

Análise de dados de telemetria de consumos domésticos para apoio na avaliação de perdas aparentes e reais em sistemas de distribuição de água

Analysis of household telemetry consumption data for helping in the assessment of apparent and real losses in water distribution systems

Silva, M.A.^{a*}, Loureiro, D.^b, Amado, C.^c, Arregui, F.^d

^a CEMAT, Instituto Superior Técnico, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal e Laboratório Nacional de Engenharia Civil, Avenida do Brasil 101, 1700-066 Lisboa, Portugal. E-mail: maria.jose.silva@tecnico.ulisboa.pt

^b Laboratório Nacional de Engenharia Civil, Avenida do Brasil 101, 1700-066 Lisboa, Portugal. E-mail: dloureiro@lnec.pt,

^c CEMAT, Instituto Superior Técnico, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal. E-mail: conceicao.amado@tecnico.ulisboa.pt,

^d ITA, Universitat Politècnica de València, Edificio 5C 1a planta Campus de Vera, 46022 Valencia, Espanha E-mail: farregui@ita.upv.es.

*Autor para correspondência

Recibido: 05/05/2020

Aceptado: 22/12/2020

Publicado: 31/01/2021

Citar como: Silva, M.A., Loureiro, D., Amado, C., Arregui, F. 2021. Analysis of household telemetry consumption data for helping in the assessment of apparent and real losses in water distribution systems. *Ingeniería del agua*, 25(1), 37-47. <https://doi.org/10.4995/la.2021.13641>

RESUMO

O aumento da cobertura de clientes domésticos com telemetria tem levado a uma maior disponibilidade de dados de consumo instantâneo nas Entidades Gestoras (EG) de sistemas de distribuição de água. A análise destes dados, para além de melhorar a gestão dos sistemas de faturação e de gestão de clientes, permite também melhorar a estimativa das perdas de água em cada EG. Com o objetivo de avaliar as perdas aparentes, através da análise de perfis do consumo total, e as perdas reais, por via do estudo do consumo doméstico mínimo, foram realizadas campanhas de telemetria em 275 clientes domésticos referentes a 10 EG. Os perfis de consumo, relevantes para aferição do erro global do parque de contadores, revelaram que os consumos predominantes variam entre 240 e 1000 L/h, sendo o volume medido abaixo de 120 L/h (e.g., devido a fugas domiciliárias) praticamente insignificante. A estimativa do consumo médio de água no período de consumo mínimo (entre as 2 e as 5 h da manhã) foi 1.2 L/(utilizador·h) para contadores DN15 e 2.0 L/(utilizador·h) para contadores DN20. Utilizando análise de *clusters*, identificaram-se ainda quatro padrões de consumo distintos, diferenciados pelas escalas de consumo, tendências e picos ao longo do dia. Este trabalho demonstrou o potencial das campanhas de telemetria para recolha de dados de consumo. Além disso, forneceu uma metodologia e conjunto inicial de estatísticas sobre o consumo doméstico em cada EG e no conjunto das EG determinante para a gestão técnica dos sistemas de distribuição de água.

Palavras chave | consumos domésticos mínimos, padrão de consumo, perfil de consumo, sistemas de distribuição de água, telemetria domiciliária.

ABSTRACT

The availability of instantaneous data on water consumption in water utilities of water distribution systems has been increasing due to the increase in the number of customers with smart meters. The analysis of these data allows to improve the billing systems and the customers' management and to improve the water losses estimate in each water utility. Aiming to evaluate the apparent losses, through the analysis of the water consumption profiles, and real losses, through the study of the minimum night flow, 275 domestic customers of 10 water utilities were measured using a smart meter. The water consumption profiles useful for the global error measurement of the flowmeters park identified the predominant flow between 240 L/h and 1000 L/h, with an insignificant measured volume below 120 L/h (e.g., due to customer leakages). The estimate of the mean water consumption in the minimum consumption

period (between 2 h and 5 h a.m.) was 1.2 L/(customer·h) for DN15 flowmeters and 2.0 L/(customer·h) for DN20 flowmeters. Using clustering, four water consumption patterns were also identified, with differences in the consumption scales, trends and peaks during the day. This work demonstrated the potential of smart metering campaigns to collect consumption data. Moreover, it also proposes a methodology and an initial set of statistics about the domestic consumption in each water utility and in the set of water utilities, which is crucial for the technical management of water distribution systems.

Key words | *minimum night flow, smart metering, water consumption pattern, water consumption profile, water distribution systems.*

INTRODUÇÃO

A disponibilidade de dados instantâneos de consumo de água tem aumentado nas Entidades Gestoras (EG) de sistemas de distribuição de água devido à instalação nos últimos anos de contadores equipados com sistemas de telemedição. A análise destes dados, para além de melhorar a gestão dos sistemas de faturação e de gestão de clientes, permite também melhorar a estimativa das perdas de água em cada EG, nomeadamente através da análise dos perfis e dos padrões de consumo e dos consumos mínimos domésticos (Loureiro et al., 2014a; Loureiro et al., 2014b; Loureiro et al., 2015).

Os perfis de consumo correspondem à distribuição do volume total medido segundo classes de caudal. Este conhecimento, conjugado com resultados de ensaios de contadores para estimativa dos erros de medição, permite estimar o erro global dos contadores, assim como a sua taxa de degradação (Arregui et al., 2006; Arregui et al., 2016; Arregui et al., 2018). Consequentemente, possibilita estimar os erros de medição como subcomponente das perdas aparentes (Vermersch et al., 2016).

