



TÉCNICO
LISBOA

Detecção de Fugas em Redes de Distribuição de Água

João André Gonçalves Pereira de Oliveira Fontes

Dissertação para obtenção do Grau de Mestre em

Engenharia Electrotécnica e de Computadores

Orientador: Professor Doutor João Manuel Lage de Miranda Lemos

Júri

Presidente: Professor Doutor João Fernando Cardoso Silva Sequeira

Orientador: Professor Doutor João Manuel Lage de Miranda Lemos

Vogal: Professor Doutor Jorge dos Santos Salvador Marques

Outubro 2016

Agradecimentos

Em primeiro lugar gostaria de expressar a minha gratidão para com o Professor João Miranda Lemos por todas as horas de orientação e pelo apoio que me deu durante o desenvolvimento desta dissertação.

Gostaria também de agradecer à minha família, em particular aos meus pais e irmã, por me terem apoiado durante todo o percurso no Instituto Superior Técnico e pela paciência para me aturar nos momentos de maior aperto e tensão, que foram difíceis não só para mim, mas também para eles.

Aos meus amigos e colegas, com quem tive a sorte e o prazer de partilhar os melhores e os piores momentos ao longo destes anos, dentro e fora do Instituto, agradeço a sua amizade e apoio sem os quais a passagem pelo Técnico teria sido muito mais difícil: Edgar, Catarina, Nuno, André, Raquel, Francisco, Diogo.

At last I would like to address some words to some of the most special people I had the luck and privilege of crossing paths with:

thank you Almina, for all the hours spent together studying, talking, laughing during what was the most difficult semester for me. Just when dropping out was a reality much closer than ever reaching the end, you were the star in the darkness that lit my life;

thank you Daniela and Melinda, for having been two pillars in my life. Your support and friendship were of the utmost importance;

thank you Chiara for all of our bilingual conversations;

thank you Jasper, Francesco, Giampiero, Piercarlo, Perla, Francesca, Malin and Freia for all the moments, for all the smiles and for all the conversations shared. Thank you for your friendship.

Finally I would like to remember someone whose life was taken before time. Thank you Metti for having shown us how life is supposed to be lived. No one lives for ever, but you will be eternal and your passion for life is something we all that had the privilege of knowing you will never forget.

This thesis has been performed in the framework of projects PTDC/EEI-PRO/0426/2014 and

UID/CEC/50021/2013, funded by the Portuguese Foundation of Science and Technology (FCT).

Abstract

This dissertation proposes a solution to the problem of leak detection in a water distribution network with a single input and a single output (SISO). Two types of leaks are considered – burst and growing leaks – for which two detection algorithms are proposed. Both algorithms are based on the demand prediction, which was performed by modelling the demand as the output of an ARMA system and computing the one-step ahead predictor that minimises the mean square error, and on residuals analysis of the predicted water consumption and the observed values. A leak is assumed to exist when the result of the residual analysis is greater than a predefined threshold.

This dissertation also proposes a method for leakage location and detection, requiring flow meters, in distribution networks that can have one or more inputs and of one or more outputs (SIMO, MISO and MIMO). In the case of a SISO network, it is composed of a set of branches and nodes. The method proposed takes advantage of the network representation using its incidence matrix, not requiring that it is simulated using an hydraulic modelling software. It is assumed that the network under study is fully characterised, *i.e.*, the weights of all branches are known.

The leak location system was trained to recognise the effects of a leak in a given pipe in the flow of water in other pipes of the network. For this purpose, a K-means based clustering method was used. It is assumed that not all pipes of the network are metered.

Keywords

Water Networks, System modelling, Prediction, Leakage location, Leakage detection

Resumo

A presente dissertação propõe uma solução para o problema de detecção de fugas numa rede de distribuição de água com um consumidor e um fornecedor – *pipeline*.

Consideram-se dois tipos de rupturas – abruptas e progressivas – para cada um dos quais é proposto um algoritmo de detecção. Ambos os algoritmos têm por base a predição do consumo de água, feita através da modulação do consumo como a saída de um sistema Auto Regressivo de Média Móvel e o cálculo do preditor um passo à frente que minimiza o erro quadrático médio, e a análise dos resíduos entre a predição e as observações de caudal feitas a cada instante. Assume-se a existência de uma fuga quando o resultado da análise dos resíduos ultrapassa um determinado limiar.

Nesta dissertação propõe-se também um método de detecção e localização de fugas, com base em medições de caudal, em redes de distribuição de água que podem ser de um ou mais fornecedores e de um ou mais consumidores, mas que se caracterizam por serem compostas por vários ramos e nós. O método proposto tira partido da representação da rede através da sua matriz de incidência, não sendo por isso necessária a sua simulação com ferramentas de modelação hidráulicas, e pressupõe que a rede está completamente caracterizada, *i.e.* é conhecido o peso de cada ramo.

O sistema de localização de fugas foi treinado de forma a reconhecer os efeitos de uma ruptura em cada um dos ramos da rede no caudal dos restantes ramos que são monitorizados, tendo sido usado um método de *clustering* baseado no algoritmo *K-means* para este efeito. Assumiu-se que apenas alguns dos ramos da rede são monitorizados.

Palavras Chave

Redes de água, Modelação, Predição, Localização de fugas, Detecção de fugas

Índice

1	Introdução	1
1.1	Motivação	3
1.2	Objectivo	4
1.3	Revisão da Literatura	4
1.4	Contribuições	6
1.5	Estrutura da Dissertação	6
2	Distribuição em Rede SISO	
	(Caso em Simulação)	9
2.1	Modelo de Consumo	11
2.2	O Preditor	12
2.3	Detecção de Fugas	14
2.3.1	Detecção de Fugas Abruptas	14
2.3.2	Detecção de Fugas Progressivas	15
2.4	Simulação	17
2.4.1	Simulação sem Fuga	17
2.4.2	Simulação de Fuga Abrupta	20
2.4.3	Simulação de Fuga Progressiva	22
2.5	Simulações de Monte Carlo	25
2.5.1	Fuga Abrupta	25
2.5.2	Fuga Progressiva	27
3	Distribuição em Rede SISO	
	(Caso Real)	31
3.1	Falha na ocorrência dos dados	32
3.2	Identificação do Modelo	32
3.3	Validação do Modelo	33
3.4	Predição	34
3.5	Fins-de-Semana e Feriados	36
3.6	Resultados	37

3.7	Simulação de Fugas	39
4	Distribuição em Rede MIMO	41
4.1	Descrição da rede	42
4.2	Análise da rede	42
4.3	Detecção de Fuga	45
4.4	Simulação de Fuga	45
4.5	Localização de Fuga	47
4.6	Resultados	49
5	Conclusões	53
5.1	Discussão e Conclusões	55
5.2	Trabalho Futuro	56
A	Resultados das Simulações de Monte Carlo para fuga abrupta	63
B	Resultado das Simulações de Monte Carlo para Fuga Progressiva	69
C	Coeficientes dos Modelos de Consumo para cada período do ano	75
D	Resultados da Validação dos Modelos: Autocorrelação do resíduo e consumos real e obtido através do modelo identificado	79

Lista de Figuras

2.1	Mapa de Polos e Zeros do sistema Σ	12
2.2	Diagrama de blocos do modelo do caudal consumido	12
2.3	Saída do sistema Σ . Sinal real (azul) e estimativa (laranja).	17
2.4	Consumo simulado. Sinal real (azul) e estimativa (laranja).	18
2.5	Consumo Médio real (azul) e estimado (laranja)	19
2.6	Integral do erro de predição 1 passo à frente, $I(t)$	19
2.7	Consumo Instantâneo real (azul) e estimado (laranja), usando \hat{Q}	20
2.8	Integral do erro de predição 1 passo à frente com fuga abrupta	21
2.9	Consumo Médio real e estimado em situação de fuga abrupta	21
2.10	Sinal de Alarme para fuga abrupta, $F(t)$	22
2.11	Consumo simulado. Sinal real (azul) e estimativa (laranja)	23
2.12	Integral do erro de predição após filtragem, $I'(t)$	23
2.13	Sinal de Alarme para fuga progressiva, $F'(t)$	24
2.14	Curvas ROC para fuga abrupta	26
2.15	Atraso na detecção em função do limiar	27
2.16	Curvas ROC para fuga progressiva	28
2.17	Atraso na detecção em função do limiar	29
3.1	Autocorrelação do resíduo (em cima) e consumos real e obtido através do modelo identificado (em baixo)	34
3.2	Compromisso entre capacidade de detecção e flexibilidade do modelo	36
3.3	Consumo de uma semana típica (Segunda-feira a Domingo)	37
3.4	Consumo real e previsto de uma semana com feriado (8 de Dezembro)	37
3.5	Consumo real e previsto dos dias feriado	38
3.6	Integral do erro de predição	38
3.7	Curvas ROC para caso real	39
3.8	Atraso de detecção em função do limiar	40
4.1	Rede	43
4.2	Rede com fuga em C_{12}	46

Lista de Figuras

D.1	1 Setembro 2015 a 6 Setembro 2015	81
D.2	7 Setembro 2015 a 27 Setembro 2015	81
D.3	21 Dezembro 2015 a 3 Janeiro 2016	82
D.4	4 Janeiro 2016 a 31 Janeiro 2016	82
D.5	1 Fevereiro 2016 a 14 Fevereiro 2015	83
D.6	15 Fevereiro 2015 a 8 Março 2016	83

Lista de Acrónimos

ARIMA	Autoregressive Integrated Moving Average
SVM	Support Vector Machines
RDA	Redes de Distribuição de Água
IST	Instituto Superior Técnico
EPA	Agência de Protecção Ambiental do governo dos Estados Unidos da América
SISO	Single-Input Single-Output
MIMO	Multiple-Input Multiple-Output
TP	Verdadeiros Positivos
TN	Verdadeiros Negativos
FP	Falsos Positivos
FN	Falsos Negativos
ROC	Receiver Operating Characteristic

1

Introdução

Contents

1.1	Motivação	3
1.2	Objectivo	4
1.3	Revisão da Literatura	4
1.4	Contribuições	6
1.5	Estrutura da Dissertação	6

Neste capítulo apresenta-se o problema das fugas em redes de distribuição de água como motivação para o desenvolvimento desta dissertação. De seguida apresentam-se os objectivos da mesma e, posteriormente, após uma breve revisão da literatura, que permite que o leitor tenha uma noção geral sobre o trabalho desenvolvido na área são enumeradas, as principais contribuições desta dissertação.

1.1 Motivação

Cerca de 70% da superfície terrestre é ocupada por água. No entanto, apenas 3% da água existente no planeta é potável, dos quais dois terços estão inacessíveis à humanidade. Como consequência, mais de mil milhões de pessoas não têm acesso a água potável e cerca de dois mil e setecentos milhões são afectadas pela sua escassez durante pelo menos um mês do ano. Enquanto o número de pessoas afectadas pela escassez de água continua a aumentar, também o consumo de água *per capita* tem aumentado nos locais onde está disponível. Assim, é da maior importância garantir que a água potável distribuída pelas redes urbanas é aproveitada ao máximo, minimizando as perdas no seu transporte.

As rupturas de canos são o maior responsável pela perda de água nos sistemas de distribuição, mas os seus efeitos vão mais além. Uma ruptura num cano implica desperdício de energia a bombear a água perdida e na manutenção da pressão na rede, bem como o desperdício de outros recursos materiais usados no processo de extracção e tratamento da água.

Em Portugal, durante a década de 1990 o volume de água perdido anualmente era de cerca de 50 milhões de metros cúbicos, o que correspondia a cerca de 25% do volume total de água fornecido à rede. Actualmente a quantidade de água desperdiçada nas redes de distribuição varia muito de país para país. Apesar de a maioria dos países se situar no intervalo de 20% a 40%, há países onde as perdas chegam aos 70%.

Por outro lado, as fugas de água nos canos das redes de distribuição em ambiente urbano constituem um perigo para a saúde pública. Em condições normais uma ruptura tem como consequência a saída de água de circulação, devido à pressão nos canos. No entanto, quando há um abaixamento da pressão numa zona da rede com fuga, a água presente no espaço circundante pode ser absorvida pelo cano através do ponto de ruptura, pondo em circulação diversos contaminantes nocivos à saúde tais como vírus e bactérias.

Enquanto que as fugas em grandes condutas são, geralmente, facilmente detectáveis e até localizáveis, já que os seus efeitos se fazem notar à superfície, a maioria das fugas não é sequer detectada até que atinjam proporções que permitam ver os seus efeitos a "olho nu", conduzindo a enormes perdas que poderiam ser evitadas caso de a detecção e localização fossem feitas mais atempadamente.

Por todos os motivos anteriormente explicados é imperativo que as redes de distribuição de

1. Introdução

água em todo o mundo sejam monitorizadas "em tempo real" e que métodos de detecção e localização de fugas sejam desenvolvidos e implementados.

1.2 Objectivo

O objectivo do trabalho realizado nesta dissertação consiste no seguinte:

Dadas observações dos caudais do consumo de água e variáveis auxiliares, pretende-se detectar e localizar a existência de fugas numa rede de distribuição de água em ambiente urbano.

Começa por se considerar o caso da rede do Instituto Superior Técnico (IST) em que há apenas um ponto de medida do caudal. Em seguida, consideraram-se situações mais complexas relativas à detecção e localização de fugas numa rede de distribuição da qual se conhece a estrutura.

Neste trabalho, a detecção de fugas é feita com base em medidas do caudal uma vez que é esta a situação existente no Campus da Alameda do IST. Uma alternativa, fora do âmbito desta dissertação, seria fazer a detecção com base em sensores de pressão, que têm a vantagem de ter um custo inferior aos sensores de caudal.

