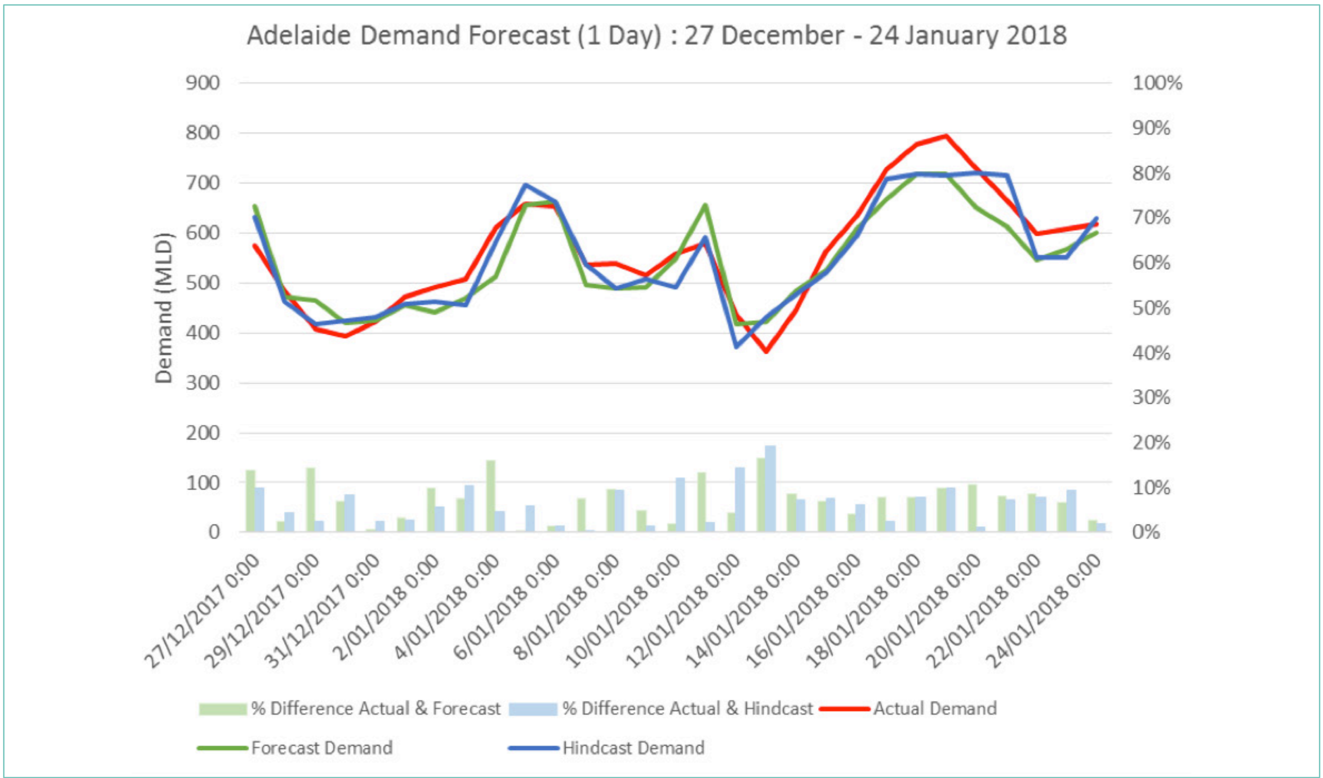

Figure 6: Demand Forecast Tool accuracy for Adelaide (December/January 2018)

Customer categories and diurnal curves

In the existing DFT, the same diurnal curves were used for a category throughout the year, regardless of season. Limited changes had been made by SA Water since the original implementation of the tool to incorporate seasonal diurnal curves or review the curves configured in the model (which were based on model calibration data from 2012).

Figure 7 shows a comparison of the annual average UoA smart meter data and the residential diurnal curves used in the DFT before this enhancement project (RES – standard residential, RES EL051 - residential for the EL51 Pressure Zone). The main differences were as follows:

The overnight usage (12am to 6am) is significantly lower for the UoA smart meter data than the current curves (both RES and RES EL051). The demand estimated by the RES curve may be more than double the residential demand during overnight period.

The UoA Smart Meter Data has higher usage for most of the day, particular at the morning and evening peaks. The RES curve underestimate the demand from the UoA smart meter data curve by up to 50% at 7am and 25% at 8pm.

The RES curve begins the morning peak 1 hour later and does not account for the longer duration of the evening peak. An error in the timing of the peaks will cause incorrect timing of supply.

The RES EL051 curve better matches the UoA smart meter data then the RES curve for the evening peak, but is worse than the RES curve for the morning peak.