No que se refere às perdas reais de água, uma das práticas mais conhecidas para a monitorização de fugas é através da monitorização dos consumos domésticos mínimos, que em geral ocorre durante o período noturno (Farley e Trow, 2003; Alegre et al., 2005). Esta prática pressupõe que os caudais medidos à entrada de um setor de uma rede de distribuição de água nos períodos em que a população se encontra mais inativa correspondam predominantemente às perdas reais. No entanto, é inevitável que ocorram alguns consumos pontuais, como consumos reais (domésticos e não domésticos) ou perdas a jusante dos contadores dos clientes. Assim, é necessário saber estimá-los de modo fiável, para uma melhor avaliação da componente de perdas reais na rede pública (Fantozzi e Lambert, 2012).

Os dados instantâneos de consumo de água permitem ainda a obtenção do padrão diário de consumo de cada cliente, calculado através da mediana em cada instante de tempo. Estes podem ser utilizados para agrupar os clientes de acordo com o seu comportamento de consumo de água, de modo a identificar as horas de maior e menor consumo, assim como clientes com consumos atípicos. Além disso, poderão também permitir identificar as habitações com usos exteriores, por exemplo, para rega de jardins, que correspondem muitas vezes a consumos significativos (Loureiro et al., 2015; Loureiro et al., 2016). Estes consumos significativos condicionam em certos casos o funcionamento hidráulico dos setores de rede, causando por exemplo problemas de pressão mínima. Quando ocorrem durante o período noturno podem interferir na análise dos consumos noturnos para monitorização de fugas. No entanto, em geral, as EG não dispõem nos sistemas de faturação e de gestão de clientes de dados sobre a existência de usos exteriores em cada cliente. Deste modo, a evidência desta característica através da análise das séries temporais de consumo telemedido é também relevante para caracterizar melhor os clientes de cada EG. Consequentemente, as EG poderão melhorar a gestão das suas redes a partir dos dados de telemedição, dado o maior conhecimento que terão das necessidades de água dos clientes. Sendo o número exato de diferentes tipos de comportamento de consumo de água existentes desconhecido, pretende-se que este estudo ajude também a atingir este objetivo.

Deste modo, a análise dos dados de telemedição através dos perfis e dos padrões de consumo e dos consumos mínimos domésticos permite uma melhor gestão das perdas de água, assim como um maior conhecimento sobre o modo como esta é utilizada. Neste artigo serão discutidas metodologias para análise de consumos domésticos telemedidos e apresentadas estatísticas relevantes para a avaliação de perdas aparentes e reais em sistemas de distribuição de água, tendo por base dados obtidos por campanhas de telemedição em 275 clientes domésticos referentes a 10 EG. Estas mesmas metodologias e os respetivos resultados encontram-se descritos nas secções seguintes. Igualmente, as principais conclusões, tal como o trabalho futuro, também se encontram descritos de seguida.

METODOLOGIA

Tendo como base o objetivo de estudar uma amostra de clientes telemedidos de diversas EG, nesta secção será apresentada a metodologia utilizada para o cálculo dos perfis de consumo, dos consumos mínimos domésticos e dos padrões de consumo destes clientes.

Perfis de consumo doméstico

A análise da distribuição do volume medido segundo classes de caudal, usualmente designada por perfil de consumo, foi efetuada com base nas seguintes classes de caudal (L/h): [0, 6], [6, 12], [12, 18], [18, 24], [24, 30], [30, 60], [60, 120], [120, 240], [240, 500], [500, 1000], [1000, 1500], [1500, 2000], [2000, 3000], [3000, +∞]. A definição destas classes teve como objetivo garantir a cobertura tanto dos consumos mais baixos, como dos consumos mais elevados. Para cada classe de caudal, foi contabilizado o volume medido que passou nessa gama. Uma vez que o período de campanha pode variar de cliente para cliente, tendo a campanha uma duração aproximada de uma semana em cada cliente, os perfis de consumo foram normalizados ao dia, permitindo uma comparação direta entre os diferentes clientes.

A diversidade da amostra no que diz respeito ao tipo de alojamento e diâmetro nominal (DN) dos contadores levou a que fosse ainda efetuada uma análise mais detalhada. Deste modo, os perfis de consumo foram analisados por tipo de alojamento (apartamento ou vivenda) e por DN do contador (DN15 ou DN20), para cada EG.

Consumo doméstico mínimo

O período de consumo mínimo, em geral durante o período noturno, pode variar entre EG e também ao longo do ano devido, por exemplo, à rega ou ao enchimento de reservatórios domiciliários ou de piscinas durante a noite. Neste trabalho, a definição deste período foi realizada com base na análise do padrão diário de cada cliente. Este padrão diário caracterizado por 96 passos de tempo com um intervalo constante de 15 minutos, foi obtido através do cálculo da mediana do consumo em cada passo de tempo (menos sensível a eventos anómalos do que a média), ao longo do período de campanha. Por exemplo, o consumo no intervalo de tempo 8:45 - 9:00 foi obtido através do cálculo da mediana do consumo de todos os dias da campanha no mesmo intervalo de tempo. Observando os padrões diários de todos os clientes de cada EG, analisou-se o consumo no período com menor consumo.

Para cada cliente de cada EG, começou-se por selecionar os dados relativos ao período de consumo mínimo para todos os dias da campanha de telemedição e calculou-se a sua média. Ao conjunto das médias de todos os clientes de uma EG, retiraram-se as médias dos 5% de clientes que apresentam médias mais elevadas, de modo a eliminar consumos atípicos, e calculou-se a média dos restantes 95% de clientes. Este valor é estatisticamente designado por média truncada à direita a 5%. Neste caso, considerou-se que esta média corresponde a uma estimativa do consumo noturno doméstico, que inclui o consumo de água para usos domésticos e as perdas domiciliárias. A opção pelo uso da média truncada à direita prende-se com a sua robustez, uma vez que os 5% de clientes com médias mais elevadas e que, portanto, não representam a generalidade, são excluídos do cálculo da média. Assim, pode-se considerar que o valor obtido é uma estimativa robusta do consumo de água para usos domésticos e perdas domiciliárias.

