THE INFLUENCE OF RESIDENTIAL WATER TANKS ON WATER CONSUMPTION: A CASE STUDY FROM PALMAS, TOCANTINS, BRAZIL

DANIELA BONAZZI SODEK, JOSÉ GILBERTO DALFRE FILHO & ANDRÉ LUIŚ SOTERO SALUSTIANO MARTIM
Department of Water Resources, School of Civil Engineering, Architecture and Urban Design, State University of Campinas, Brazil

ABSTRACT
Water losses in a water distribution system refers to the volume of water produced but not consumed (real loss), or consumed but not charged (apparent loss). The International Water Association (IWA) has a method to quantify and identity water losses through the water balance. However, quantifying the apparent loss to compose the water balance can be a challenge. Although installed flowmeters are essential to monitor consumption, they are not enough to estimate apparent loss. For instance, if the meter is operating out of the correct range, it represents an apparent loss that is difficult to evaluate. Among other causes, in residences where private water tanks are used to store water, the use of float valves can cause higher under-measurements. This indirect water feed is the most common case in Brazilian residences and led to this study, based on the analysis of two years consumption data from a district meter area of Palmas, Tocantins, Brazil. The database presents consumption information for 183 directly fed properties and 102 indirectly fed properties, flowmeter operating range and the size of storage, if present. Consumption patterns according to the presence and size of storage were evaluated and showed that the larger the storage the higher the consumption. The data was also analysed statistically by the two-sample t-test and showed that customers without a water tank consumed, on average, 13% more than those with a water tank. When replacing flowmeters, water companies should prioritize customers with water tanks due to the largest investment return. By understanding the available consumption data, water companies can estimate and possibly reduce apparent water loss, by adopting the most adequate measures.

Keywords: apparent loss, water consumption, flowmeter under-registration, private water tank.

1 INTRODUCTION
Flowmeters are instruments used to measure flow per unit of time. They are essential in water distribution systems to measure the flow produced, distributed and consumed. Information provided by the flowmeters enables water companies to charge customers for the volume consumed and planning for future investments. It is a valuable way to provide information to aid on water loss programmes and consumption reduction [1]. In Brazil, most of the residential connections have flowmeters, in 2016, about 92% of the connections in the country were metered [2]. However, having a flowmeter installed does not guarantee data quality as there are many factors that impact on metering accuracy.

Water loss is known to be divided into real loss, when it is produced but not consumed, or apparent loss, when it is consumed but not charged. Identifying water loss can be a very complex process as it is influenced by many factors [3]. The International Water Association (IWA) has a well-established method to evaluate water loss which starts with a water balance based on data collected from water meters. As a result, a summary of consumed volumes and water loss is presented and the volume produced can be compared with the consumed or lost volumes [1]. The quality of the water balance depends on the quality of the data available, however the purpose of it is not only quantify the total lost volume, but also to identify its provenience [4]. Apparent loss refers to data acquisition errors, metering errors and
unauthorised consumption and it can represent a financial loss sometimes greater than
the real loss [5]. Metering errors and unauthorised consumption are the main sources of
apparent losses.

When it comes to metering errors Jhonson and Vermersch [6] lists aging, inappropriate
meter installation and impact of customers in-house installation as the main reasons for
apparent losses. For each apparent loss cause, the authors suggest a range of solutions; for
instance, regarding aging, the suggestion is to have an aging and usage curve and a meter
replacement policy, for inappropriate installation, they suggest the definition of standard
procedure and drawings, training and replacement of meters installed incorrectly; for the
impact of customer’s in-house installation, they suggest laboratory tests and correct selection
and sizing of meters according to the consumption profile. Flowmeter under-registration is
only one factor that can contribute to increase apparent loss and impacts not only the revenue,
but also the planning for water distribution system expansion [1]. The under-registration is
hard to be determined or precisely measured, so, often a default value is assumed.