1.3 Revisão da Literatura

Apesar de a detecção e localização de fugas em tempo útil ser da maior importância para as companhias de água, esta é uma tarefa que se tornou cada vez mais difícil dada a crescente complexidade das redes de distribuição. A constante renovação das redes também constituiu um problema, já que os métodos tradicionalmente usados na detecção de fugas, sejam eles acústicos [10] ou electromagnéticos [3], têm um fraco desempenho quando aplicados aos canos de materiais plásticos. O facto de as fugas apenas poderem ser detectadas se não se encontrarem a uma distância superior a 2 metros do local de inspecção para métodos electromagnéticos e a 250 metros para métodos acústicos [5] e a necessidade de conhecer com exactidão o comprimento do cano entre dois pontos de detecção, o que raramente é possível, levou a que fosse necessário desenvolver técnicas para a detecção e localização de fugas em redes de distribuição mais eficientes e de maior alcance.

Os primeiros métodos propostos para superar a limitação anterior consistiam na utilização de um localizador de canos para a cálculo do seu comprimento, nomeadamente utilizando o radar [7] [8] [26], tendo sido posteriormente proposto um método automatizado para o cálculo do comprimento. Estes processos eram, no entanto, bastante morosos e falíveis.

Em contraste com os métodos anteriormente usados, foi proposto em [29] um método acústico de identificação de fugas que não requer conhecimento das características físicas do cano.

Actualmente as companhias de água têm desenvolvido e investido em estratégias, de maior ou menor complexidade, de análise das variáveis hidráulicas, como a pressão ou o fluxo, nas

suas redes. O objectivo destas estratégias é reduzir o tempo decorrido desde o início da fuga e a sua detecção, bem como o tempo de resposta, isto é, o tempo decorrido desde a detecção até ao início da intervenção [14]. O desenvolvimento destas estratégias só foi possível com o desenvolvimento e massificação das tecnologias de transmissão de dados em tempo real.

Actualmente a maioria dos métodos que se encontra na literatura implica modelar e fazer uma simulação hidráulica extensiva da rede usando ferramentas como o EPANET, disponibilizado pela Agência de Protecção Ambiental do governo dos Estados Unidos da América (EPA) (<http://www.epa.gov/nrmrl/wswrd/dw/epanet.html>). De uma forma simples, quando o estado actual da rede apresenta um grande desvio em relação ao estado expectável, obtido através da simulação, é accionado um alarme que indica a ocorrência de uma falha [27]. As técnicas de detecção e localização de fugas baseadas em modelos hidráulicos foram alvo de estudo extensivo em [9].

Os modelos Autoregressive Integrated Moving Average (ARIMA) exploram a propriedade que, em condições normais, os caudais na rede em cada instante estão relacionados com os caudais em instantes anteriores. Este método tanto pode ser explorado de forma a detectar falhas como a reconstruir sinais enviados por sensores que tenham sido corrompidos [21]. Outras soluções utilizam redes neuronais para criar um modelo probabilístico do caudal futuro com base nos registos disponíveis até ao momento, o qual é usado para identificar comportamentos anormais [13]. Foram também propostos métodos que utilizam Support Vector Machines (SVM) [30]. Neste conjunto de métodos as medições de variáveis hidráulicas como a pressão e o fluxo de água fornecidas por sensores são analisados de forma a identificar e localizar fugas.

Outros métodos que podem ser encontrados na literatura que podem ser aplicados a redes de distribuição em tempo real para detecção de fugas incluem análise de séries temporais [16], filtros de Kalman [15] e a análise de transitórios [5].

Em [20] o problema da detecção e localização de fugas é formulado como um problema de estimação de mínimos quadrados, mas a estimação de parâmetros em modelos não lineares de redes de distribuição de água não é uma tarefa simples já que, na maioria dos casos, o problema é indeterminado devido ao reduzido número de sensores de pressão presentes nas redes.

Alternativamente, em [17] é estudada qual a disposição dos sensores de pressão na rede, aplicado à cidade de Barcelona, que melhor potencia a detecção de fugas e em [18] é proposto um método de análise de resíduos em tempo real utilizando um limiar de decisão que tem em conta as incertezas do modelo e o ruído das medições. Mais recentemente [24] propõe uma estratégia de disposição de sensores de pressão óptima numa rede de forma a maximizar um conjunto de parâmetros de diagnóstico.

Outras técnicas utilizadas baseadas em sensores de pressão incluem a utilização de redes neuronais artificiais [1] [25] [2], treinadas com os dados de pressão e fluxo de água resultantes das simulações, SVM's [4], Clustering [23] e *Fuzzy Pattern Recognition* [28].

1. Introdução

Em [22] é proposto um método com base em medições de caudais, em vez de pressões. O método consiste na análise de resíduos, sendo estes definidos como a diferença entre os caudais nominais (sem fuga) e os caudais (eventualmente) com fugas. Apesar de o uso de caudalímetros ser possível nas redes de transporte, nas Redes de Distribuição de Água (RDA) a quantidade de canos existente é muito maior e estes sensores têm a desvantagem de serem não só mais caros, mas também de terem um custo de instalação e manutenção superior.

Recentemente foi proposta uma metodologia de análise de resíduos baseada em simulação hidráulica e sensores de pressão [19]. Os resíduos são definidos como a diferença entre as medições e as estimativas obtidas através da simulação sem fugas, a partir dos quais se identifica, no caso de existir, a fuga. Para a localização da fuga, foram também simuladas todas as possíveis localizações de fugas na rede e armazenados os dados daí resultantes, criando assim uma "assinatura" para cada fuga. Os resíduos calculados são comparados com a assinatura de cada fuga, sendo assim determinada a localização da fuga. O método proposto foi testado, com sucesso, na cidade de Barcelona.

Na cidade de Lisboa a conjugação de vários dos métodos anteriormente apresentados permitiu reduzir as perdas de água para aproximadamente 10%, dados de 2011, [6], o que se traduz numa poupança financeira de cerca de 6 milhões de euros e numa redução de impacte ambiental de 1800 toneladas de dióxido de carbono, 17.2 toneladas de cloro e redução energética de 4.7 GWh.

1.4 Contribuições

As principais contribuições desta dissertação consistem em algoritmos para a detecção e localização de fugas em redes de distribuição de água em tempo real utilizando sensores de caudal e sem recorrer a simulações hidráulicas da mesma, tirando proveito da matriz de incidências pela qual a rede é representada e aplicando técnicas de aprendizagem automática.

A maior contribuição desta dissertação consiste na aplicabilidade dos métodos desenvolvidos à monitorização da rede do campus da Alameda do Instituto Superior Técnico (IST), que permitem fazer a detecção de fugas de forma automática.

1.5 Estrutura da Dissertação

A presente dissertação está organizada da seguinte forma.

No capítulo 1 começa por se explicar as motivações e os objectivos a alcançar com o desenvolvimento do trabalho. Após uma revisão da literatura disponível sobre o tema, apresentam-se as principais contribuições desta dissertação.

No capítulo 2 considera-se um caso em simulação onde se considera um fornecedor e um cliente – Single-Input Single-Output (SISO). Define-se o modelo do consumo e seu algoritmo

de predição, e apresentam-se os métodos de detecção para os vários tipos de fugas possíveis. Por fim, o desempenho dos algoritmos desenvolvidos é avaliado através de simulações de Monte Carlo.

No capítulo 3 aplicam-se os métodos desenvolvidos no capítulo 2 a um caso real, com dados relativos ao consumo de água no *campus* da Alameda do IST.

No capítulo 4 trata-se o caso de distribuição de água através de uma rede Multiple-Input Multiple-Output (MIMO) explorando a representação da rede pela sua matriz de incidências. Além do objectivo comum às fases anteriores do trabalho de detecção de fugas na rede, pretende-se localizar com um certo grau de certeza a ocorrência da fuga através da aplicação de técnicas de aprendizagem automática.

No capítulo 5 apresentam-se as conclusões retiradas do trabalho desenvolvido e aponta-se o trabalho que deverá ser desenvolvido no futuro neste tema.

2

Distribuição em Rede SISO (Caso em Simulação)

Contents

2.1	Modelo de Consumo	11
2.2	O Preditor	12
2.3	Detecção de Fugas	14
2.4	Simulação	17
2.5	Simulações de Monte Carlo	25

2. Distribuição em Rede SISO (Caso em Simulação)

Neste capítulo começa por se considerar um exemplo de simulação, a propósito do qual se definem o modelo do consumo, o algoritmo de predição e o algoritmo de detecção de fugas utilizados. Estes conceitos serão em seguida aplicados a um caso real com dados relativos ao consumo de água em campus do IST.

2.1 Modelo de Consumo

Começa por se considerar um caso simples, em simulação, com apenas um consumidor. O consumo de água $Q(t)$, em que $t \in \mathbb{N}$ representa o tempo discreto, é modelado por um sistema Auto Regressivo de Média Móvel (ARMA) com função de transferência

$$\frac{C(z)}{A(z)}, \quad (2.1)$$

e por um filtro, que traduz a sazonalidade diária, à saída do qual se soma uma constante, \bar{Q} , que representa o consumo médio. O filtro força a periodicidade do consumo, com um período de 24 horas. Por fim é somado um escalão de amplitude d , proporcional a \bar{Q} , que representa a existência de uma fuga.

Uma vez que os dados do consumo são disponibilizados a cada quinze minutos pela empresa que abastece o Campus da Alameda do IST, será esse o período de amostragem usado, o que se traduz em quatro leituras por hora, resultando em 96 registos diários. Assim, a função de transferência do filtro de periodicidade é

$$\frac{1}{1 - z^{-96}}. \quad (2.2)$$

Por forma a ter-se um sinal com uma forma similar ao consumo diário real, as suas condições iniciais devem ser especificadas em 96 pontos. A entrada do sistema é ruído branco, $e(t)$, de média nula e variância σ^2 .

Os polinómios $C(z)$ e $A(z)$ em 2.1 são ambos de segundo grau e de Hurwitz, e os zeros do segundo polinómio foram escolhidos de forma a obter uma resposta oscilatória:

$$\begin{aligned} A(z) &= [(z - (0.4 + j0.4))][z - (0.4 - j0.4)] \\ C(z) &= z^2 \end{aligned} \quad (2.3)$$

2. Distribuição em Rede SISO (Caso em Simulação)

A figura 2.1 mostra o mapa de polos e zeros do sistema Σ .

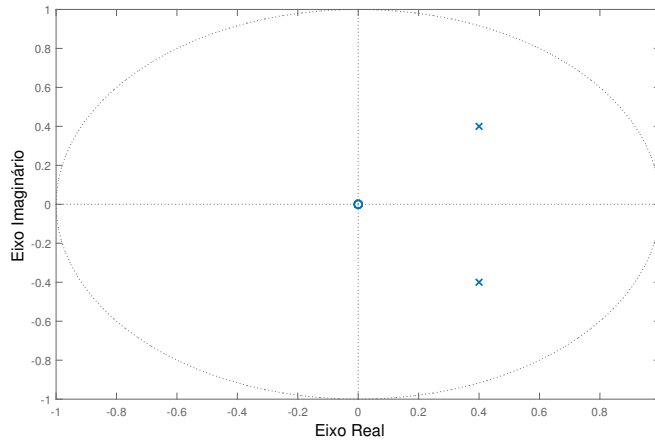


Figura 2.1: Mapa de Polos e Zeros do sistema Σ

Na figura 2.2 está representado o diagrama de blocos sistema anteriormente descrito.

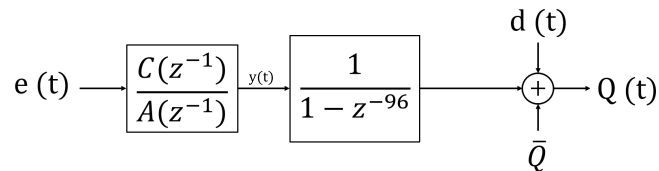


Figura 2.2: Diagrama de blocos do modelo do caudal consumido

Nesta fase não é feita distinção entre dias úteis, fins-de-semana e dias feriados, tal como não é tido em conta o período do ano (Verão/Inverno, período lectivo, período de exames, período de férias, etc).

2.2 O Preditor

Por forma a detectar a existência de uma fuga foi projectado um observador-preditor para o consumo, que dá a sua previsão no instante t dadas as observações até ao instante $t - 1$, $\hat{Q}(t|t - 1)$. Para tal, calculou-se o preditor de erro quadrático médio mínimo [12] para o sistema Σ obtido através de

$$\hat{y}(t + m|t) = \frac{G_m^*(q^{-1})}{C^*(q^{-1})} y(t), \quad (2.4)$$

onde $C^*(q^{-1})$ é o polinómio recíproco de $C(q)$, dado por

$$C^*(q^{-1}) = q^{-n} C(q), \quad (2.5)$$

e $G_m^*(q^{-1})$ e $F_m^*(q^{-1})$ satisfazem

$$\frac{C^*(q^{-1})}{A^*(q^{-1})} = F_m^*(q^{-1}) + q^{-m} \frac{G_m^*(q^{-1})}{A^*(q^{-1})}. \quad (2.6)$$

Os operadores q e q^{-1} são, respectivamente, os operadores avanço e atraso.

O preditor óptimo de $y(t)$ é então dado por

$$\hat{y}(t+m|t) = \alpha y(t-m+1) + \beta y(t-m), \quad m = 1 \quad (2.7)$$

e os coeficientes do preditor, α e β , são obtidos por divisão longa dos polinómios $C(z)$ e $A(z)$.