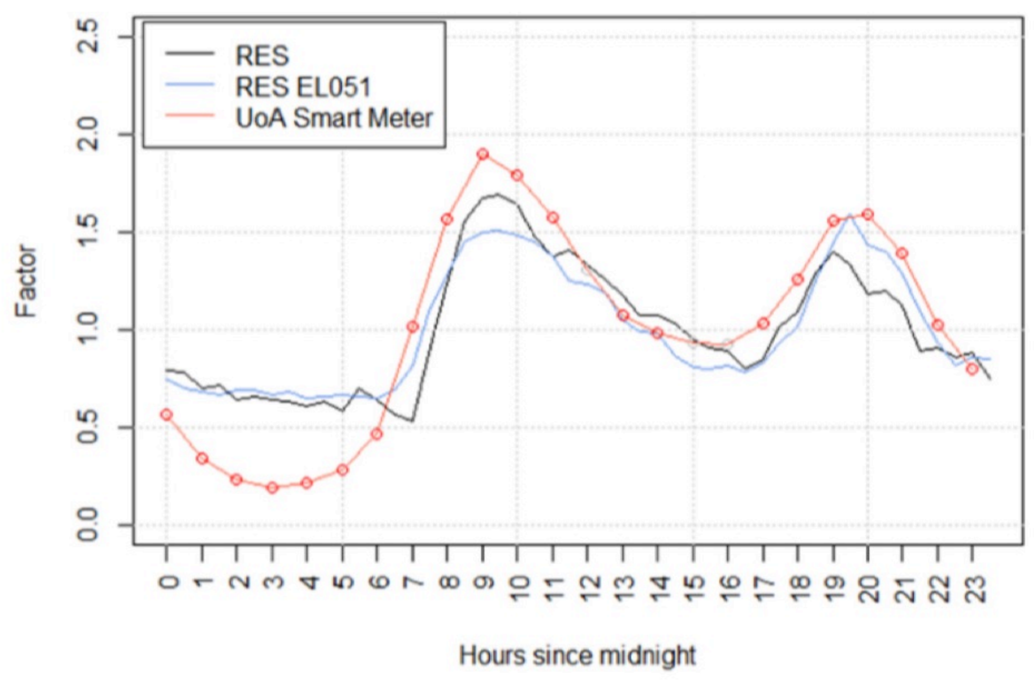

Figure 7: Residential diurnal curve comparison

The lower night time/early morning usage seen in the smart meter data compared to the diurnal patterns used in the DFT is likely because the current patterns used in the DFT account for some network losses. Whereas the smart meter data is directly from the household connection. The original DFT patterns were developed using typical design curves per customer category and also utilising network flow data from a calibration day, where network demand total includes non-revenue water.

A separate 'losses' category is included as part of the DFT and the live operational hydraulic model, so any demand attributed to loss (and pattern over time) is already accounted for. Residential diurnal patterns closer to the shape seen from the household smart meters (lower night-time and higher peaks) was therefore a recommendation of the project team. The effect of this recommendation is seen in the next section of this paper.

The residential end-use data available from UoA was also used by the project team to identify the differences in seasonal consumption patterns for a typical residential household in Adelaide. Figure 8 highlights the different seasonal patterns (for Summer/Winter) for different household incomes.

As seen in Figure 8, UoA also used the household water use data to investigate trends for different household types using demographics. These included low, medium and high-income households. The results showed significant differences in when households use water (i.e. peak periods) and how this changes seasonally. This data could also be applied in the DFT to apply high or low-income consumption patterns to particular customer categories and/or areas. SA Water is working to implement some of these recommendations into the DFT and NOM.