Regarding aging it is responsibility of the operating company to verify if the flowmeter
function well. The meters should be taken for analysis every five years [7]. Although
flowmeter performance is expected to decay with age, Du Plessis and Hoffman [8] present
the results of lab tests performed on some flowmeters from South Africa where 71% of the
meters were within the 2% accuracy for a new class C meter, in spite of 80% of them having
more than 10 years of use. The meters were removed during routine meter replacement
programs and taken to laboratory for testing. Tests were performed on 91 meters aged
between 5 and 35 years, all with a nominal flow of 1.5 m³/h Class C. The pressure on the test
was the municipal recorded pressure and the flow rate varied between 0.9 and
1.8 m³/h. However, it should be noted that performance for lower flows could present
different outcome.

Concerning installation, researches has shown that flowmeters installed with the turbine
axis with a 45° angle against the vertical axis can undermeasure by 8% and it could rise up
to 15% if the meter is installed at a 90° degree angle [9]. Pipeline fittings upstream of
flowmeters may also impact the measurement, for instance, the experiment presented by
Martim et al. [10] showed that electromagnetic flowmeter with an upstream gate valve can
undermeasure about 3%.

Relating to the in-house installation, researches has also shown that in properties with
water tank, flowmeter can undermeasure due to the common use of float valves (or ball valve)
at the inlet of the tanks. Float valves are often used in residential water tank to control water
levels. These valves open and close slowly allowing a flow smaller than the minimum flow
carried out a case study in Malta aiming on to quantify these under-registered volumes by
monitoring the inlet and outlet of water tanks. The research had class D new meters with data
logger installed on the inlet and outlet of the roof tanks of three houses. The results showed
that the inlet flowmeter under-registered up to 9% of the flow measured by the outlet meter.
The test was also performed with a 5-year old in-situ flowmeter on the inlet and a new meter
on the water tank outlet and found that around 92% of outflows from the water tank were not
measured by the inlet meter. The authors have also performed the same test replacing the
float valve with a solenoid valve to control flows into the water tank and the results showed
that there was no under-registration. The authors recommend the use of solenoid valve as a
solution for the under-registration. Pereira [12] has performed a similar experiment in the
city of Campinas, Brazil. The author compares the performance of a conventional float valve
with the performance of a high flow float valve. Under-registration was also detected when
using conventional float valve. Despite the use of high flow float valve be a possible solution,
the intervention of the final user could compromise the results. The use class C flowmeters was recommended on the study.

International standards for cold potable water flowmeters used to classify the meters in metrological classes. Although some of the Brazilian current standards still refers to metrological classes, for example [13]–[15], a newer standard NBR16043 [16] (based on ISO 4064-1 2005) is also in place and currently being updated to comply with ISO4064-1 2014 [17], regulation No. 295 [18] by the National Metrological Institute (INMETRO) is also in place for the transitional period. The Brazilian standard presents three metrological classes A, B or C while the international (ISO4064-1 1993) had a further class D. The new international standard presents two accuracy classes, type 1 and 2, for each class a maximum permitted error (MPE) is specified for the upper and lower flow rate zones [17]. However, most of the articles referenced in this work still uses the old metrological classes, being class A the less accurate and class D the most accurate. Basically, the more accurate the meter, the smaller will be the start up flow it can measure. Meters class B are commonly used for domestic purposes in Brazil. Independently of the feed (direct or indirect), the magnitude of the flow rate consumed relates to the error curve of the meter, to accurately measure the consumption. As shown on the typical error curve in Fig. 1, between flow start-up and minimum flow (Qmin or Q1) there is no consumption been registered, between Qmin and the transition flow (Q1 or Q2) it will register with up to ±5% error, between Q2 and the nominal flow (or permanent) (Qn or Q3) the flow meter will measure with up to ±2% error. In residences with float valves at the entrance of water tanks, the flow will be below minimum more frequently which means under-registration will be greater.