A variância do erro de predição é dada por

$$(1 + f_1^2 + \dots + f_{m-1}^2) \sigma^2 \quad (2.8)$$

em que $f_i, i = 1, \dots, n-1$ são os coeficientes de $F_m^*(q^{-1})$:

$$F_m^*(Q^{-1}) = 1 + f_1 q^{-1} + \dots + f_{m-1} q^{-m-1}. \quad (2.9)$$

Uma vez que $y(t)$ não está directamente acessível, é necessário aplicar o filtro inverso a $Q(t)$ e subtrair o seu valor médio, obtendo-se uma reconstituição do sinal à saída do sistema Σ , $y_{rec}(t)$, dado por

$$y_{rec}(t) = Q(t) - \bar{Q}(t) - Q(t-96) + \bar{Q}(t-96) \quad (2.10)$$

$$\Leftrightarrow y_{rec}(t) = \begin{cases} Q(t) - \bar{Q}(t) & , \text{ se } t \leq 96 \\ Q(t) - \bar{Q}(t) - Q(t-96) + \bar{Q}(t-96) & , \text{ se } t \geq 97 \end{cases} \quad (2.11)$$

em que $\bar{Q}(t)$ é o valor médio do consumo.

Substituindo $y(t)$ por $y_{rec}(t)$ na equação (2.7) tem-se

$$\hat{y}(t+m|t) = \alpha y_{rec}(t-m+1) + \beta y_{rec}(t-m). \quad (2.12)$$

Uma vez que o consumo médio $\bar{Q}(t)$ é desconhecido, é necessário estimá-lo. A estimativa de $\bar{Q}(t)$ foi calculada como sendo o valor esperado de $Q(t)$. Supõe-se que \bar{Q} é constante por troços, variando esparsamente no tempo, e que o seu valor é conhecido para situações normais (quer dizer, em situações em que não há fuga), por exemplo na forma de uma tabela em função do período do ano.

Sabe-se que a entrada do sistema Σ tem média nula. Logo, também $y(t)$ será de média nula e, conseqüentemente,

$$\mathbb{E}[\hat{Q}(t)] = \mathbb{E}[y(t) + \bar{Q}] = \bar{Q}. \quad (2.13)$$

Assim, usa-se o valor médio das observações da saída,

$$\hat{Q}(T) = \frac{1}{T} \sum_{i=1}^T Q(t), \quad (2.14)$$

2. Distribuição em Rede SISO (Caso em Simulação)

como estimador de $\bar{Q}(t)$. É possível provar que este é um estimador óptimo, na medida em que é não-enviesado e a sua variância atinge o limiar inferior de *Cramer Rao*:

$$\mathbb{V}[\hat{Q}] \geq \frac{1}{I_f(\bar{Q})}, \quad (2.15)$$

em que a Informação de Fisher, $I_f(\bar{Q})$, é definida como

$$I_f(\bar{Q}) = -\mathbb{E}\left[\frac{\partial^2 l(Q, \bar{Q})}{\partial \bar{Q}^2}\right], \quad (2.16)$$

e $l(Q, \bar{Q})$ é o logaritmo natural da função de probabilidade [11].

É, no entanto, importante notar que a estimativa \hat{Q} só é válida quando feita em períodos nos quais se sabe não existirem fugas.

Passando o sinal $\hat{y}(t)$ pelo filtro de periodicidade definido na equação (2.2) e somando $\hat{Q}(t)$ obtém-se a estimativa do consumo

$$\hat{Q}(t+m|t) = \begin{cases} \hat{y}(t+m|t) + \hat{Q}(t) & , \text{ se } t \leq 96 \\ \hat{y}(t+m|t) + \hat{Q}(t) + Q(t+m-96) - \hat{Q}(t+m-96) & , \text{ se } t \geq 97 \end{cases}, \quad (2.17)$$

onde se substituiu a previsão do consumo em momentos do passado, $\hat{Q}(t+m-96)$, pelas observações feitas, $Q(t+m-96)$. Uma vez que se definiu o valor médio do caudal como sendo constante, a diferença

$$\hat{Q}(t) - \hat{Q}(t+m-96) \quad (2.18)$$

é desprezável. Finalmente, $\hat{Q}(t+m|t)$ é dado por

$$\hat{Q}(t+m|t) = \begin{cases} \hat{y}(t+m|t) + \hat{Q}(t) & , \text{ se } t \leq 96 \\ \hat{y}(t+m|t) + Q(t+m-96) & , \text{ se } t \geq 97 \end{cases}. \quad (2.19)$$

2.3 Detecção de Fugas

Uma vez conhecidos os consumos de água real e previsto num determinado intervalo de tempo, é possível detectar a existência de fugas na rede de distribuição, por comparação dos dois.

Existem dois tipos de fugas que podem ocorrer: fugas abruptas e fugas progressivas. O primeiro caracteriza-se por um aumento instantâneo do consumo, no momento em que ocorre a falha, e por um caudal constante. O segundo é caracterizado por um aumento progressivo e lento do caudal de água perdido.

Os métodos de detecção a seguir apresentados estão ilustrados sob a forma de pseudo-código no Algoritmo 1.

2.3.1 Detecção de Fugas Abruptas

A medida do desvio entre o consumo esperado numa situação sem fuga, $\hat{Q}(t)$, e o consumo efectivamente observado, $Q(t)$, em cada instante t , é definida como o sendo o erro quadrático,

i.e.,

$$r(t) = (Q(t) - \hat{Q}(t))^2. \quad (2.20)$$

Integrando o erro ao longo dos últimos N instantes de tempo obtem-se um sinal I que será próximo de zero na ausência de fugas e que terá um salto após a ocorrência de uma fuga:

$$I(T) = \frac{1}{N} \sum_{T-N}^T r(t), \quad N = 96. \quad (2.21)$$

O horizonte de integração está intimamente ligado ao desempenho do detector de fugas. Quanto maior for N mais se filtram as discrepâncias entre a predição e o valor real do consumo, tornando o sistema mais sensível à ocorrência de fugas, mas aumentando também o tempo decorrido desde a sua ocorrência até à sua detecção. A escolha deste parâmetro não obedece a nenhuma regra e foi feita por tentativa e erro, observando-se os resultados da simulação de fugas de várias intensidades.

O sinal de alarme, F , que indica a existência de uma fuga será accionado quando $I(t)$ ultrapassar um certo limiar ϵ :

$$F(t) = \begin{cases} 0, & \text{se } I(t) < \epsilon \\ 1, & \text{se } I(t) \geq \epsilon \end{cases}. \quad (2.22)$$

2.3.2 Detecção de Fugas Progressivas

A detecção de rupturas progressivas é feita de forma semelhante, com a particularidade de $r(t)$ ser filtrado por um filtro de mediana, reduzindo a sensibilidade a *outliers*. Substituindo, na equação (2.21), $r(t)$ pelo sinal filtrado obtém-se o sinal que permite activar o alarme que indica a existência de uma ruptura, quando este ultrapassa o limiar, ϵ' :

$$I'(T) = \frac{1}{N} \sum_{T-N}^T R(t), \quad N = 192, \quad (2.23)$$

onde $R(t)$ é o resultado da operação de filtragem de $r(t)$. Uma vez que se pretende que o sistema seja sensível a fugas de intensidade lentamente crescente no tempo, o horizonte de integração tem de ser alargado em relação ao caso anterior. Assim, escolheu-se integrar o resíduo numa janela de dois dias, correspondente a 192 amostras.

Uma vez que este tipo de falhas se caracteriza por ser de intensidade crescente, a sua detecção implica a utilização de um horizonte de integração superior ao usado em 2.21. Novamente, a escolha de N influencia o tempo decorrido desde a ocorrência até à detecção da fuga: quanto maior o seu valor, maior o tempo decorrido. Para valores de N demasiado pequenos deixa de ser possível fazer a detecção deste tipo de rupturas uma vez que não é tida em consideração a crescente discrepância entre a precisão e o consumo real.

O sinal de alarme é então dado por

$$F'(t) = \begin{cases} 0, & \text{se } I'(t) < \epsilon' \\ 1, & \text{se } I'(t) \geq \epsilon' \end{cases}. \quad (2.24)$$

2. Distribuição em Rede SISO (Caso em Simulação)

Algoritmo 1 Detecção de Fugas em Simulação

```
1:  $\hat{Q}(T) \leftarrow \frac{1}{T} \sum_{i=1}^T Q(t)$ 
2:  $Q_c(t) \leftarrow Q(t) - \hat{Q}(t)$ 
3:  $y_{rec}(t) \leftarrow Q_c(t) - Q_c(t - 96)$ 
4:  $\hat{y}(t + m) \leftarrow \alpha y_{rec}(t - m + 1) + \beta y_{rec}(t - m)$ 
5:  $\hat{Q}(t + m) \leftarrow \hat{y}(t + m) + \hat{Q}(t) + Q(t + m - 96) - \hat{Q}(t + m - 96)$ 
6:  $r(t) \leftarrow [Q(t) - \hat{Q}(t)]^2$ 
7:  $R(t) \leftarrow r(t)$  filtrado por filtro de mediana
8:  $I(T) \leftarrow \frac{1}{N} \sum_{T-N}^T r(t), N = 96$ 
9:  $I'(T) \leftarrow \frac{1}{N} \sum_{T-N}^T R(t), N = 192$ 
10: if  $I(T) \geq \epsilon$  then
11:    $F(T) \leftarrow 1$ 
12: else
13:    $F(T) \leftarrow 0$ 
14:
15: end if
16: if  $I'(T) \geq \epsilon$  then
17:    $F'(T) \leftarrow 1$ 
18: else
19:    $F'(T) \leftarrow 0$ 
20:
21: end if
```

2.4 Simulação

Os métodos descritos em 2.2 e 2.3 foram implementados usando o *Matlab*[®]. Usou-se a função *randn* para gerar o vector de entrada do sistema Σ , $e(t)$.

As funções de transferência de Σ e do filtro 2.2 foram representadas por meio dos coeficientes dos seus polinómios. Todos os sinais foram calculados analiticamente.

As simulações foram efectuadas para um período de 10 dias e, nos casos em que existem, as fugas foram mantidas até ao final do tempo de simulação.

Começou por se testar o preditor do consumo, através de simulação sem fuga, tendo-se posteriormente simulado fugas abruptas e progressivas por forma a testar os detectores. Os resultados apresentados em 2.4.2 e 2.4.3 não têm como objectivo avaliar o desempenho dos detectores desenvolvidos, o que é feito em 2.5, mas sim demonstrar os seus princípios e ilustrar as suas limitações.

Com o intuito de facilitar a visualização optou-se por, nalguns gráficos, não abranger todo o período de simulação.

2.4.1 Simulação sem Fuga

Inicialmente o sistema foi simulado sem a presença de perturbações por forma a validar o preditor e a calibrar os limiares de decisão.

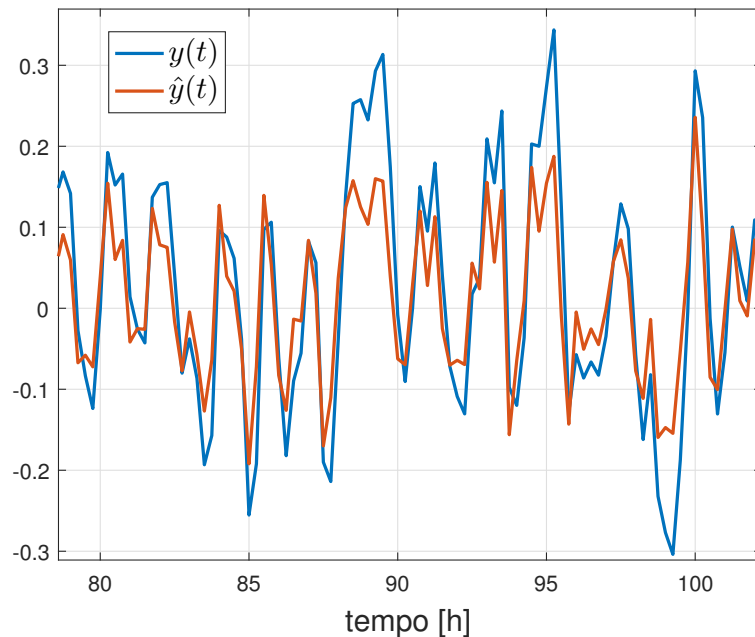


Figura 2.3: Saída do sistema Σ . Sinal real (azul) e estimativa (laranja).

Na figura 2.3 estão representadas as previsões um passo à frente e os valores reais para a saída do sistema Σ . A previsão segue de perto o sinal real, apesar não conseguir atingir os

2. Distribuição em Rede SISO (Caso em Simulação)

valores mais extremos.

O consumo, $Q(t)$, está representado na figura 2.4. Pode-se verificar que a previsão está bastante próxima do consumo real, pelo que o erro será pequeno.