De modo a obter uma estimativa apenas para as perdas reais domiciliárias, para cada utilizador de cada EG, calculou-se ainda o mínimo dos dados relativos ao período de consumo mínimo de todos os dias da campanha de telemedição. De seguida, foi calculada a média truncada à direita a 5% do conjunto dos mínimos dos utilizadores de uma EG. O valor obtido foi considerado como uma estimativa das fugas domiciliárias.

Padrões de consumo doméstico

A análise dos padrões de consumo foi efetuada utilizando *clustering* que pressupõe, em geral, que a dimensão das séries temporais de consumo de água seja igual para todos os clientes. Uma vez que tal pode não se verificar, foi necessário encontrar um intervalo comum a todos os clientes, tendo-se optado pelo dia. Por conseguinte, para esta análise, foi utilizado o padrão diário de cada utilizador, calculado usando a mediana do consumo em cada passo de tempo, tal como descrito anteriormente.

A análise dos padrões de consumo começou por uma análise de *clustering*, de modo a identificar os tipos de padrões existentes na base de dados. O primeiro passo da análise de *clustering* passa pela escolha da distância a utilizar para medir a dissimilaridade entre os objetos. De seguida, é necessário escolher o algoritmo de *clustering* a utilizar e avaliar a partição obtida, por exemplo, através da utilização de índices construídos para o efeito. Neste caso, optou-se por utilizar a distância Dynamic Time Warping (DTW), uma vez que esta distância compara duas séries temporais e tenta encontrar um caminho ótimo entre as duas (Sardá-Espinosa, 2017). Deste modo, permite reconhecer formas diferentes, mesmo se as séries temporais apresentarem um *shift* no tempo ou diferentes escalas. Os pontos inicial e final das séries temporais deverão coincidir, mas os restantes pontos podem estar desfasados no tempo.

Como algoritmo de *clustering* elegeu-se o *clustering* hierárquico aglomerativo com o método de Ward (Ward, 1963). A vantagem da utilização do *clustering* hierárquico prende-se com o facto de não ser necessário especificar o número de *clusters* a priori, sendo este possível de inferir a partir do dendrograma obtido. A avaliação da partição obtida pode ser efetuada através de um grande número de índices disponíveis, tendo-se optado pela utilização dos quatro seguintes: o índice de Dunn, *silhouette*, o índice DB (Davies-Bouldin) e o índice CH (Calinski-Harabasz) (Arbelaitz et al., 2013).

Depois de avaliar o algoritmo de *clustering*, espera-se que as séries temporais incluídas em cada *cluster* sejam similares. Assim, o cálculo de um protótipo de todas as séries temporais de cada *cluster* pode ser de grande importância para representar os *clusters* e, em particular, para sumariar as características mais importantes das séries incluídas num mesmo *cluster*. Uma das formas mais simples de calcular o protótipo é através da mediana em cada instante de tempo.

RESULTADOS

A base de dados de telemedição utilizada neste trabalho foi recolhida no âmbito do projeto iPerdas 2016 – Iniciativa nacional para a gestão eficiente de perdas (iperdas.lnec.pt). Um conjunto de contadores Itron – Aquadist foi disponibilizado às EG com um rácio de acordo com a Portaria nº 21/2007 de 26 de Setembro, Diário da República: 20 contadores DN15 (resolução de 0.1 L/pulso, $Q_1=6.3$ L/h, $Q_2=10$ L/h, $Q_3=2.5$ m³/h, $Q_4=3.125$ m³/h); e 10 contadores DN20 (resolução de 1 L/pulso, $Q_1=10$ L/h, $Q_2=16$ L/h, $Q_3=4.0$ m³/h, $Q_4=5.0$ m³/h). Adicionalmente, 10 *data loggers* também foram disponibilizados às EG, sendo que cada *data logger* pode ser ligado a dois contadores em simultâneo.

Sempre que ligados aos contadores, os *data loggers* enviam um impulso quando uma quantidade de água equivalente à resolução do contador é consumida. Por outras palavras, considerando que se está no instante de tempo 0, se o *logger*, ligado a um contador DN15, envia um impulso no instante t , isto significa que, entre 0 e t , foram consumidos 0.1 L de água. Assim, o intervalo de tempo entre registos consecutivos de eventos é aleatório. Deste modo, os dados foram normalizados para um passo de tempo fixo, por exemplo, de 15 minutos, e convertidos em L/h. Dado o modo de recolha dos dados, não é possível saber se existem eventos não registados. Consequentemente, após a normalização dos dados para um passo de tempo de 15 minutos e L/h, não há valores omissos.

Os dados de telemedição utilizados foram recolhidos em 275 utilizadores de 10 EG de sistemas de distribuição de água. No caso dos perfis de consumo e dos consumos mínimos domésticos, os dados de uma das EG não foram considerados para análise, ficando apenas 261 utilizadores de 9 EG. Esta opção deveu-se a características, quer em escala, quer em comportamento, que diferenciam esta EG e que, portanto, enviesavam os resultados das restantes. Para cada utilizador, foram recolhidos dados durante cerca de uma semana, entre junho de 2016 e abril de 2017.

Nas subsecções seguintes serão apresentados os resultados obtidos com estes dados em termos de perfis de consumo, consumos mínimos domésticos e padrões de consumo.