![Figure 1: Typical flowmeter error curve (based on NBR16043 [16]).](image)

One possible solution to reduce under-registration is presented by Fantozzi [19]. The author presents a case study carried in Maddalena, Italy, where 33 meters class C aged between 1 and 7 years were tested with and without an Unmeasured Flow Reducer (UFR). UFR are simple equipment installed on the water main in-line with water meter. It regulates the water flow to only let flow past in batches when it is high enough to be measured by the water meter. According to the author for this case study the use of UFR reduced 94% of unmeasured flows below start-up flow, 31% of unmeasured flows at start-up flowrate and
14% of unmeasured flows at transition flow. In residences where a private water tank is coupled with an old flowmeter, it may not register more than 50% of the consumed flow [20]. Criminisi et al. [20] presents a case study undertaken in Palermo, Italy, to identify possible solutions to under-registration in residences with private water tanks. The under registration is considered even more relevant as the use of float valve will make a relevant part of the incoming flow to be lower than the start-up rate. The study concluded that the use of UFR could reduce under-registration by about 9% when the old meters where kept. When combining the installation of UFR with the replacement of old meters by new class C turbine meters the under-registration dropped to almost zero.

The best solution to reduce under-registration where private water tanks seems to be a combination of UFR installation and flowmeter replacement. Fontanazza et al. [21] presents a case study showing the effect of flowmeter age and private water tank on apparent losses. The study was performed in a small district meter area (DMA) of Palermo (Italy) where properties have private water tanks. Consumption on this DMA was monitored before and after UFR installation, this data was used to verify the hydraulic model produced with EPANET. The model was used to investigate apparent losses evaluating four different scenarios: do nothing, common flowmeter substitution, installation of UFR and a combination of changing flowmeters and installing UFR. The last scenario, proved to be the best solution. According to the authors the analyses also shown that properties with water tanks that were full most of the time and with low network pressure would benefit the most form UFR installation as the flowmeters would be often operating with low incoming flows. According to the authors, the use of a model can be an efficient method to identify areas with high apparent loss.

In Brazil, most of the residences have private water tank to store water in case of a water shortage and the used of class B flowmeters is common. Researches has shown that this combination is prone to have biggest under-registration, increasing apparent losses. The objective of this case study is to identify a pattern of consumption of residences with private water tank using information from two years consumption database of a sector form Palmas city in Brazil. As a result, an estimated percentage will be determined to aid the water company to better estimate under-registration for customer with private water tank, as well as helping planning the meter replacing programme.

2 METHODOLOGY

The city of Palmas is the capital of Tocantins Estate in Brazil, it was founded in 1989, aiming at being a centre of social economic development for the Brazilian North region. Its geographical location is presented in Fig. 2. With a territorial area of 2,218,942 km² and an estimated population of 291,855 people [22], Palmas per capita water consumption is of 150 L/inhab./day. In 2016, the average monthly salary was about $US 950,00 dollars. The case study sector has a complete consumption data base for a period of 2 years (2016 and 2017). It is a newly populated area, horizontal and mainly residential, with similar socioeconomic standard. The data provided by the water company presents the costumer category (residential, commercial or industrial), installed flowmeter brand, flowmeter nominal flow rate, flowmeter maximum flow rate, volume of residential water tank and monthly consumption per costumer. The data was presented in shapefile format, which can be visualised with Geographical Information System (GIS) software. It holds geometric location and also contained database tables with the consumption information. This database format has facilitated sharing the data between software as Microsoft Excel and Minitab® Statistical Software, used in this study.
For the purpose of evaluating the effect of the storage size in consumption, only residential customers with data regarding storage size, installed flowmeter brand, nominal and maximum flow and consumption data for all of the 24 months were considered. It was not the intention of this work to verify the data accuracy, but instead, to use the data as presented. Meter accuracy from the readings were also accepted to be the same throughout the period and among properties. Customers with “0 litres” on the database were understood to have no water tank and therefore deemed to be directly feed from the water distribution system. Customers with information regarding the tank size registered only with “litres” or “0” were considered uncertain and disregarded in the study. Those listed with the tank volume different than zero, have a water tank and were considered to have an indirect feed. The database has in total 757 connections, being 393 direct, 266 indirect fed and 98 uncertain. Data screening was performed using Microsoft Excel.