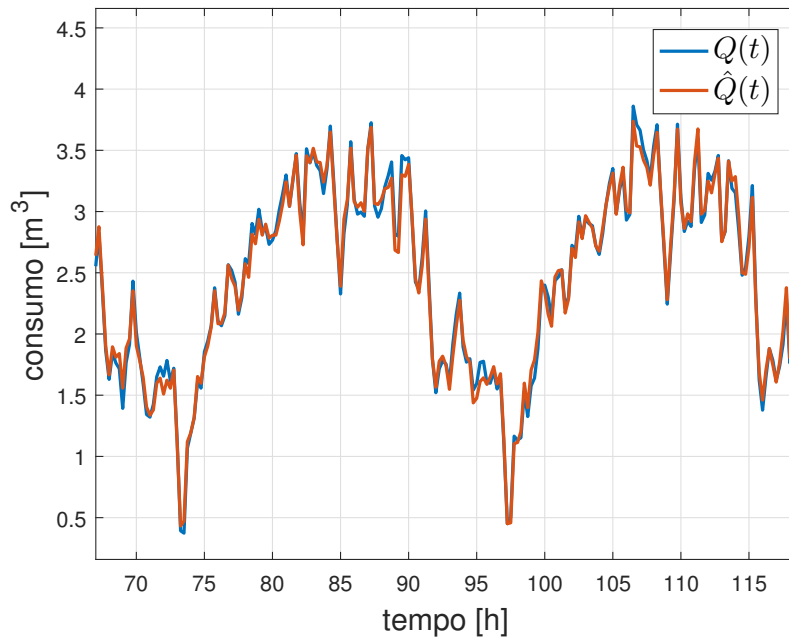


Figura 2.4: Consumo simulado. Sinal real (azul) e estimativa (laranja).

Na figura 2.5 está representado o real valor médio do consumo, \bar{Q} , e a sua estimativa, $\hat{\bar{Q}}$. De notar que o desvio do valor real é pequeno, tal como se previa, pelo que o erro introduzido pela estimativa do valor médio será pouco significativo.

Na figura 2.6 está representado o integral do erro, $I(t)$, numa janela de 96 amostras conforme a equação (2.21). Tal como se esperava, o seu valor é muito próximo de zero, o que corrobora as constatações anteriores de que a estimativa do valor médio introduz um erro pouco significativo e que a previsão do consumo é muito próxima do seu valor real.

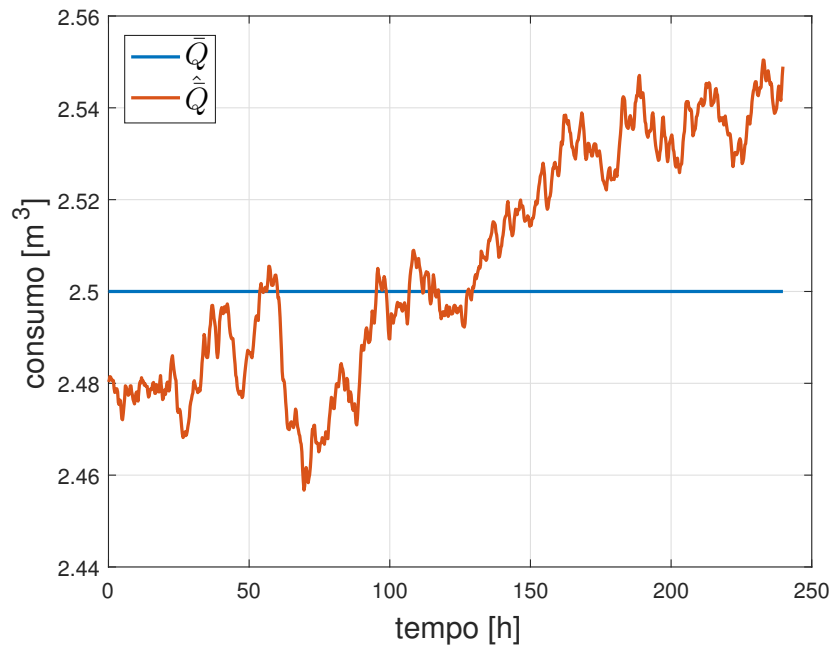
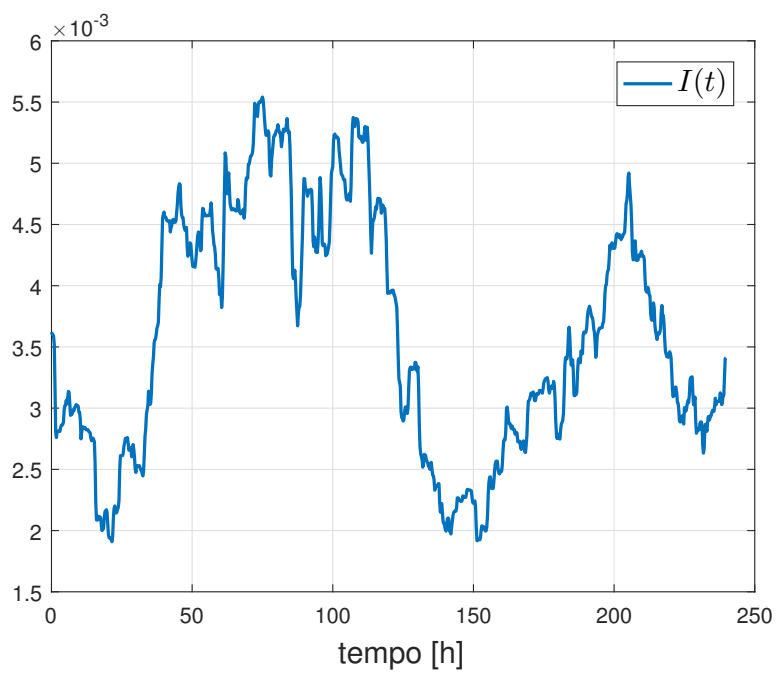


Figura 2.5: Consumo Médio real (azul) e estimado (laranja)

Figura 2.6: Integral do erro de predição 1 passo à frente, $I(t)$

2.4.2 Simulação de Fuga Abrupta

A simulação sem fugas permitiu inferir qual o limiar de detecção, ϵ , a usar. Para ilustrar o método de detecção de uma fuga abrupta, simulou-se um período de 10 dias com uma fuga a surgir ao quinto dia. O valor da fuga foi definido como sendo 20% do valor médio do consumo.

Novamente, representa-se a estimativa e o valor real do consumo, na figura 2.7.

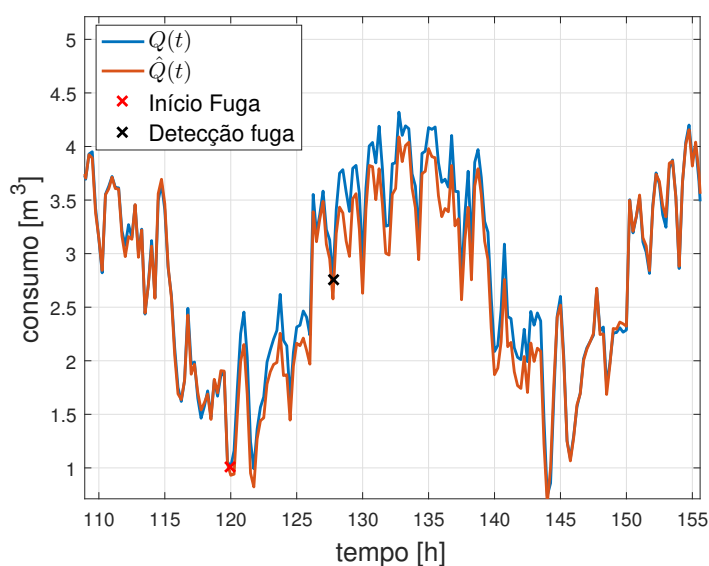


Figura 2.7: Consumo Instantâneo real (azul) e estimado (laranja), usando \hat{Q}

Pode-se ver que nas 24 horas após o início da fuga (marcado com uma cruz vermelha) a estimativa segue a tendência dos acontecimentos reais, *i.e.*, o consumo previsto aumenta a partir do dia seguinte àquele em que se dá a fuga. Isto deve-se ao facto de se ter assumido que o consumo é aproximadamente periódico, com período de um dia, o que é imposto pelo filtro (2.2). Tal como se pode verificar pela equação (2.19), a estimativa do consumo tem em consideração o consumo no dia anterior, o que leva a que a previsão siga a tendência de aumento do consumo causada pela fuga.

Apesar de a fuga ter sido detectada mesmo sob as condições anteriores, este não é um comportamento desejável para o sistema.

Na secção 3.4 propõe-se um método para mitigar este problema, obtendo-se então um previsão abaixo do consumo real nos casos em que há fugas, tal como seria expectável, mesmo após as primeiras 24 horas após o início da fuga.

O integral do erro de predição 1 passo à frente, a partir do qual se infere a existência de fugas, está representado na figura 2.8. Tal como anteriormente o sinal é próximo de zero até ao início da fuga (assinalado com uma cruz vermelha), momento a partir do qual o seu valor começa a aumentar devido à discrepância entre o consumo real e estimado. É possível notar que a partir de um determinado instante (marcado com uma cruz cor de rosa) o seu valor volta a estar abaixo do

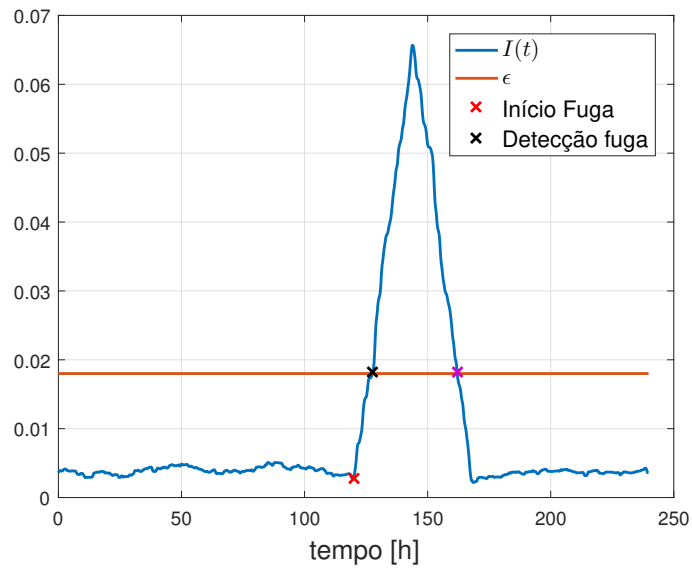


Figura 2.8: Integral do erro de previsão 1 passo à frente com fuga abrupta

limiar ϵ e progressivamente aproxima-se de zero. Este fenómeno deve-se ao facto anteriormente explicado, de que a estimativa segue a tendência do consumo real.

Nas figuras 2.9 e 2.10 estão representados, respectivamente, o consumo médio (real e estimado) e o sinal de alarme, $F(t)$.

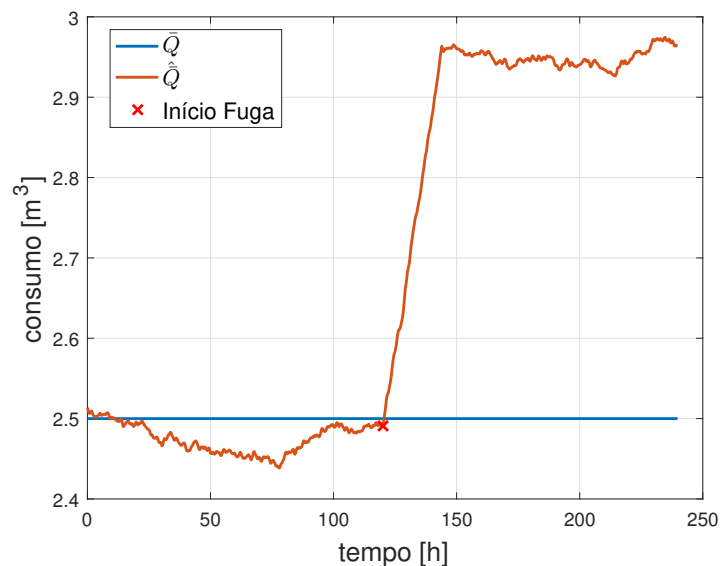


Figura 2.9: Consumo Médio real e estimado em situação de fuga abrupta

Também a estimativa do consumo médio é afectada pela fuga. No entanto, os seus efeitos não se propagam à estimativa do consumo $Q(t)$ já que, em (2.18), se optou por desprezar a diferença entre os valores médios.

2. Distribuição em Rede SISO (Caso em Simulação)

O sinal de alarme gerado pode ser observado na figura 2.10. A partir do momento em que a fuga é detectada (marcado com uma cruz preta) o sinal tem o valor 1. Este valor mantém-se até que o integral do erro de predição seja novamente inferior ao limiar de detecção.

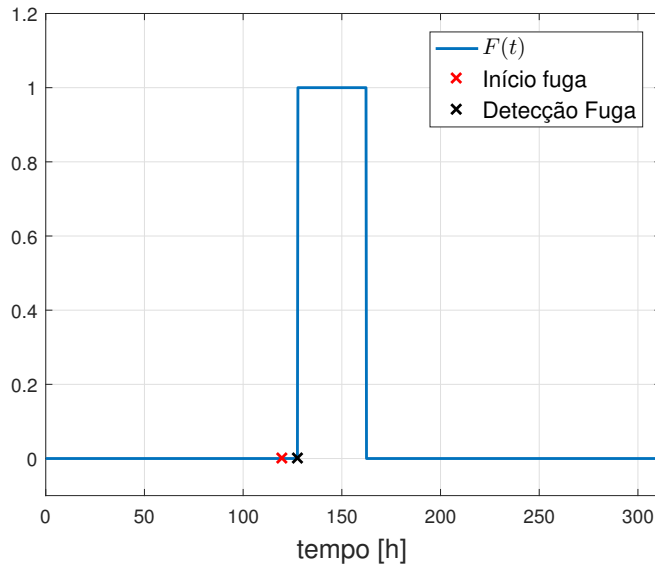


Figura 2.10: Sinal de Alarme para fuga abrupta, $F(t)$

2.4.3 Simulação de Fuga Progressiva

O volume de água perdido aquando da existência de uma ruptura progressiva foi definido como uma progressão aritmética, em que a cada quarto de hora o volume aumenta da razão da progressão. A razão da progressão foi definida como sendo uma percentagem do consumo médio.