Perfis de consumo

Considerando a análise da distribuição do volume medido segundo as classes de caudal mencionadas na Metodologia, verificou-se que as classes predominantes foram]240, 500] e]500, 1 000] L/h. Verificou-se, ainda, que a percentagem de volume

medido nas classes baixas de caudal (i.e., abaixo de 120 L/h), tipicamente devido a usos de baixa intensidade e de longa duração, é praticamente desprezável, pelo que as fugas domiciliárias não são em geral expressivas. Na Figura 1 encontram-se representados os perfis de consumo global das EG para contadores DN15 e DN20 e respetivas bandas de tolerância para os erros de medição dos contadores utilizados. Os dados de 261 clientes de nove EG foram utilizados para o cálculo destes perfis de consumo. Verifica-se que estes são semelhantes ao perfil tipo 1 apresentado em Vermersch et al. (2016), correspondente a apartamento sem jardim ou piscina e sem reservatório domiciliário. Deste modo, o perfil de consumo global adotado corresponde a uma boa aproximação inicial que cada EG deve aferir. Além disso, verifica-se ainda que estes contadores se encontram bem dimensionados, face às suas características metrológicas e respetivas bandas de tolerâncias apresentadas na Figura 1 para o DN15 e DN20.

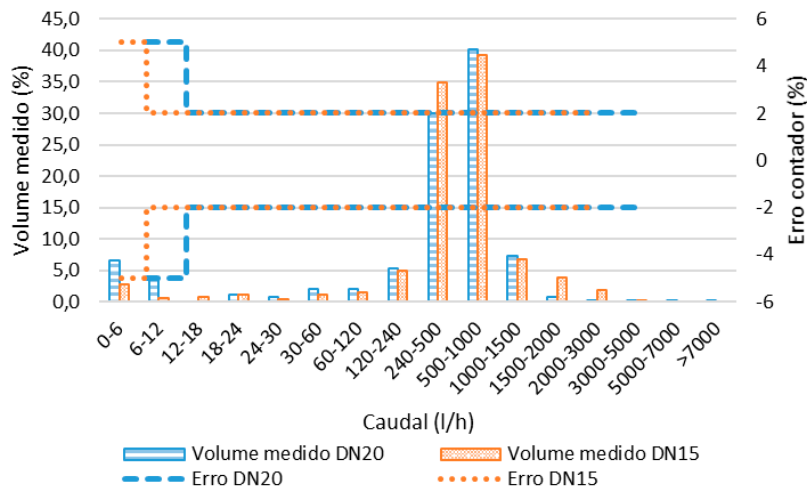


Figura 1 | Perfis de consumo por tipo de contador (DN15 e DN20) e respetivas curvas de erro dos contadores para uma amostra de 41 clientes DN20 e para uma amostra de 55 clientes DN15.

Na Tabela 1 apresentam-se os consumos médios diários por tipo de contador (DN15 e DN20) e por tipo de alojamento (apartamento e vivenda) para cada EG. É ainda indicada a dimensão da amostra existente em cada caso (entre parêntesis). Na generalidade das EG, verificam-se consumos médios diários mais elevados em vivendas do que em apartamentos. Este facto torna-se mais perceptível na amostra de contadores DN15, devido à baixa dimensão da amostra de contadores DN20. Contudo, em termos médios, a mesma conclusão é válida para estes últimos contadores. Conclui-se ainda que o consumo em contadores DN20 é ligeiramente superior ao consumo em contadores DN15.

Tabela 1 | Consumos médios diários (L/(utilizador-dia)) por contadores DN15 e DN20 e por tipo de alojamento (apartamento e vivenda). O número de utilizadores é apresentado entre parêntesis.

Entidade Gestora	Consumo médio diário DN15		Consumo médio diário DN20		Consumo médio diário global
	Apartamento	Vivenda	Apartamento	Vivenda	
EG #1	292.7 (15)	-	395.6 (6)	313.4 (1)	321.7
EG #2	-	-	-	4 926.2 (14)	4 926.2
EG #3	261.2 (34)	496.2 (17)	-	-	324.0
EG #4	293.9 (28)	-	340.5 (4)	-	299.7
EG #5	348.5 (5)	468.1 (27)	-	-	449.5
EG #8	356.6 (10)	356.3 (3)	-	-	356.6
EG #9	280.5 (1)	584.0 (24)	314.8 (5)	944.9 (2)	555.0
EG #10	303.2 (11)	437.6 (10)	314.0 (1)	-	410.1
EG #11	242.1 (17)	353.1 (7)	182.2 (7)	-	277.7
EG #12	404.8 (6)	522.4 (5)	339.2 (9)	462.7 (6)	418.1
Média ponderada	291.1	490.5	311.0	553.3 ^(*)	377.3 ^(*)

^(*) Excluindo a EG #2, por ser considerada um *outlier* na amostra.

Consumos mínimos domésticos

Para nove das 10 entidades, os valores mínimos da série mediana foram obtidos entre as 2:00 e as 5:00 da manhã, inclusive, tendo este sido definido como o período de consumo mínimo doméstico. Na EG #2 verificou-se que o período de consumo mínimo não correspondia ao período noturno, tendo-se verificado que este se registava entre as 13:00 e as 17:00, inclusive, durante o período de campanha. De facto, nas habitações desta EG ocorrem regas durante a noite, levando a que o consumo mínimo se verifique durante o dia.

A análise dos consumos, durante o período de consumo mínimo, foi feita tendo em conta o DN dos contadores (DN15 e DN20) e o tipo de alojamento (apartamento e vivenda), para cada EG. Assim, a média global ponderada dos consumos médios das EG no período de consumo mínimo foi de 1.2 L/(utilizador·h) para contadores DN15 e de 2.0 L/(utilizador·h) para contadores DN20. No que diz respeito à estimativa das fugas domiciliárias, em vez dos consumos médios, foram considerados os consumos mínimos no período de consumo mínimo e foi calculada a média global ponderada dos mesmos. A estimativa das fugas domiciliárias no período de consumo mínimo foi de 0.1 L/(utilizador·h) para contadores DN15 e de 0.4 L/(utilizador·h) para contadores DN20. Para o cálculo destes valores, excluiu-se o valor da EG cujo período de consumo mínimo não corresponde ao período noturno, uma vez que também apresenta um valor de consumo médio no período de consumo mínimo muito elevado, 57.25 L/(utilizador·h), quando comparado com os restantes. Verifica-se, assim, que os consumos médio e mínimo no período de consumo mínimo são mais elevados em contadores DN20 do que em contadores DN15, tal como tinha acontecido com o consumo médio diário apresentado anteriormente. Em termos comparativos, note-se que, no Reino Unido, o valor médio estimado no período de consumo mínimo é de 1.7 L/(utilizador·h), enquanto que no Canadá chega a 3.0 e na Malásia a 5.0 (Fantozzi e Lambert, 2012). Estes valores refletem diferentes hábitos de consumo, tipologias de alojamento e condições climáticas, assim como a existência de reservatórios domiciliários.