After screening, the data was transferred to Minitab® Statistical Software. Two analyses are proposed, the first being the relationship of average consumption and the water tank size and the second is to determine if the average consumption of residences without a water tank (direct feed) is greater than the average consumption of the residences with water tanks (indirect feed). Therefore, a detailed distribution analysis was performed for these two consumption groups, followed by the statistical two-sample t-test.

In descriptive statistic, there are several visual ways to describe the distribution of a certain set of data. Visual methods help to understand the sample data and to make comparison between them. Box plots are a simple, but very effective tool and it has been frequently used and widely recommended as the best visual method to present the data distribution [23]. The box plot is characterized by a box with a line inside and two whiskers. This work presents results with boxplots Tukey-style whiskers.

Figure 2: Location of Palmas in Brazil (10°11’4”S, 48°20’1”W).
Once the data is organized from the highest to the lowest value, the 25th, 50th and 75th percentiles can be determined. The lower line of the box presents the first quartile or lower quartile (Q₁) and is the 25th percentile of the sample. The line in the middle is the median, equivalent to the 50th percentile of the sample, also known as the second quartile (Q₂) and, finally, the upper line of the box is third or upper quartile (Q₃) equivalent to the 75th percentile. The interquartile rage equals to the difference between the first and third quartile (IQR = Q₃−Q₁). The whiskers are the lines that extend to the data that is no more than 1.5 x IQR from the edge of the box (Tukey style) [24], [25]. Any data outside the limits of the whiskers is considered an outlier and is individually plotted. A sample containing too many outliers may indicate that it does not follow a normal distribution. An outlier is an observation distant from the rest of the data and, therefore, treated as abnormal data, that can affect the overall observation due to its very high or very low values. Once outliers have been identified, a decision can be made whether or not to keep them on the sample [24]. In this work, the box plot was performed using the statistical software Minitab for the two consumption groups, direct and indirect feed. The outliers identified were removed before performing the hypotheses test.

A hypothesis test is used when the objective of a statistical analyses is to evaluate two exclusive statements about a population, to determine which is best supported by the sample data [24]. The two statements are called null hypothesis (H₀), and alternative hypothesis (Hₐ). The null hypothesis will be discarded only if the sample evidences suggest that H₀ is false. In this case Hₐ is true. Clearly, this case study contains two independent groups (direct or indirect feed). The null hypothesis is that there is no difference between the average consumption of the direct feed customer and the average consumption of the indirect feed customer, represented on eqn (1). The applied test was the one-tailed test, with the alternative hypothesis considering that the direct feed population mean is greater than the indirect feed population mean, represented in eqn (2):

\[ H₀: \mu₁ - \mu₂ = \Delta₀ , \quad (1) \]
\[ Hₐ: \mu₁ - \mu₂ > \Delta₀ , \quad (2) \]

where \( \mu₁ \) is mean consumption of direct feed customers, \( \mu₂ \) is the mean consumption of indirect feed customers and \( \Delta₀ \) is the assumed value for the difference, in this case study \( \Delta₀ = 0 \). A significance level (\( \alpha \)) needs to be chosen to make the decision. In this case, 0.05 (\( \alpha=5\% \)) was chosen. The significance level indicates that there is a 5% chance of concluding that there is a difference between groups, when there is actually no difference. To apply the hypothesis test, the two-sample t-test can be calculated using eqn (3). This equation is comparing the sample means to the null hypothesis and incorporating the sample size and variability in the data.

\[ t = \frac{\bar{x} - \bar{y} - \Delta₀}{\sqrt{\frac{s₁²}{m} + \frac{s₂²}{n}}} , \quad (3) \]

where \( \bar{x} \) is the average direct feed consumption, \( \bar{y} \) is the average indirect feed consumption, \( \Delta₀ \) is the value chosen for the null hypotheses (in this case zero), \( s₁ \) is the direct feed sample mean deviation \( m \) the sample size, \( s₂ \) is the indirect feed sample mean deviation and \( n \) the sample size.