Dado que o volume de água perdido resultante de rupturas progressivas cresce lentamente, foi necessário simular o seu início mais cedo. Assim, a ruptura passou a ter início ao terceiro dia de simulação. Nesta simulação volume máximo que a fuga pode tomar, ou seja, o último elemento da progressão, é 20% do consumo médio, o que se traduz numa taxa de crescimento de 0.1042 %/h.

Nas figuras 2.11 e 2.12 estão representados, respectivamente, o consumo (real e estimado) e o integral – $I'(t)$ – do erro de predição, $R(t)$.

Pode-se ver que a fuga foi detectada, já que $I'(t)$ ultrapassa o limiar pré-definido, ϵ' . Na figura 2.13 está representado o sinal de alarme $F'(t)$ que, conforme a equação 2.24, toma o valor 1 quando $I'(t)$ está acima de ϵ' e o valor 0 caso contrário.

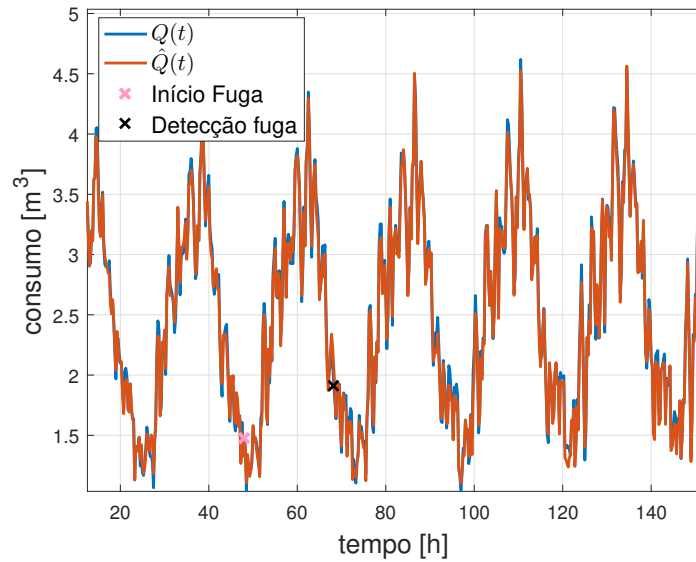
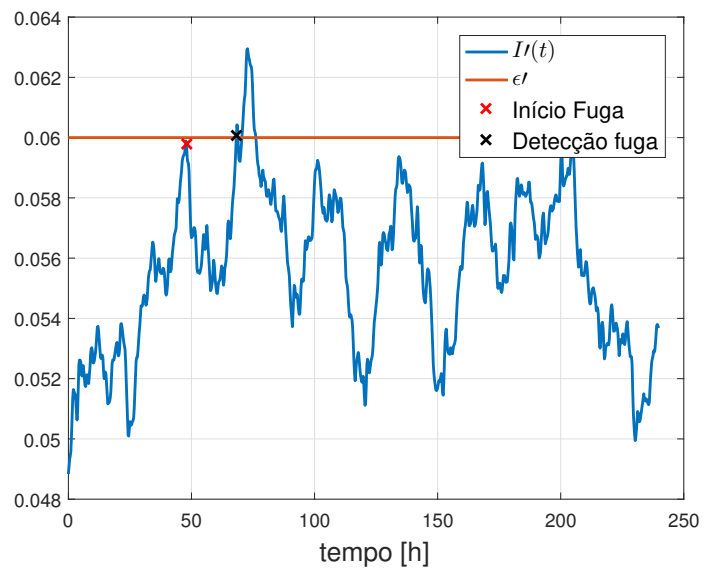


Figura 2.11: Consumo simulado. Sinal real (azul) e estimativa (laranja)

Figura 2.12: Integral do erro de predição após filtragem, $I'(t)$

2. Distribuição em Rede SISO (Caso em Simulação)

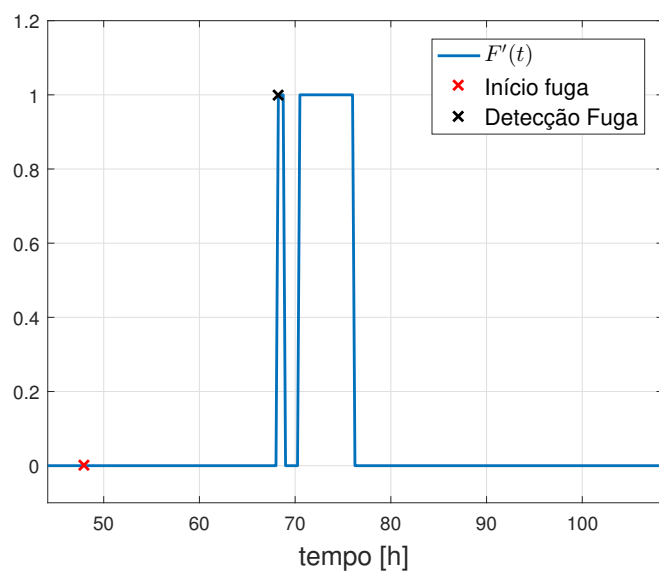


Figura 2.13: Sinal de Alarme para fuga progressiva, $F'(t)$

2.5 Simulações de Monte Carlo

De forma a quantificar o desempenho de cada um dos detectores de fugas implementados fez-se um conjunto de simulações de Monte Carlo, tendo-se variado os limiares de decisão, a intensidade da fuga, no caso de fugas abruptas, e a taxa de crescimento da fuga, no caso de fugas progressivas.

Escolheu-se como valor médio do consumo um valor próximo do valor médio dos dados reais: $\bar{Q} = 2.4 m^3/15min$. Foram feitas 10 mil simulações para cada situação.

2.5.1 Fuga Abrupta

Cada simulação foi feita para um período de 30 dias e com a fuga com a duração de três. É de notar que o período da fuga apenas totaliza 10% do tempo total de simulação. Assim, este será o valor máximo que os Verdadeiros Positivos (TP) e Falsos Negativos (FN) podem atingir. Da mesma forma, o máximo valor que Verdadeiros Negativos (TN) e Falsos Positivos (FP) pode tomar é 90%.

Na tabela 2.1 apresentam-se os resultados do desempenho do detector para o limiar óptimo, $\epsilon^* = 0.16$, o qual se caracteriza pela melhor relação entre a taxa de detecção, o tempo de atraso na detecção de fugas, e o número de falsos alarmes. As tabelas para todos os limiares podem ser consultadas no Anexo A.

F [%]	F [$m^3/15min$]	Atraso	TP [%]	FP [%]	TN [%]	FN [%]
5	0.12	18.46 [h]	0.38	0.02	89.98	9.62
10	0.24	9.28 [h]	3.81	0.01	89.99	6.19
15	0.36	4.29 [h]	5.38	0.02	89.98	4.62
20	0.48	2.19 [h]	6	0.01	89.99	4
25	0.6	1.18 [h]	6.28	0.02	89.98	3.72
30	0.72	37.66 [min]	6.43	0.01	89.90	3.57
35	0.84	19.76 [min]	6.52	0.02	89.98	3.48
40	0.96	9.85 [min]	6.57	0.01	89.99	3.43
45	1.08	5.35 [min]	6.60	0.02	89.98	3.40
50	1.2	2.68 [min]	6.62	0.02	89.98	3.40
55	1.32	1.66 [min]	6.64	0.01	89.99	3.36
60	1.44	0.82 [min]	6.65	0.02	89.98	3.35

Tabela 2.1: Resultados das Simulações de Monte Carlo para ϵ^*

Na tabela 2.2 estão registadas as percentagens de sucesso e de insucesso médias para o mesmo limiar. Considera-se como sucesso uma detecção correcta de fuga (*True Positive*) ou uma não detecção quando esta não existe (*True Negative*). Um insucesso ocorre quando a fuga não é detectada (*False Negative*) ou quando é dado um falso alarme (*False Positive*).

Tal como seria expectável, quanto maior for a intensidade da fuga, maior é a taxa de sucesso já que, para um determinado limiar, o número de falsos alarmes não se altera e o número de

2. Distribuição em Rede SISO (Caso em Simulação)

detecções acertadas cresce, dado a disparidade entre o consumo real e estimado ser maior, tornando a fuga mais evidente.

F [%]	Sucesso [%]	Insucesso [%]
5	90.37	9.63
10	93.80	6.20
15	95.63	4.64
20	95.98	4.02
25	96.26	3.74
30	96.42	3.58
35	96.50	3.50
40	96.56	3.44
45	96.59	3.41
50	96.61	3.39
55	96.62	3.38
60	96.63	3.37

Tabela 2.2: Sucessos e Insucessos na detecção resultante das Simulações de Monte Carlo para ϵ^*

Apresentam-se ainda as curvas Receiver Operating Characteristic (ROC), na figura 2.14, para cada intensidade de falha. A curva ROC é a curva que relaciona TP e FP para cada limiar de detecção ϵ . Cada curva corresponde a um nível de intensidade de fuga.

Quanto mais baixo for o limiar de detecção, maior será a ocorrência de TP. Em contrapartida, também as ocorrências de FP aumentam. Ao estabelecer-se um limiar mais elevado, o número de ocorrências de FP diminui, mas a taxa de detecção de fugas também diminui, podendo chegar ao ponto de nenhuma fuga ser detectada.

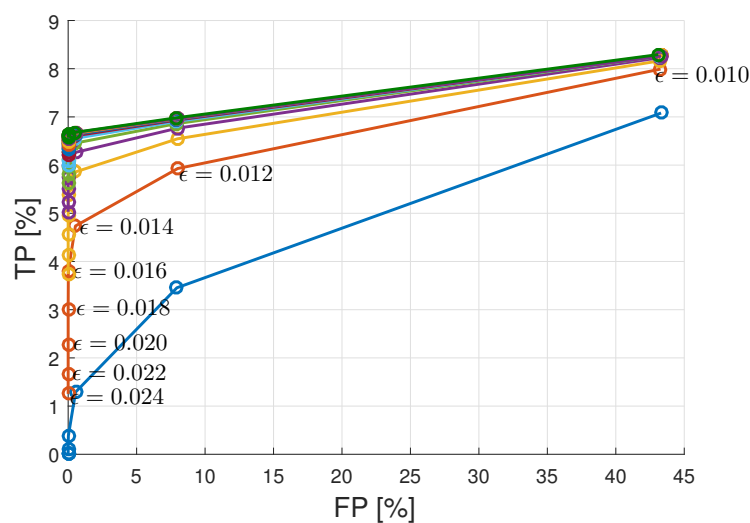


Figura 2.14: Curvas ROC para fuga abrupta

Por outro lado, o limiar de decisão também tem impacto no atraso de detecção. Quanto mais baixo for ϵ , menor será o tempo decorrido entre o início da fuga e a sua detecção.

Na figura 2.15 pode-se ver a variação do atraso de detecção em função do limiar ϵ , para cada intensidade de fuga. Naturalmente, para fugas de maior intensidade o atraso de detecção é inferior.

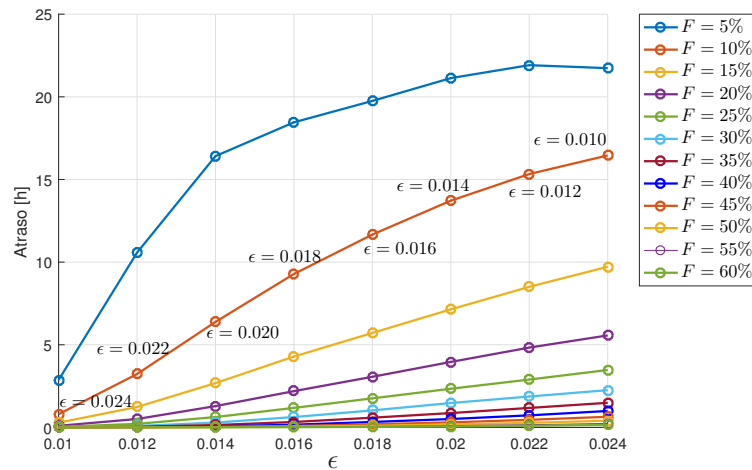


Figura 2.15: Atraso na detecção em função do limiar

2.5.2 Fuga Progressiva

Cada simulação foi feita para um período de 30 dias, com a fuga a ocorrer nos últimos 10 dias da simulação. Variou-se o limiar de decisão e a taxa de crescimento da fuga por hora, R . A taxa de crescimento foi definida em função do consumo médio, o qual não foi alterado. Começou por se testar uma taxa de crescimento de $0.5\%/h$, tendo-se incrementado em intervalos de 0.5 até 4%.

Uma vez que a fuga apenas totaliza um terço do tempo total de simulação, este é o valor máximo que TP e FN podem tomar. Da mesma forma, TN e FP nunca serão superiores a dois terços.

Na tabela 2.3 apresentam-se os resultados das simulações para o limiar óptimo, $\epsilon'^* = 0.056$, que, novamente, se caracteriza pela melhor relação entre a taxa de detecção, o tempo de atraso na detecção de fugas e número de falsos alarmes. Além de se apresentar as taxas de TP, FP, TN e FN, é ainda apresentado o volume de água perdido desde o início da fuga e até à sua detecção, Q_{fuga} .