A Figura 2 (a) apresenta os valores das médias truncadas para as EG que realizaram campanhas de telemedicação em contadores DN15, separados por apartamentos e vivendas. Apresentam-se ainda as médias globais (média ponderada das médias truncadas) por tipo de alojamento (apartamento ou vivenda). Concluiu-se que a média ponderada dos consumos médios no período de consumo mínimo dos clientes com contadores DN15 é bastante superior quando se trata de uma vivenda, 1.8 L/(utilizador·h), do que quando se trata de um apartamento, 0.9 L/(utilizador·h). Esta diferença já tinha sido verificada anteriormente no caso dos consumos médios diários, onde, em contadores DN15, o consumo em vivendas, 490.5 L/(utilizador·dia) é bastante superior ao consumo em apartamentos, 291.1 L/(utilizador·dia). No caso da EG #12, para DN15 vivendas, foi calculada a média truncada a 15%, de modo a eliminar um cliente atípico. A Figura 2 (b) permite analisar a dispersão dos consumos médios no período de consumo mínimo de todos os clientes com contadores DN15 em apartamentos. Verifica-se que existem clientes com consumos médios atípicos, o que justifica a utilização da média truncada, e que, mesmo dentro dos consumos médios típicos, existe alguma variabilidade de valores.

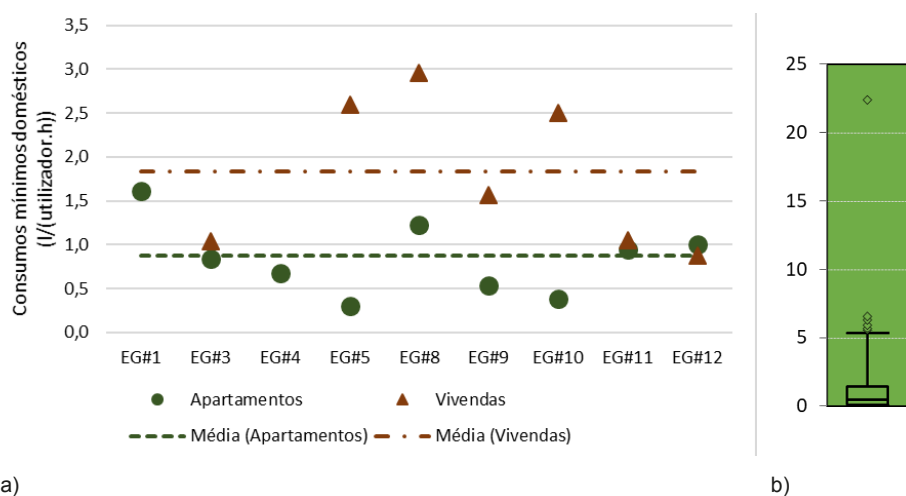


Figura 2 | Consumos médios no período de consumo mínimo dos clientes com contadores DN15: a) médias truncadas por EG e média global por tipo de alojamento (apartamento e vivenda); b) boxplot em apartamentos.

Na Figura 3 encontram-se representados os valores das médias truncadas por EG em contadores DN20 por tipo de alojamento, tais como as suas médias globais. Em oposição ao caso dos contadores DN15, verifica-se que o consumo médio no período de consumo mínimo dos clientes em vivendas, 2,0 L/(utilizador·h), é inferior ao consumo mínimo dos clientes em apartamentos, 2,3 L/(utilizador·h). Esta diferença é contrária à existente em termos de consumos médios diários, onde o consumo de uma vivenda, 553.3 L/(utilizador·dia), é superior ao de um apartamento, 311.0 L/(utilizador·dia). Contudo, as conclusões poderão estar enviesadas dada a reduzida dimensão da amostra no caso dos contadores DN20 (32 apartamentos e 9 vivendas), pelo que não é possível retirar conclusões mais detalhadas.

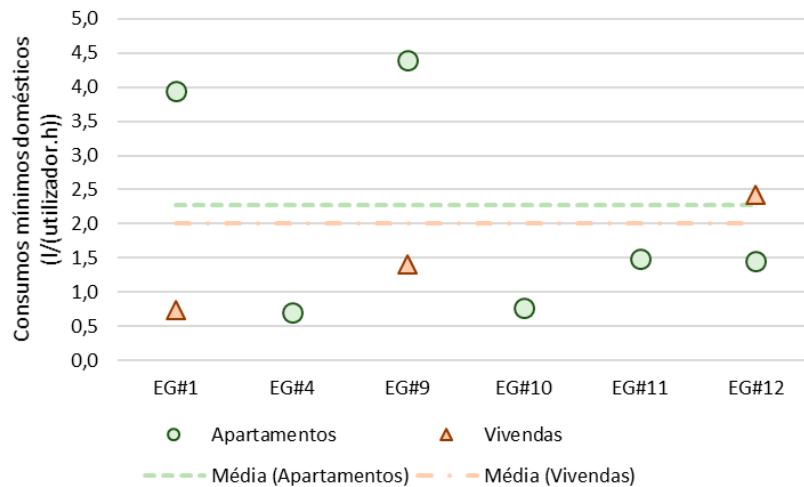


Figura 3 | Média global e média truncada dos consumos médios no período de consumo mínimo dos clientes com contadores DN20 por tipo de alojamento (apartamento e vivenda).