Once the t-value is determined the P-value can be determined based on significance level, t-values and sample size. The P-value can be seen as the limit of significance level. P-values smaller or equal to the chosen significance level (\( \alpha=0.05 \)) means the null hypotheses can be rejected for this significance level. P-values greater than the chosen significance level
(α=0.05) means it is not possible to reject the null hypothesis for the chosen significance level. The two-sample t-test was performed using Minitab and the P-value is calculated by the software, however P-value can also be read from tables.

3 RESULTS AND DISCUSSION

The data screening was performed, as described, with Microsoft Excel. The result, is the sample data for the statistical study. The sample after the screening had 183 direct feed properties and 102 indirect feed properties. This sample was then transferred to Minitab Statistical Software, to be analysed. Two analyses were performed with the data, the first looked at the relationship between storage volume and average consumption and the second at the difference between average consumption of customer with and without water tank.

For the first analysis the data was divided in groups by water tank volume the average consumption and deviation were determined. The sample presented only five customers with 750 litres water tank, being that sample too small to be considered in this analysis, however these customers were considered on the total of properties with water tank for the second analysis. As presented on Table 1, the sample contains 49 residences with 500 litres water tank, 32 residences with 1,000 litres water tanks, 8 residences with 400 litres water tank and 8 residences with 250 litres water tank.

Table 1: Statistical result for the analysis of properties versus water tank capacity.

<table>
<thead>
<tr>
<th>Water Tank Capacity</th>
<th>Sample size (N)</th>
<th>Average (m³)</th>
<th>Deviation (m³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000 litres</td>
<td>32</td>
<td>11.79</td>
<td>3.40</td>
</tr>
<tr>
<td>500 litres</td>
<td>49</td>
<td>11.65</td>
<td>4.97</td>
</tr>
<tr>
<td>400 litres</td>
<td>8</td>
<td>10.64</td>
<td>3.35</td>
</tr>
<tr>
<td>250 litres</td>
<td>8</td>
<td>8.59</td>
<td>5.79</td>
</tr>
</tbody>
</table>

When averages are compared, the properties with bigger storages presented the highest average consumption, as shown on Table 1 and Fig. 3. The average consumption of customers with 1,000 litres water tanks was 11.79 m³, and for customers with 500 litres was only 1.2% smaller (11.65 m³). However, properties with 400 litres and 250 litres presented average consumption 9.8% (10.64 m³) and 27.1% (8.60 m³) smaller than customers with 1,000 litres respectively.

![Average consumption per water tank volume](image)

Figure 3: Average consumption for 24 months according to water tank size.
The result shows that there is a tendency for customers with bigger water tanks to have greater average water consumption. One way to see this tendency is, those with a larger water tanks are, probable, properties with characteristics that will demand more water, for instance, the number of shower rooms, size of the garden, presence of swimming pool, the size of the house and the number of inhabitants. Because of that those properties will look for a storage that will keep them going in an event of a draught. In this sense, it would be expected for customer without water tanks to have even smaller average consumption, however the results of the second analysis show otherwise.

The second analysis was performed to identify if there is a difference between the average consumption of customer indirectly fed (with water tank) and directly fed (without water tank). The data was then separated in two groups and the box plots prepared with Minitab® Statistical Software to analyse the data distribution for each group. Fig. 4 presents the result of that analysis. The sample of directly fed customers presented median of 13 m$^3$, while the sample of indirectly fed presented that of 10.6 m$^3$. Regarding outliers, the directly fed sample has eight outliers and the indirectly fed sample has only one. Outliers represent customers with higher consumption rates than most customer from the sample and were not considered on the two-sample t-test.

![Boxplot of Direct and Indirect consumption](image)

**Figure 4:** Sample distribution analyses with boxplot for direct and indirect feed customers.

Having outliers shows that the distribution is not normal, however as both samples are bigger than 15, normality is not a problem and the two-sample t-test is exact [24], [25].