Apesar destas fugas se caracterizarem por um crescimento lento do caudal perdido, o facto de correrem durante muito tempo faz com que o volume de água perdido seja muito superior àquele das fugas abruptas. Enquanto que no pior caso de uma ruptura abrupta o volume de água perdido era ligeiramente inferior a $9 m^3$, o melhor caso na detecção de uma fuga progressiva representa uma perda superior a esses mesmos $9 m^3$. É, portanto, do maior interesse que a percentagem de fugas deste tipo não detectadas seja tão baixa quanto possível.

2. Distribuição em Rede SISO (Caso em Simulação)

R [%/h]	Q_{fuga} [m^3]	Atraso [h]	TP [%]	FP [%]	TN [%]	FN [%]
0.5	21.01	81.76	2.97	5.33	61.34	30.36
1.0	35.21	79.29	3.80	5.41	61.25	29.53
1.5	35.65	72.03	5.76	5.33	61.34	27.58
2.0	28.89	62.00	9.05	5.29	61.38	24.28
2.5	21.00	50.51	13.83	5.39	61.27	19.50
3.0	15.02	41.24	19.53	5.36	61.31	13.80
3.5	11.66	34.50	24.47	5.31	61.35	8.86
4.0	9.85	30.23	27.51	5.32	61.34	5.82

Tabela 2.3: Resultados das Simulações de Monte Carlo para ϵ'^*

Na figura 2.16 apresentam-se as curvas ROC para este detector, onde cada curva corresponde a uma taxa R .

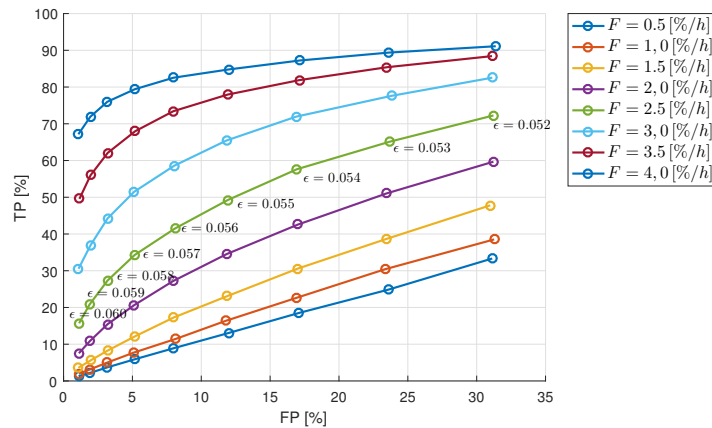


Figura 2.16: Curvas ROC para fuga progressiva

Na figura 2.17 está representado o atraso na detecção da fuga em função do limiar ϵ' . De uma forma geral, quanto mais elevado for o limiar de decisão, maior será o o atraso. Por outro lado, quanto maior for o limiar de decisão, menor será o número de falsos alarmes, à custa de um decréscimo da taxa de detecção correcta, como se pode verificar pelas curvas da figura 2.16.

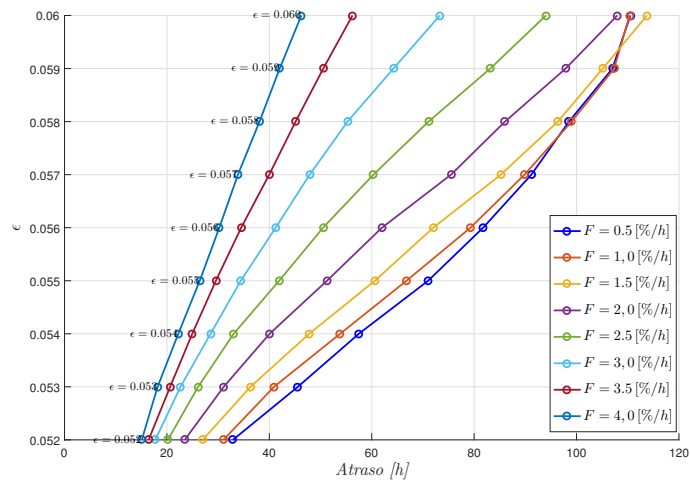


Figura 2.17: Atraso na detecção em função do limiar

2. Distribuição em Rede SISO (Caso em Simulação)

3

Distribuição em Rede SISO (Caso Real)

Contents

3.1 Falha na ocorrência dos dados	32
3.2 Identificação do Modelo	32
3.3 Validação do Modelo	33
3.4 Predição	34
3.5 Fins-de-Semana e Feriados	36
3.6 Resultados	37
3.7 Simulação de Fugas	39

3. Distribuição em Rede SISO (Caso Real)

Neste capítulo os dados usados são os do consumo de água no ponto AJA (Avenida António José de Almeida) entre 15 de Setembro de 2015 e 8 de Março de 2016.

O abastecimento de água ao Instituto Superior Técnico é feito por três pontos distintos. Enquanto que o Pavilhão de Civil e o Pavilhão de Matemática são abastecidos independentemente, todos os outros pavilhões são abastecidos por uma conduta que chega ao IST pela entrada contígua à AJA.

3.1 Falha na ocorrência dos dados

Tal como explicado previamente, os dados do consumo são fornecidos pela EPAL a cada quinze minutos, ou seja, cada leitura indica o consumo total de água no último quarto de hora.

Por vezes acontecem problemas que impedem o correcto envio dos registos para os servidores o que resulta na ausência de informação para esse período que pode ser total ou parcial. A ausência de informação é parcial quando no envio seguinte de dados está o consumo agregado de todos os instantes anteriores, ou seja, é conhecido o consumo total no período de "apagão" mas não a sua distribuição no tempo. A ausência de informação é total quando não há nenhuma informação acerca do consumo durante o período de "apagão".

Uma vez que o objectivo principal deste trabalho é explorar a previsão do consumo para detecção de fugas, foi usada uma abordagem simplificada na reconstrução dos dados: tendo em conta que o consumo diário é aproximadamente periódico, quando não há informação num dado instante usa-se a informação desse mesmo instante no dia que o antecedeu com as mesmas características. Por exemplo, se num determinado dia – suponhamos, uma sexta-feira – não houver registo às 14h15m utiliza-se a leitura das 14h15m de quinta feira. No caso de o dia em questão ser um Sábado ou Domingo, usa-se o valor do Sábado ou Domingo anterior.

3.2 Identificação do Modelo

De forma a poder detectar fugas nos dados reais é necessário que haja um modelo que dê a previsão do consumo em cada instante. No caso dos dados reais o sistema Σ é desconhecido e terá que ser identificado.

Tal como era expectável, após análise dos registos do consumo no período indicado, foi possível verificar que o consumo não tem o mesmo padrão em todos os momentos do ano. Assim, dividiu-se o ano ¹ em vários períodos, como mostra a tabela 3.1, e para cada período foi calculado o valor médio do consumo e identificado o modelo correspondente.

Foi ainda possível verificar que o consumo ao fim-de-semana é aproximadamente independente do momento ano, pelo que se identificou um modelo para para Sábado e um modelo para

¹ Não se considerou um ano completo, mas apenas o período dos dados disponíveis.

Domingo os quais, a par de \bar{Q} para cada um destes dias, serão únicos, independentemente do período do ano.

Início	Fim	Descrição
1 Set 2015	6 Set 2015	Verão
7 Set 2015	27 Set 2015	"Regresso às aulas"
28 Set 2015	20 Dez 2015	Período Lectivo
21 Dez 2015	3 Jan 2016	Festividades
4 Jan 2016	31 Jan 2016	Época de Exames
1 Fev 2016	14 Fev 2015	Pausa entre Semestres
15 Fev 2016	8 Mar 2016	Segundo Semestre

Tabela 3.1: Divisão dos vários momentos do ano

De forma a identificar o sistema Σ foi necessário aplicar o filtro inverso de periodicidade, *i.e.* a função inversa de (2.2), aos dados do consumo de forma a obter o sinal $y(t)$.

Os modelos foram identificados usando a função *arma* do *Matlab* através do sinal y , com o numerador e o denominador de segundo grau. Os coeficiente dos polinómios obtidos para cada modelo podem ser consultados no Anexo C.

Por forma a ultrapassar o problema identificado em 2.4.2 os valores médios calculados para cada período são os valores usados para $\hat{Q}(t)$ nas equações (2.11) e (2.19).

3.3 Validação do Modelo

De forma a validar que os modelos identificados representam fielmente o sistema real analisou-se o resíduo da saída dos sistemas Σ , $y(t)$. O resíduo é a diferença entre os sinais real – $y_{real}(t)$ –, obtido após aplicar o filtro inverso de periodicidade aos dados do consumo, e o sinal obtido através da excitação do modelo identificado – $y_{ID}(t)$. O modelo identificado foi excitado com ruído branco.

O sinal $y_{real}(t)$ é obtido através dos dados do consumo da seguinte forma:

$$y_{real}(t) = \begin{cases} Q(t) - Q_0(t) & , \text{ se } t \leq 96 \\ Q(t) - Q(t - 96) & , \text{ se } t \geq 97 \end{cases} \quad (3.1)$$

O resíduo é então dado por

$$\delta(t) = y_{real}(t) - y_{ID}(t). \quad (3.2)$$

O modelo identificado representa fielmente o modelo real se a autocorrelação do resíduo for similar à autocorrelação de ruído branco. Assim, calcularam-se os coeficientes de autocorrelação de $\delta(t)$ para cada modelo e calculou-se o intervalo de confiança a 99%. Na figura 3.1 estão representados os coeficientes de autocorrelação para vários atrasos (em cima) e os consumos real e calculado a partir do modelo identificado (em baixo) para o período em que o consumo é mais regular, definido como "período lectivo" na tabela 3.1.

3. Distribuição em Rede SISO (Caso Real)

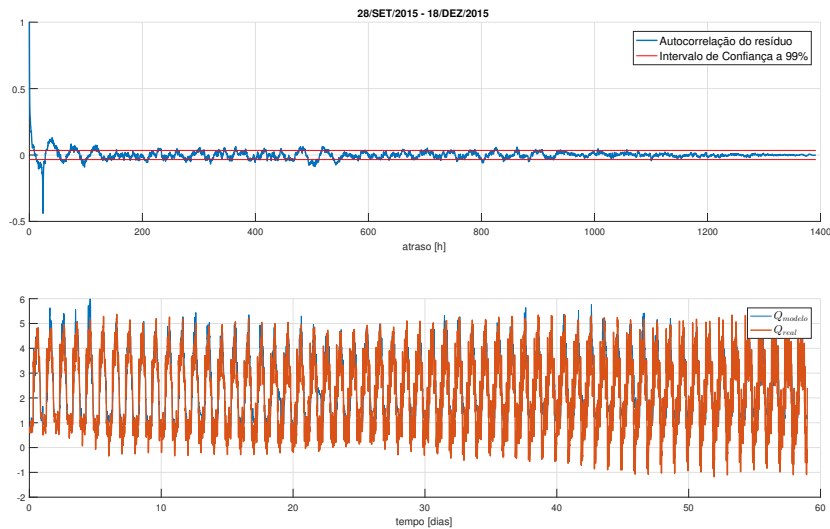


Figura 3.1: Autocorrelação do resíduo (em cima) e consumos real e obtido através do modelo identificado (em baixo)

No anexo D podem ser consultadas as figuras para os restantes períodos.

Pode-se ver que, em geral, os coeficientes de autocorrelação do resíduo estão dentro do intervalo de confiança a 99%, pelo que o modelo identificado constitui uma representação fiel da realidade, o que se pode também concluir por inspeção visual dos consumos.

No entanto, devido à escassez de dados ou à grande irregularidade dos consumos nalguns períodos, nem todos os modelos identificados são tão fiéis à realidade quanto o anteriormente mencionado, sendo os mais críticos os correspondentes a períodos de pausa; férias de Verão e interregno entre 1^o e 2^o semestres. Apesar de haver margem para futuramente melhorar alguns dos modelos, através da obtenção de mais dados para os períodos em causa, os modelos obtidos são satisfatórios para o desenvolvimento desta dissertação.

3.4 Predição

No caso apresentado em simulação, na secção 2, o sistema Σ era um caso particular do modelo ARMA, em que o polinómio do numerador é

$$C(z^{-1}) = 1. \quad (3.3)$$

O modelo identificado para os dados reais é um modelo ARMA puro e o sistema é da forma

$$\frac{C(z)}{A(z)} = \frac{c_2 z^2 + c_1 z + c_0}{a_2 z^2 + a_1 z + a_0}. \quad (3.4)$$

Multiplicando por z^{-2} tem-se

$$\frac{C^*(z^{-1})}{A^*(z^{-1})} = \frac{c_2 + c_1 z^{-1} + c_0 z^{-2}}{a_2 + a_1 z^{-1} + a_0 z^{-2}}. \quad (3.5)$$

Assim, o preditor óptimo para $y(t)$ um passo à frente será da forma

$$\hat{y}(t+1|t) = \alpha(q^{-1})y_{rec}(t) + \beta(q^{-1})y_{rec}(t-1) + \gamma(q^{-1})y_{rec}(t-2), \quad (3.6)$$

em que os coeficientes α , β e γ são obtidos através da divisão longa dos polinómios de Σ .

A recuperação de $y(t)$ é feita de forma análoga à usada na simulação mas, para melhorar o desempenho, as condições iniciais usadas são observações reais de um dia típico dentro do período vigente:

$$y_{rec}(t) = \begin{cases} Q(t) - Q_0(t) & , \text{ se } t \leq 96 \\ Q(t) - \bar{Q}(t) - Q(t-96) + \bar{Q}(t-96) & , \text{ se } t \geq 97 \end{cases}, \quad (3.7)$$

onde $Q_0(t)$ são as condições iniciais. Uma vez que o consumo médio foi tabelado para cada período do ano, $\bar{Q}(t)$ é constante dentro de cada período.