Padrões de consumo

Relativamente aos *clusters* obtidos para a tipificação dos padrões de consumo, a análise dos índices de validação dos *clusters* indicou que o número ótimo de *clusters* era quatro (Figura 4). Analisando os resultados da partição em quatro *clusters*, é possível concluir que as séries temporais do primeiro *cluster* (7 séries) têm tendência a ter consumos consideráveis durante todo o dia, mesmo no período noturno. Estes clientes caracterizam-se por um consumo mensal elevado. O segundo *cluster* (9 séries) tem séries temporais com elevados consumos, aproximadamente constantes, durante 2 ou 3 horas. Estas séries temporais apresentam ainda elevados consumos no período noturno. A série temporal do terceiro *cluster* tem um consumo muito elevado (cerca de 2000 L/h) entre as 18h45 e as 19 h, o que torna este cliente um *outlier* e um caso que a EG deverá analisar. Contudo, acredita-se que as 258 séries do quarto *cluster* não deverão ter todas o mesmo comportamento. Assim, aplicou-se uma análise de *subclustering*, ou seja, foi aplicado o método de Ward apenas a estas 258 séries e foi efetuada uma nova avaliação de *clusters*. Neste caso, o número ótimo de *subclusters* não é tão claro. Contudo, dois foi o número mais frequente de acordo com o conjunto de índices utilizado (Figura 5). Para comparar a forma das séries temporais destes dois *subclusters*, utilizou-se a mediana e os quartis como protótipos de cada um dos *subclusters* (Figura 6). Embora o padrão das medianas e dos quartis de ambos os *subclusters* sejam similares, a escala é diferente. Deste modo, verifica-se que estes dois grupos de clientes têm um comportamento de consumo de água idêntico, mas a quantidade de água consumida é diferente (cerca de 1.5 L/h no *cluster* 1 e cerca de 6 L/h no *cluster* 2), sendo este o fator de separação. Se tivessem sido considerados cinco *clusters* na primeira análise, o quarto *cluster* teria sido dividido nestes dois *subclusters*, mantendo-se os restantes *clusters* como primeiro, segundo e terceiro. Apesar de esta separação não ser tão clara, a escala é também um fator importante para a análise. Assim, acredita-se que a base de dados deverá ser dividida em cinco *clusters*.

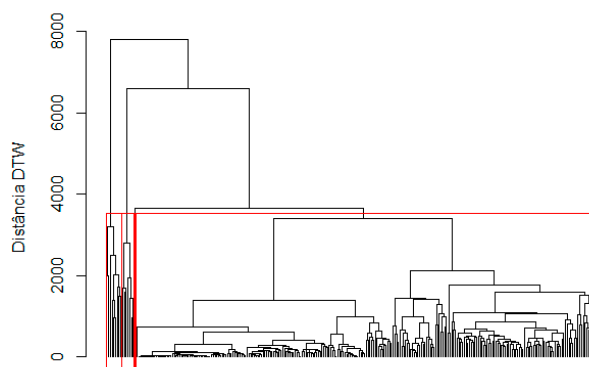


Figura 4 | Dendrograma da análise de *clusters* obtido com a distância DTW aplicada ao conjunto de todas as séries temporais e respetiva partição em quatro *clusters*.

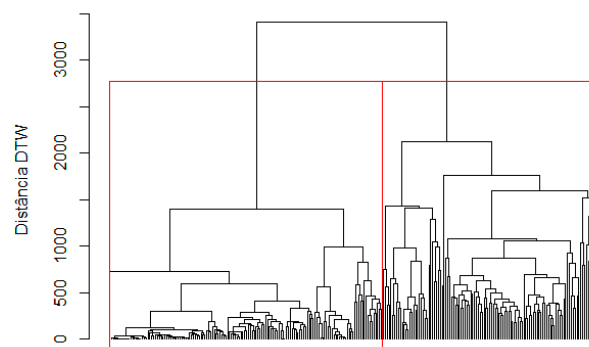


Figura 5 | Dendrograma da análise de *clusters* obtido com a distância DTW aplicada ao subconjunto de séries temporais e respetiva partição em dois *clusters*.

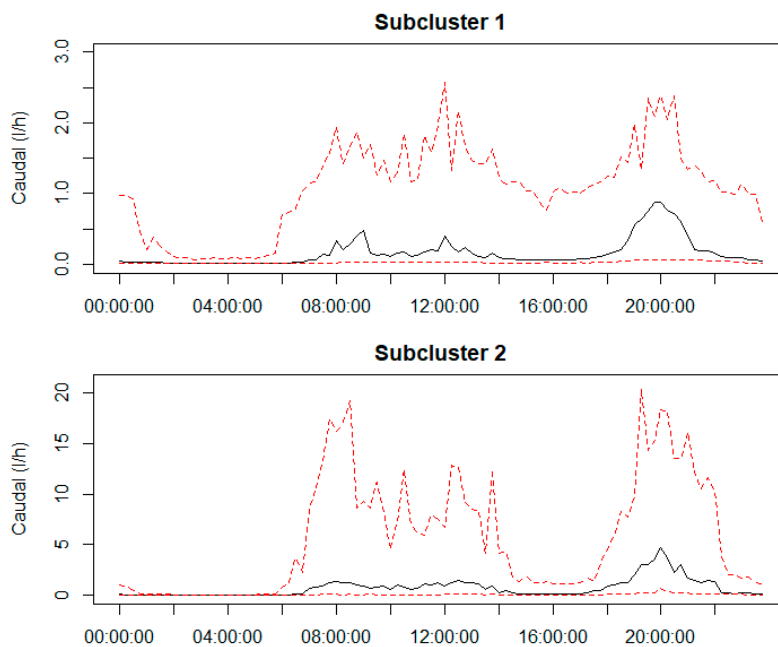


Figura 6 | Mediana (linha contínua preta), 1º e 3º quartis (linhas tracejadas vermelhas) dos *subclusters* 1 e 2 obtidos na análise de *subclustering*.