The final sample used to perform the statistical test has 175 directly fed properties and 101 indirectly fed properties. The descriptive statistic was performed and its results are shown
The average volume of the directly fed sample is of 12.85 m$^3$ and 4.02 m$^3$ deviation, while for the indirectly fed sample is of 11.18 m$^3$, 4.19 m$^3$ deviation.

Table 2: Results of descriptive statistics for the two samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>N</th>
<th>Mean</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>175</td>
<td>12.85</td>
<td>4.02</td>
</tr>
<tr>
<td>Indirect</td>
<td>101</td>
<td>11.18</td>
<td>4.19</td>
</tr>
</tbody>
</table>

The result of the descriptive statistics shows that the average consumption of properties without water tanks is in average 1.67 m$^3$ greater than those with a water tank. There is 99% confidence that the real difference is bigger than 0.458 m$^3$. However, a hypothesis test is needed to confirm if the difference is significant for the considered samples. This difference can be due to flowmeter under registration in properties with float valves on the inlet of water tank, as the inlet flows can be very low.

For the two-sample t-test, the null hypothesis is the difference between the mean consumption of the directly fed customer and the mean consumption of the indirectly fed customer equal to zero (H$_0$: $\mu_1-\mu_2=0$), meaning there is no difference in the average consumption of properties directly and indirectly fed. The alternative hypothesis is that the difference between the mean consumption of directly fed customers and the mean consumption of indirectly fed customers is greater than zero (H$_1$: $\mu_1-\mu_2>0$), meaning the average consumption of customers directly fed is greater than the average consumption of customers indirectly fed.

The results of the two-sample t-test shows a calculated t-value for the samples of 3.23 and degree of freedom (DF) of 201 which gives a P-value of 0.001. As P-value is smaller than the significance level of 0.05, it can be concluded that the null hypotheses can be discarded and the alternative hypotheses can be accepted. That means that the difference between the average consumption of the directly fed properties and the indirectly fed properties is significant at both levels 95% and 99%. According to the tests, there is 98% confidence that the real difference is between 0.458 m$^3$ and 2.88 m$^3$.

In statistics, the power function is the probability of concluding that there is a real difference when this difference actually exists. The power function depends on the sample size and deviation, one practical way to improve the power function is to increase the sample size [25]. Ideally the power function should be higher than 90%. For the samples provided on this study the t-test has detect a difference of 1.67 m$^3$ and the power for the observed difference is 74.4%. Although the power is lower than 90%, according to Minitab® Statistical Software [25] no concern is needed as the test has detected a difference between the samples.

4 CONCLUSIONS
The rapid population growth experienced by the cities on the last decades has put into perspective many aspects of sustainability, among others, access to clean water and the need of an efficient distribution system. Understanding water loss and how it can be reduced for each system poses as a challenge as well as duties for water companies. In the case study presented, the average of two years consumption of customer with and without private water tank were compared statistically using the two-sample t-test. The statistical test showed a significant difference between the average consumption of customer with direct and indirect feed. The difference of 1.67 m$^3$ was observed, meaning that customers with private water tank for this study measured in average 13% less consumption. This result can be used as an estimative of under-registration when preparing a water balance. The average consumption
of properties according to the water tank volume was also analysed, concluding that properties with bigger water tanks do consume more water.

Furthermore, when replacing flowmeters water companies should prioritise customer with water tanks as it will give the largest investment return, especially if replacing old class B with more accurate meters. Another alternative could be the installation of UFR to further reduce under-registration. The use of solenoid valves or high flow float valve could also be considered, however as it an in-house installation it cannot be installed by the water companies, and its maintenance would be the customers responsibility.

ACKNOWLEDGEMENTS
Authors would like to thanks CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for the scholarship and Palmas water company, BRK Ambiental, for providing the valuable data that enabled this work to be undertaken.

Portions of information contained in this paper are printed with permission of Minitab Inc. All such material remains the exclusive property and copyright of Minitab Inc. All rights reserved.

REFERENCES