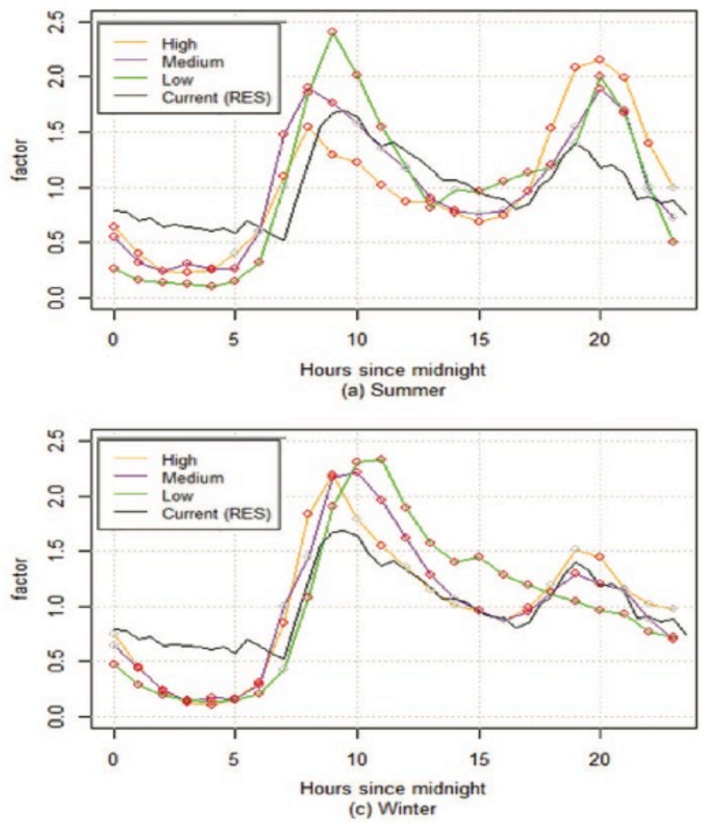

Figure 8: Summer/Winter diurnal curves

Demand forecast tool validation

As part of the project, 'Accuracy Parameters' were developed to track the performance of the demand forecast for each region and also the water treatment plant forecasts. These accuracy parameters calculated mean average error and mean average percentage error. These parameters are calculated daily and on an average 30 day rolling period. This data helps the business understand the accuracy of the tool in real terms and on an ongoing basis (i.e. difference in megalitres).

Network flow data has been configured as verification data feeds to the live hydraulic model, also allowing for ongoing tracking of performance of the hydraulic model, but also the demand forecast data fed into the model. Other live data used as verification feeds to the hydraulic model include storage levels and pump/valve status (i.e. running/closed).

Offline from the live hydraulic model and daily demand forecasting validation, diurnal curves developed from smart meter data by UoA were tested in the hydraulic model to measure improvements and identify changes that should be made to some of the demand data assumptions.

Figure 9 shows mass balance flows to the large EL51 Pressure Zone over 24-hours for a typical autumn day. Data produced from the hydraulic model and from network flow meters is shown. The majority of customers in this zone are residential.

Configuring the hydraulic model with the demand forecast for this autumn day and the original residential diurnal curves, flow to the EL51 zone did not align well with data from the network flow meters. The diurnal curves developed from the UoA's smart meter study were also tested and this resulted in a better match with network flow meter data. The magnitude and timing of peak hour also showed better alignment.