De forma a corrigir a falha apontada na secção 2.4.2 propõe-se que as observações passadas sejam centradas em relação ao consumo médio nas 24 horas que as antecedem. Assim, $Q(t-96) - \bar{Q}(t-96)$ é substituído na equação (2.11) por:

$$Q(t-96) - \frac{1}{96} \sum_{t-96*2}^{t-97} Q(\tau). \quad (3.8)$$

Desta forma o valor médio das observações passadas não influenciará a previsão até ao terceiro dia de fuga permitindo que, numa situação de fuga, a previsão não siga a tendência do aumento do consumo médio por mais tempo, permitindo um horizonte temporal para a sua detecção mais alargado. O sinal $y(t)$ é então recuperado a partir de $Q(t)$ através de

$$y_{rec}(t) = \begin{cases} Q(t) - Q_0(t) & , \text{ se } t \leq 96 \\ Q(t) - \bar{Q} - [Q(t-96) - \frac{1}{96} \sum_{t-96}^{t-1} Q(\tau)] & , \text{ se } 97 \leq t \leq 192 \\ Q(t) - \bar{Q} - [Q(t-96) - \frac{1}{96} \sum_{t-192}^{t-97} Q(\tau)] & , \text{ se } t \geq 193 \end{cases}. \quad (3.9)$$

Uma vez recuperado y é feita a previsão um passo à frente, usando a equação (3.6). Dado que para $t \leq 192$ não há 96 observações anteriores a $Q(t-95)$ usa-se as 96 observações que precedem $Q(t)$ para calcular o valor médio a subtrair. A previsão do consumo é obtida a partir de \hat{y} através de

$$\hat{Q}(t+1|t) = \begin{cases} \hat{y}(t+1|t) + Q_0(t+1) & , \text{ se } t \leq 96 \\ \hat{y}(t+1|t) + \bar{Q} + Q(t-95) - \frac{1}{96} \sum_{t-95}^t Q(\tau) & , \text{ se } 97 \leq t \leq 192 \\ \hat{y}(t+1|t) + \bar{Q} + Q(t-95) - \frac{1}{96} \sum_{t-191}^{t-96} Q(\tau) & , \text{ se } t \geq 193 \end{cases}. \quad (3.10)$$

É necessário no entanto ter presente que esta modificação tem um custo associado: o método de predição torna-se menos flexível. Se existir uma sequência de dias dentro de um dos períodos durante a qual o consumo médio é aproximadamente constante mas significativamente diferente do valor tabelado, anteriormente esta discrepância seria corrigida a partir do segundo dia. Com a

3. Distribuição em Rede SISO (Caso Real)

modificação anteriormente sugerida perde-se esta capacidade de adaptação a flutuações dentro de um período.

Na figura 3.2 pode-se ver uma comparação dos dois métodos na situação anteriormente descrita.

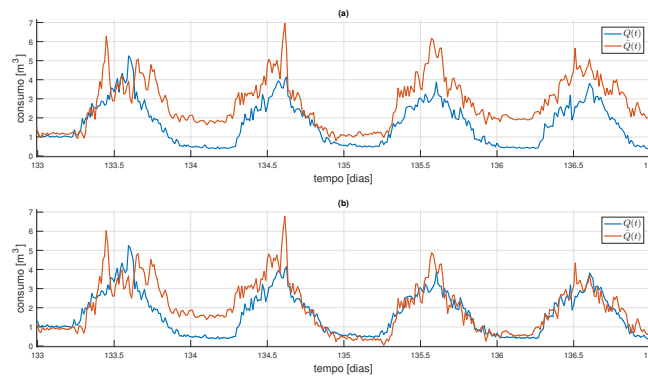


Figura 3.2:

- (a) Consumo Real e Previsto usando a equação (3.10)
- (b) Consumo Real e Previsto usando a equação (2.11)

Em cima estão representados os consumos reais e estimados após a implementação da alteração e em baixo estão representados os mesmos sinais antes da alteração. É notória a diferença entre os dois casos, em particular a partir do dia 134 onde o consumo estimado acompanha de perto de o consumo real na figura (b), enquanto que na figura (a) a previsão está sempre acima do real, devido à discrepância do valor médio real face ao tabelado.

3.5 Fins-de-Semana e Feriados

Tal como dito anteriormente, os modelos e os valores médios usados para prever o consumo aos Sábados e Domingos são independentes do período do ano. Por esse motivo, as condições iniciais usadas na equação (3.7) não são observações reais de um dia típico dentro do período do ano vigente mas sim observações reais do consumo em cada um destes dias.

Relativamente aos dias feriado, foi possível verificar que o consumo nestes dias é, naturalmente, bastante reduzido comparativamente a um dia útil, não fazendo sentido que as condições iniciais usadas para estes dias sejam observações de um dia útil. Assim, para estes dias as condições iniciais usadas na equação (3.7) são observações de um qualquer dia feriado para o primeiro feriado e, para os seguintes, são as observações do feriado anterior. No entanto, ao contrário do que acontece com os fins-de-semana, o modelo usado em cada período adequa-se razoavelmente a estes dias especiais, pelo que o modelo usado para o sistema Σ será o modelo do período vigente.

3.6 Resultados

Começa por se apresentar os resultados da predição do consumo para uma semana típica (segunda feira a Domingo), na figura 3.3. É possível ver que a previsão para Sábado está longe de ser perfeita, o que se deve à irregularidade do consumo de água nestes dias, mais ainda do que aos Domingos.

Pode ainda ver-se o efeito do filtro de periodicidade no dia de quinta feira. No final do dia anterior o consumo teve um comportamento ímpar que o filtro tentou replicar, levando que a predição não acompanhe o consumo real tão de perto como vinha acontecendo.

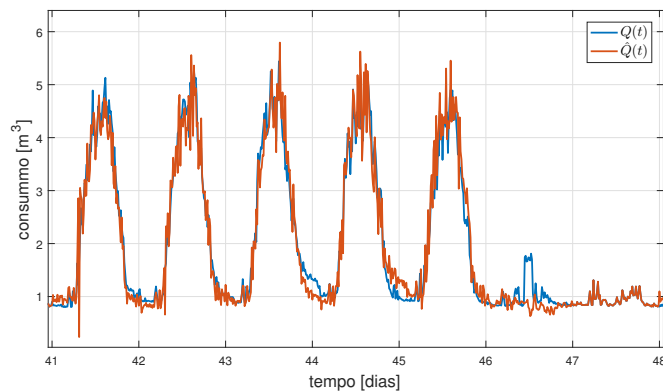


Figura 3.3: Consumo de uma semana típica (Segunda-feira a Domingo)

Na figura 3.4 apresenta-se o resultado da predição para uma semana atípica, onde ocorreu o feriado de 8 de Dezembro. A previsão para o dia 8 (3ª feira) não é feita com o mesmo rigor que nos dias úteis. Tal deve-se, essencialmente, às condições iniciais usadas. Uma vez que não foi possível ter acesso a dados de consumo de outros "8 de Dezembro" as condições iniciais usadas são observações de um outro dia feriado, degradando assim a qualidade da previsão feita.

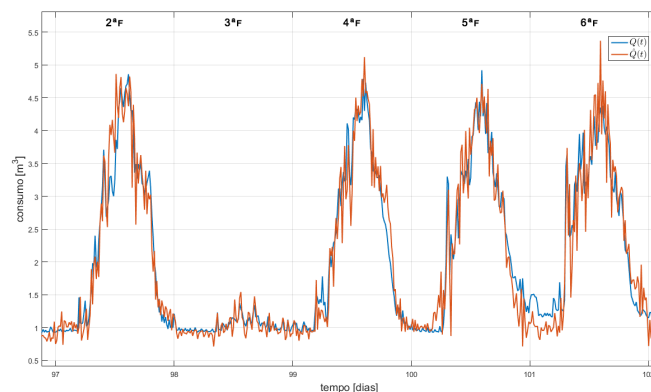


Figura 3.4: Consumo real e previsto de uma semana com feriado (8 de Dezembro)

3. Distribuição em Rede SISO (Caso Real)

Na figura 3.5 apresenta-se a previsão do consumo para todos os feriados dentro do período que se está a tratar, definido no capítulo 3.

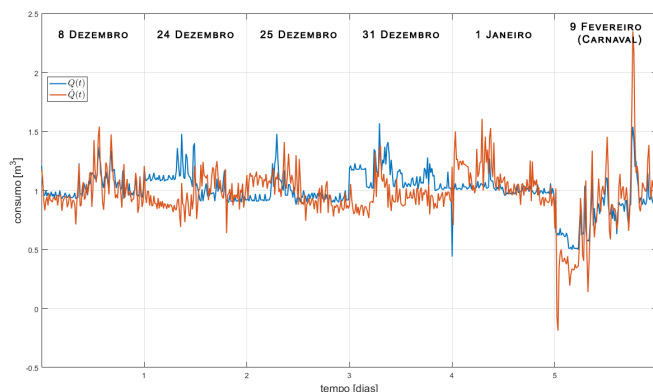


Figura 3.5: Consumo real e previsto dos dias feriado

Se, eventualmente, existisse um banco de dados mais extenso, onde se pudessem encontrar observações do consumo nos vários feriados em vários anos, as condições iniciais usadas poderiam ser baseadas em observações passadas do dia em questão e, provavelmente, a previsão seria muito mais rigorosa.

O problema da escassez de dados está ainda patente na figura 3.6, que mostra os sinais $I(t)$ e $I'(t)$. É possível verificar que nalguns períodos os integrais tomam valores muito elevados, como consequência da discrepância entre os consumos real e estimado, a qual se deve ao facto de o modelo usado não mimetizar fielmente os acontecimentos reais, devido à falta de dados disponíveis daqueles períodos para fazer a sua identificação.

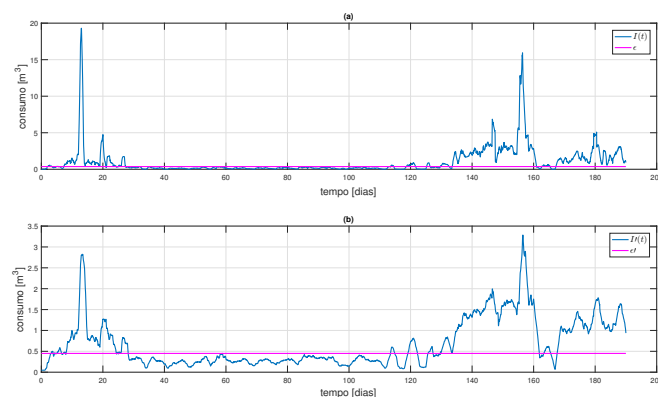


Figura 3.6:

- (a) Integral do Erro de Predição 1 passo à frente, $I(t)$;
- (b) Integral do Erro de Predição 1 passo à frente após filtragem com filtro de mediana, $I'(t)$

Outro factor que contribui fortemente para esta discrepância é o facto de se assumir que o con-

sumo é aproximadamente periódico, o que na realidade é uma aproximação muito grosseira, em particular durante a semana de acolhimento dos novos alunos.

3.7 Simulação de Fugas

Por todos os aspectos explicados anteriormente, o mecanismo de detecção de fugas apenas será eficaz no período de 21 de Setembro a 20 de Dezembro 2015. Fora deste período apenas seria possível detectar fugas com dimensão tão grande que seriam prontamente detectadas a "olho nu".

De forma a testar o desempenho do mecanismo de detecção de fugas em situação real, simularam-se fugas com a duração de 5 dias, dentro deste período, fazendo-se variar a sua intensidade, instante de início e limiar de detecção. Para estas simulações apenas se consideraram os dias úteis, uma vez que o consumo médio nos outros dias é, comparativamente, muito baixo, levando a uma grande disparidade entre o consumo previsto e real. Por isso, o caso em que os dias úteis estão todos concatenados é o pior caso para o detector.

A intensidade da fuga foi variada entre 5% e 60% em intervalos de 5%. Consideraram-se 5 instantes distintos para o início da fuga, não havendo sobreposição de períodos.

Apesar de as curvas ROC terem por objectivo caracterizar o desempenho teórico do detector, optou-se por apresentá-las também para as simulações sobre os dados reais, na figura 3.7, já que constituem uma forma simples de visualizar os resultados. Novamente é possível ver que quanto maior for a tolerância menor será o número de falsos alarmes, à custa de uma diminuição de TP e do aumento do atraso na detecção.

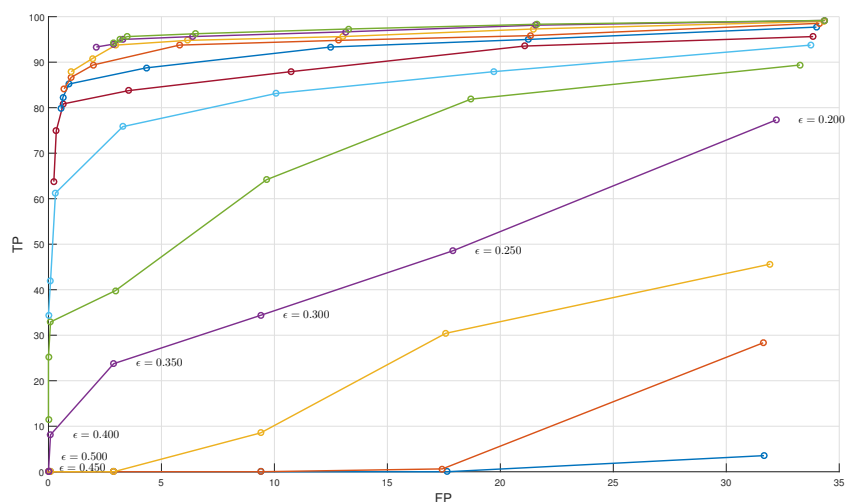


Figura 3.7: Curvas ROC para caso real

Na figura 3.8 está representado o atraso na detecção em função da tolerância. Pode-se

3. Distribuição em Rede SISO (Caso Real)

verificar que, efectivamente, para uma tolerância maior o atraso é maior. Por uma questão de melhor visualização das curvas definiu-se como 100 h o atraso nos casos em que a fuga não foi detectada.