Considerando a partição em cinco *clusters*, o consumo no período noturno presente nas séries temporais dos dois primeiros *clusters* indicam que estes clientes devem ter consumo exterior. De facto, a análise de informação externa (imagens de satélite do alojamento) indica que 12 das 16 séries temporais estão associadas a moradias com piscinas e/ou jardins. Nas restantes 4 séries não foi possível concluir sobre a existência ou não de consumo exterior. Contudo, genericamente, pode concluir-se que as séries temporais dos *clusters* 3 e 4 têm consumo exterior. As séries temporais dos quarto e quinto *clusters* também podem ter consumo exterior, mas este poderá não ter sido detetado quando os dados foram recolhidos ou poderá ser residual.

Na Tabela 2 apresenta-se o resumo das características de cada um dos cinco *clusters*, tais como a sua dimensão e o comportamento do consumo que os caracteriza. Na Figura 7, encontram-se representados cinco exemplos de séries temporais, pertencentes a cada um dos *clusters*, em escala logarítmica. A correspondência entre as séries e os *clusters* a que pertencem pode ser encontrada na Tabela 2. Por exemplo, quando no *cluster* 1 se indica C5_42516 significa que, na Figura 7, a série com legenda C5_42516 foi incluída no *cluster* 1.

Tabela 2 | Caracterização dos *clusters* obtidos numa amostra de 275 clientes.

Cluster	Caracterização
1	Número de séries: 7 Tendência para ter consumos durante todo o dia, incluindo durante a noite Série exemplo: C5_42516
2	Número de séries: 9 Algumas horas com consumos bastante elevados (cerca de 400 L/h) aproximadamente constantes, mesmo durante a noite Elevados consumos Série exemplo: C2_2G01
3	Número de séries: 1 Consumo muito elevado de 2000 L/h entre 18:45-19:00 Série exemplo: C9_2633
4	Número de séries: 144 Comportamento normal com consumos reduzidos (próximos de 0) durante a noite e consumos mais elevados (cerca de 4 L/h) em determinados períodos do dia Consumo de água muito baixo (cerca de 1.5 L/h) Série exemplo: C11_110201
5	Número de séries: 114 Comportamento normal com consumos baixos (próximos de 0) durante a noite e consumos elevados (cerca de 10 L/h) em determinados períodos do dia Consumo de água baixo (cerca de 6 L/h) Série exemplo: C1_38753

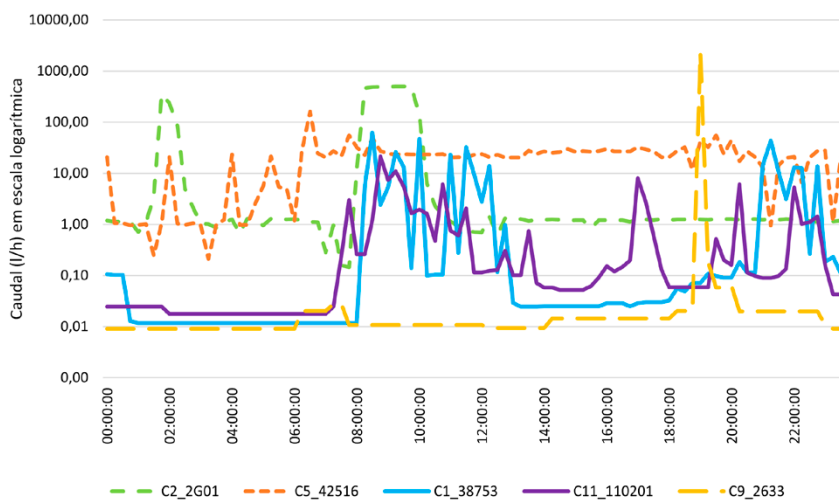


Figura 7 | Exemplos de séries temporais presentes em cada *cluster*.

CONCLUSÕES

As campanhas de telemedição realizadas no decurso do projeto iPerdas permitiram a obtenção de dados de 275 clientes de 10 entidades gestoras participantes com parque de contadores. Estes dados foram utilizados para o estabelecimento de uma metodologia com vista à obtenção dos perfis e padrões de consumo e consumo mínimo doméstico.

Relativamente aos perfis de consumo, estes permitem a contabilização do volume medido segundo determinadas classes de caudal para apoio na estimativa do erro global do parque de contadores. Neste estudo, estes permitiram concluir que as classes predominantes são]240, 500] e]500, 1000] L/h e que as fugas (caudal abaixo de 120 L/h) são praticamente irrelevantes para o global dos dados analisados.

Com base nos dados de telemedição, analisaram-se também os consumos mínimos domésticos para apoiar na estimativa de perdas reais em sistemas de distribuição de água. Para a globalidade das EG, verificou-se que o consumo mínimo doméstico foi de 1.2 L/(utilizador·h) para contadores DN15 e de 2.1 L/(utilizador·h) para contadores DN20. Contudo, é de realçar que a amostra selecionada para telemedição é reduzida e pode não ser representativa de todas as tipologias típicas de consumos existentes nas EG. Deste modo, é recomendado que as EG continuem com estudos deste tipo.