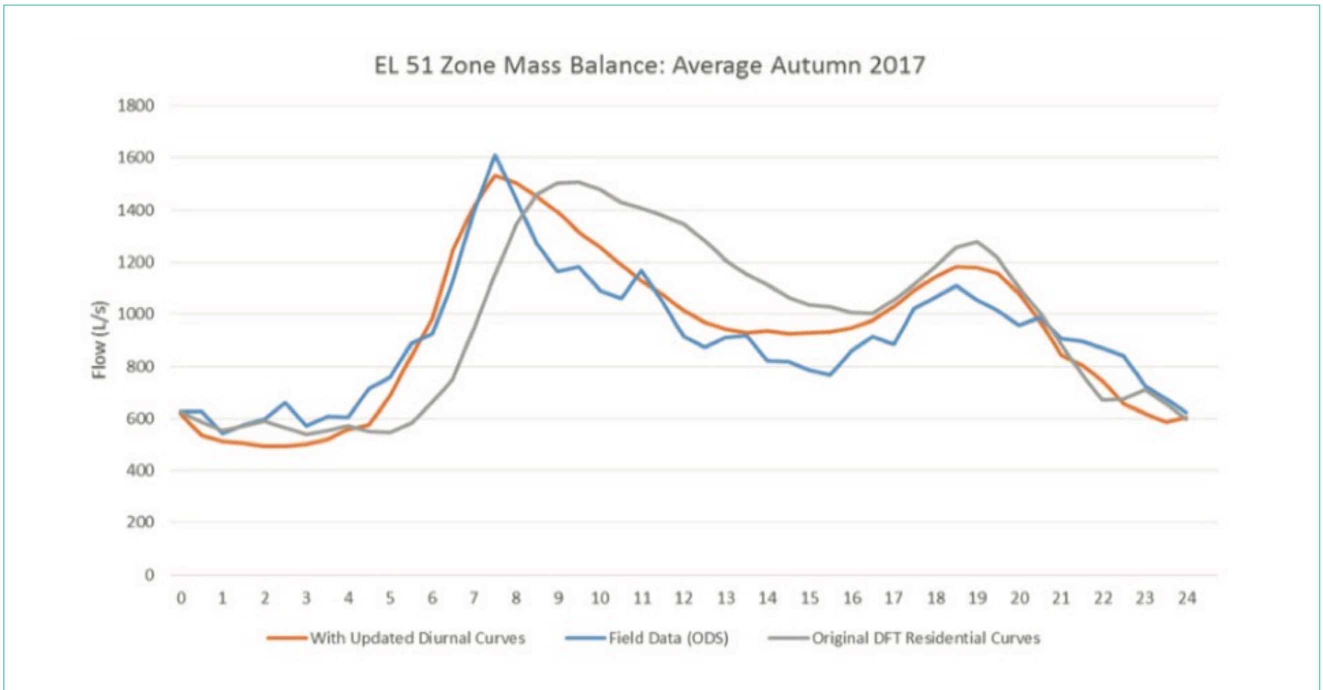

Figure 9: Improved forecast (DFT and NOM)

Some minor adjustments were made to the residential diurnal curve produced from the UoA data sets along with a review of the 'losses' pattern applied to the model. An updated residential diurnal pattern for the EL51 was developed that followed the trends UoA developed with diurnal curves from smart meter data, and also provided a good fit with field data over a range of days.

The project team provided updated diurnal curves for SA Water to review and implement in the demand forecast tool and network operations model. The demand forecast tool and the network operations model have the flexibility for users to review, update and test different data sets such as these residential diurnal curves either offline or committed into the live models.

This flexibility to change input data gives the business the flexibility to change data assumptions in the demand forecast tool if the models (DFT/NOM) are no longer tracking well against field data, if new monitoring sites are added (i.e. more major customers) or if further analysis is carried out for different days/seasons/scenarios, etc.

BENEFITS

The following list outlines some of the benefits already observed from the expanded and enhanced DFT:

- Improved incident response and scheduled outage response planning
- Support of operational planning during critical high demand periods
- Forecasting use of desalination plant for critical supply shortages and high power prices
- Improved data security and integration of operations planning tools
- Consistent data used for forecasting, modelling and operations planning
- Improved water routing and supply capability with reduced costs
- Improved water supply resilience and outage response capability
- Adoption of latest available technology
- Collaborative, agile approach, flexible to new innovative ideas
- Returned business intelligence to core systems
- Enhanced tools already used daily to support operational decision-making
- Providing a machine learning use case for further initiatives within SA Water

A robust and reliable demand forecast tool underpins operational decision making at SA Water and is used as an input into other decision support tools used by the team. Overall the decision support tools lead to real world benefits including:

- Reduction in the number of repair or shutdown issues that escalate to significant or major events
- Reduction in the overall risk of major water supply shutdown events
- Reduction in the number of customers impacted by events
- Reduction in the duration that customers are impacted by events
- Improved understanding of available time for repair before customers become impacted
- Improved speed and accuracy of decision making for planned outages
- Improved speed and accuracy of decision making for unplanned event response
- Improved speed and accuracy of root cause analysis
- Reduction in frequency and duration of general network 'problem solving' investigations

CONCLUSION

Expansion and enhancements of the DFT were a key objective of the project undertaken for SA Water. The project achieved more accurate metropolitan water demand forecasts using a consistent methodology, including all WTP and supply zones, and increased granularity on how the demand forecast is broken down per supply area. Additional accuracy of parameters to track performance, improved data inputs and parameters based on long-term field data were all delivered as part of this project. This resulted in improved performance and forecast accuracy of the decision support tools, particularly the DFT and NOM.

ACKNOWLEDGEMENT

The authors would like to thank SA Water for its support and guidance throughout the project and all the Stantec, SA Water, University of Adelaide and Bentley staff involved in the delivery, testing and final go-live implementation of the expansion and enhancements to the Decision Support Tools for SA Water.

REFERENCES

Arbon, N., Thyer, M., McDonald, D.H., Beverley, K. and Lambert M. (2014), Understanding and predicting household water use for Adelaide Goyder Institute for Water Research. Goyder Institute for Water Research, Technical Report

MWH. (2011). Demand Forecast Tool Functional Design Report

THE AUTHOR

Alana Duncker

Alana has over 16 years' experience in the water industry, she is highly skilled in computer modelling, design and optimisation of water distribution networks for planning and operations, as well as project technical leadership, project and team management.

Working at Stantec these last 6 years, Alana has had a focus on analytics, live hydraulic modelling, operations optimisation and the development of decision support tools for utilities. Alana has been involved since 2012 in the development and delivery of a series of Decision Support tools for the SA Water Operations group, including a live hydraulic model for the Adelaide metropolitan network (Network Operations Mode), a short term 7 day demand forecast tool (Demand Forecast Tool), a bulk water optimisation tool (Distribution Optimisation Tool) and an analytics and dashboarding tool (Network Status Display). Alana continues to work with SA Water to expand and enhance these award winning tools.