No que diz respeito aos padrões de consumo, estes foram obtidos através de uma análise de *clustering* das séries temporais de consumo de água. A análise de *clusters* levou à partição da base de dados em cinco *clusters*. O primeiro *cluster* inclui sete séries temporais com tendência para ter consumo durante todo o dia, incluindo no período noturno; são séries com consumos mensais elevados. O segundo *cluster* é composto por nove séries temporais e caracteriza-se por algumas horas (2 a 3) com um consumo elevado aproximadamente constante, mesmo durante a noite. Uma única série constitui o terceiro *cluster*, caracterizando-se por um consumo muito alto (2000 L/h) entre as 18:45 e as 19:00. Os dois últimos incluem 258 séries temporais (144 e 114 no quarto e no quinto *clusters*, respetivamente). Ambos são constituídos por séries com um comportamento normal de consumo de água, caracterizado por um baixo consumo durante a noite e alto no início e fim do dia. A diferença reside na escala de consumo, que é maior no quinto *cluster*. Para além de distinguir os grupos de clientes pela escala de consumo, esta análise permitiu ainda identificar clientes com consumo exterior ou apenas com consumo interior.

Em termos de consumo interno e externo, o comportamento das séries temporais dos dois primeiros *clusters* indicam que estas séries deverão estar associadas a clientes com consumo exterior de água. Informações externas sobre os dados permitiram verificar que quase todas as séries destes dois *clusters* (12 das 16) estão associadas a moradias com piscinas e/ou jardins. Relativamente às restantes quatro, não foi possível explicar estes consumos.

A análise dos dados de telemedição permite assim conhecer como é que os clientes usam a água e, conseqüentemente, estimar melhor as perdas de água e sensibilizar para uma utilização mais eficiente da mesma. Este trabalho permitiu às EG ter conhecimento sobre o potencial da telemedição para melhorar a gestão de perdas de água e conhecer melhor as suas tipologias de consumo doméstico.

Relativamente a trabalho futuro, considera-se que os padrões de consumo deverão ser analisados separadamente para dias úteis e sábados, domingos e feriados, de acordo com a época do ano. A utilização dos conhecimentos obtidos neste estudo poderá ser ainda útil para um estudo de amostragem dos clientes onde deverá ser instalada telemedição. A instalação de telemedição noutros clientes permitirá ainda aumentar a amostra e, como tal, melhorar as estimativas de perfis de consumo e consumos mínimos domésticos por tipologia e tipo de cliente. Além disso, será possível inferir a quantidade de água que é consumida em cada setor de rede, calcular as perdas de água e priorizar os setores de rede.

AGRADECIMENTOS

Os autores agradecem a colaboração das entidades gestoras participantes no projeto Inicativa Nacional para gestão eficiente de perdas (iPerdas 2016). O primeiro autor agradece à Fundação para a Ciência e Tecnologia (FCT), Portugal, pela Bolsa de Doutoramento SFRH/BD/131382/2017. Agradece-se ainda o financiamento da FCT, através do projeto UID/Multi/04621/2019 do CEMAT/IST-ID, Centro de Matemática Computacional e Estocástica, Instituto Superior Técnico, ULisboa.

REFERÊNCIAS

- Alegre, H., Coelho, S.T., Almeida, M.C., Vieira, P. 2005. *Guia Técnico nº3: Controlo de perdas de água em sistemas de adução e distribuição*. IRAR, INAG, LNEC, Lisboa, Portugal.
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J.M., Perona, I. 2013. An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1), 243-256. <https://doi.org/10.1016/j.patcog.2012.07.021>
- Arregui, F., Cabrera Jr., E., Cobacho, R. 2006. *Integrated water meter management*. IWA Publishing, Londres.
- Arregui, F., Balaguer, M., Soriano, J., García-Serra, J. 2016. Quantifying measuring errors of new residential water meters considering different customer consumption patterns. *Urban Water Journal*, 13(5), 463-475. <https://doi.org/10.1080/1573062X.2014.993999>
- Arregui, F., Gavara, F.J., Soriano, J., Pastor-Jabaloyes, L. 2018. Performance analysis of ageing single-jet water meters for measuring residential water consumption, *Water*, 10, 612. <https://doi.org/10.3390/w10050612>
- Fantozzi, M., Lambert, A. 2012. Residential night consumption—assessment, choice of scaling units and calculation of variability. *Proceedings of the IWA Water Loss Conference*, Fevereiro, Manila, Philippines, 26-29.
- Farley M., Trow S. 2003. *Losses in water distribution networks*, IWA Publishing, Londres.
- Loureiro, D., Vieira, P., Makropoulos, C., Kossieris, P., Ribeiro, R., Barateiro, J., and Katsiri, E. 2014a. Smart metering use cases to increase water and energy efficiency in water supply systems. *Water Supply*, 14(5): 898-908. <https://doi.org/10.2166/ws.2014.049>
- Loureiro, D., Alegre, H., Coelho, S., Martins, A., and Mamade, A. 2014b. A new approach to improve water loss control using smart metering data. *Water Supply*, 14(4): 618-625. <https://doi.org/10.2166/ws.2014.016>
- Loureiro, D., Rebelo, M., Mamade, A., Vieira, P., Ribeiro, R. 2015. Linking water consumption smart metering with census data to improve demand management. *Water Supply*, 15(6), 1396-1404. <https://doi.org/10.2166/ws.2015.086>
- Loureiro, D., Mamade, A., Cabral, M., Amado, C., Covas, D. 2016. A comprehensive approach for spatial and temporal water demand profiling to improve management in network areas. *Water Resources Management*, 30(10), 3443-3457. <https://doi.org/10.1007/s11269-016-1361-3>
- Sardá-Espinosa, A. 2017. Comparing time-series clustering algorithms in R using the dtwclust package.
- Vermersch, M., Carteado, F., Rizzo, A., Johnson, E., Arregui, F., Lambert, A. 2016. *Guidance notes on apparent losses and water loss reduction planning*. Disponível em: <https://www.mcast.edu.mt/guidance-notes-app-loss/>, acessado em: 14/12/2020.
- Ward, J.H. Jr. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236-244. <https://doi.org/10.1080/01621459.1963.10500845>