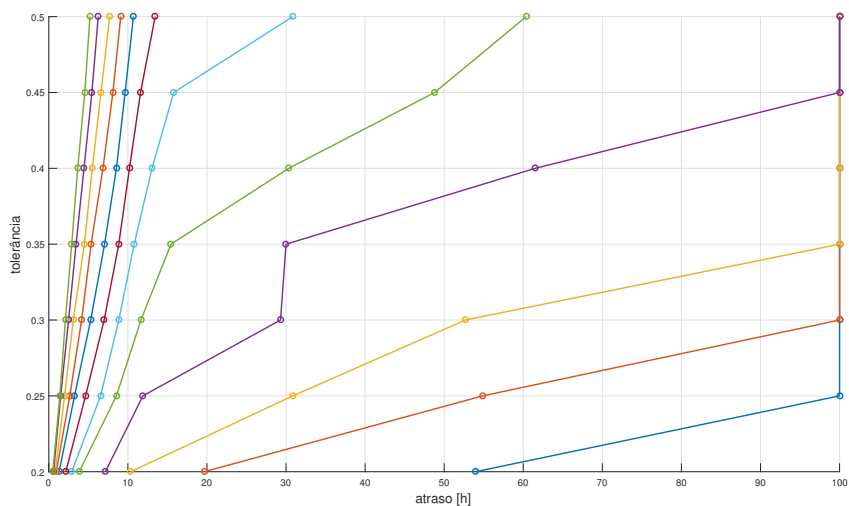


Figura 3.8: Atraso de detecção em função do limiar

Pode então concluir-se que o detector de fugas projectado cumpre o objectivo, já que detecta com grau de precisão aceitável fugas com caudal superior a $500 L/15min$ (aproximadamente 20% do consumo médio) que, em Fevereiro de 2016, era o patamar de perdas "aceitáveis", *i.e.*, é o volume de água consumido quando seria expectável que fosse praticamente nulo e cuja origem não é possível localizar.

Quanto ao atraso de detecção, apesar de uma fuga com caudal inferior a $500 L/15min$ poder demorar até 60 horas a ser detectada, este é um comportamento aceitável já que se trata de rupturas que não são visualmente detectáveis e cujo atraso de detecção estaria na ordem de meses. De facto, em Novembro de 2014, no Campus Tecnológico de Loures, teve início uma rutura cujo caudal estimado terá começado entre 5% e 10% do valor médio do consumo no período considerado nestas simulações e que foi crescendo ao longo do tempo a qual apenas foi intervencionada durante o mês de Fevereiro de 2015.

4

Distribuição em Rede MIMO

Contents

4.1	Descrição da rede	42
4.2	Análise da rede	42
4.3	Detecção de Fuga	45
4.4	Simulação de Fuga	45
4.5	Localização de Fuga	47
4.6	Resultados	49

4. Distribuição em Rede MIMO

Neste capítulo trata-se o caso de distribuição de água através de uma rede MIMO (do inglês *Multiple Input Multiple Output*). A rede é associada a um grafo descrito pela sua matriz de incidência e é completamente caracterizada quando, além da matriz de incidência, se conhece a distribuição de água por cada ramo. Além do objectivo comum às fases anteriores do trabalho, o de detecção de fugas na rede, pretende-se agora localizar com um certo grau de certeza a ocorrência da fuga. Para tal, são aplicadas técnicas de aprendizagem automática não supervisionada.

4.1 Descrição da rede

Uma rede pode ser descrita por um grafo, um objecto matemático definido como sendo um conjunto de nós (ou vértices) e de ramos, em que cada ramo liga dois nós. Os ramos podem ser direccionados, quando o fluxo se faz apenas num sentido, ou não-direccionados, quando o fluxo é bidireccional.

Para o trabalho presente assume-se que não há refluxo de água nos canos - representados pelos ramos - e, portanto, a rede é representada por um grafo direccionado. Os nós da rede podem representar pontos de abastecimento, pontos de consumo ou apenas pontos de redistribuição de água.

A rede usada neste capítulo é uma rede desenhada apenas para o desenvolvimento do trabalho. De forma a estabelecer um paralelismo físico com o trabalho desenvolvido nos capítulos anteriores, a rede pode ser vista como uma representação do conjunto de condutas que levam a água desde os pontos de abastecimento do Campus da Alameda do IST até aos diferentes pavilhões. Um exemplo de rede está representada na figura 4.1: $P1$, $P2$ e $P3$ são nós de consumo e $A1$ e $A2$ são nós de abastecimento. A rede associada ao grafo da figura 4.1 não representa a rede do IST, tendo sido escolhida por forma a ilustrar os conceitos e métodos usados nesta dissertação.

4.2 Análise da rede

A matriz de incidência da rede, M , é obtida aplicando a *Lei dos Nós* de forma análoga à que KCL (*Kirchoff Current Law*) é aplicada numa rede eléctrica. A sua dimensão é de $V \times B$, onde V é o número de nós e B o número de ramos da rede. Cada entrada da matriz pode tomar três valores:

- $[M]_{v,b} = +1$, quando o ramo C_b entra no nó N_v ;
- $[M]_{v,b} = -1$, quando o ramo C_b sai do nó N_v ;
- $[M]_{v,b} = 0$, quando o ramo C_b não liga o nó N_v a nenhum outro.

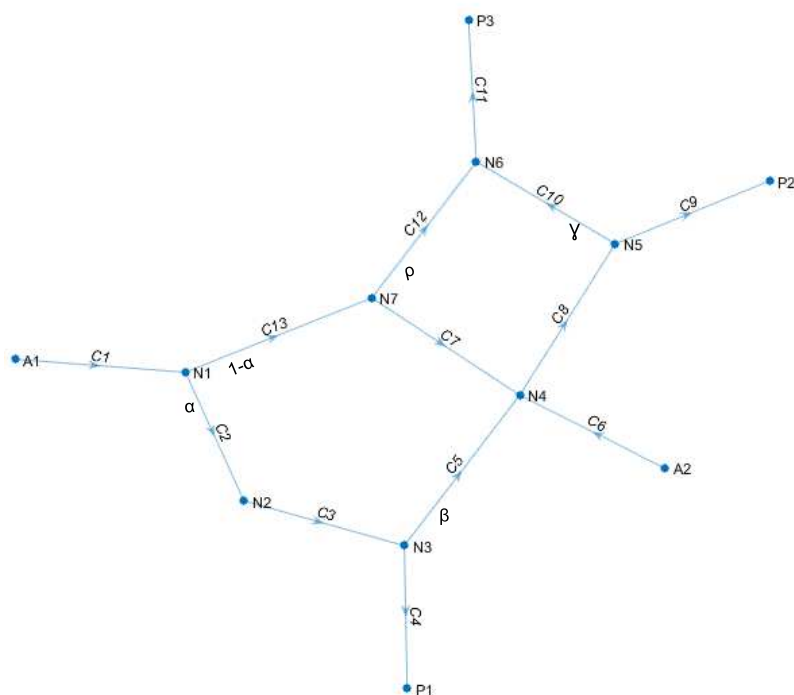


Figura 4.1: Rede

A matriz de incidência da rede considerada é então

$$M = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.1)$$

Se, além da matriz de incidência, for conhecida a distribuição de água pelos ramos de saída de cada nó torna-se possível calcular o caudal em cada ramo conhecendo-se apenas os caudais de entrada na rede.

Seja C o vector de incógnitas – o caudal em cada ramo – de dimensão $B \times 1$. O caudal em cada ramo da rede pode ser calculado resolvendo o sistema de equações lineares

$$MC = 0 \quad (4.2)$$

e impondo as restrições relativas aos caudais de entrada da rede e à divisão do caudal em cada nó.

4. Distribuição em Rede MIMO

Para o caso da rede considerada nesta dissertação o caudal em cada ramo é obtido resolvendo

$$\begin{aligned}
 MC &= 0 \\
 \text{s.t.} \\
 C_1 &= S_1, \\
 C_6 &= S_2, \\
 C_2 &= \alpha C_1, \\
 C_5 &= \beta C_3, \\
 C_{10} &= \gamma C_8, \\
 C_{12} &= \rho C_{13},
 \end{aligned} \tag{4.3}$$

onde S_i especifica o caudal injectado na rede no ponto de entrada i e α, β, γ e ρ especificam os pesos dos ramos. Uma vez que a matriz de incidência impõe que o caudal à entrada é igual ao caudal à saída em cada nó, não é necessário explicitar o peso de todos os ramos da rede.

Durante o desenvolvimento do trabalho começou por se considerar a resolução do problema descrito pela equação (4.3) usando o ambiente *cvx*, um sistema de optimização baseado em *Matlab* desenvolvido pelo Professor Stephen Boyd, da Universidade de Stanford, para resolução de problemas de optimização convexos. No entanto, dada a sua elevada complexidade computacional e o volume de dados a tratar, tal revelou-se impraticável.

Definiu-se então uma matriz, R , de dimensão $B \times E$, onde E é o número de entradas da rede (pontos de abastecimento), que estabelece a relação entre o caudal em cada ramo e o caudal injectado na rede que pode ser calculada, indirectamente, a partir da matriz de incidência.

Seja M' a matriz de incidência modificada por forma a incluir o peso de cada ramo de saída de um nó ou seja, substituindo $[M]_{v,b} = -1$ pela fracção do caudal de entrada do nó abastecedor que percorre o ramo b . Para a rede considerada a matriz M' é

$$M' = \begin{bmatrix}
 1 & -\alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -(1-\alpha) \\
 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & -(1-\beta) & -\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -(1-\gamma) & -\gamma & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & -(1-\rho) & 0 & 0 & 0 & 0 & -\rho & 1 \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{bmatrix}. \tag{4.4}$$

Multiplicando M' por C obtém-se o conjunto de equações que permite calcular o caudal em cada ramo. Resolvendo cada uma das equações em ordem aos caudais de entrada da rede e

reescrevendo-as na forma matricial obtém-se a matriz R , que para a rede representada na figura 4.1 é

$$R = \begin{bmatrix} 1 & 0 & 0 \\ \alpha & 0 & 0 \\ \alpha & 0 & 0 \\ \alpha(1-\beta) & 0 & 0 \\ \alpha\beta & 0 & 0 \\ 0 & 1 & 0 \\ (1-\rho)(1-\alpha) & 0 & 0 \\ (1-\rho)(1-\alpha) + \alpha\beta & 1 & 0 \\ (1-\gamma)[(1-\rho)(1-\alpha) + \alpha\beta] & 1-\gamma & 0 \\ \gamma[(1-\rho)(1-\alpha) + \alpha\beta] & \gamma & 0 \\ \gamma[(1-\rho)(1-\alpha) + \alpha\beta] + \rho(1-\alpha) & \gamma & 0 \\ \rho(1-\alpha) & 0 & 0 \\ (1-\alpha) & 0 & 0 \end{bmatrix}. \quad (4.5)$$

O vector de caudais $C = [C_1 \cdots C_B]^T$ pode agora ser calculado, em cada instante de tempo, simplesmente através da multiplicação de R por u :

$$C(t) = Ru(t), \quad (4.6)$$

onde $u(t) = [S_1(t) S_2(t)]^T$ é o vector dos caudais de entrada.

4.3 Detecção de Fuga

Assumindo que não há perdas em nenhum ponto da rede, o somatório do caudal em cada uma das saídas tem que ser igual ao somatório dos caudais injectados na rede. Assim, conclui-se que na presença de uma fuga os dois somatórios anteriores não serão iguais: a diferença entre o somatório do caudal das saídas e o caudal injectado na rede é o caudal de fuga.

Seja C^* o conjunto dos caudais observados. Seja Ω o conjunto dos caudais injectados na rede e Ψ o conjunto dos caudais de saída da rede.

O sinal de alarme de fuga é então dado por ¹

$$F(t) = \begin{cases} 0, & \text{se } \sum_i C_i^* = \sum_j C_j^*, \quad \forall i \in \Omega, \quad \forall j \in \Psi. \\ 1, & \text{c.c.} \end{cases} \quad (4.7)$$

4.4 Simulação de Fuga

A simulação de uma fuga é feita adicionando um nó (e um ramo) ao ramo que se pretende simular como tendo uma fuga. O caudal no ramo de fuga, C_F , é definido como sendo uma fracção do caudal que percorria o ramo afectado antes da ruptura: para uma fuga no nó b

$$C_F = f C_b, \quad f \in [0, 1]. \quad (4.8)$$

¹O símbolo \in indica os índices dos elementos do conjunto.

4. Distribuição em Rede MIMO

Ao acrescentar-se um nó e um ramo à rede, a matriz de incidência terá de ser modificada por forma acomodar as alterações. Além disso, a matriz R também se altera e as suas entradas dependem não só da intensidade da fuga, f , mas também da sua localização. Por exemplo, para uma fuga em C_{12} , conforme representado na figura 4.2, a equação (4.6) passaria a ser

$$\begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \\ C_8 \\ C_9 \\ C_{10} \\ C_{11} \\ C'_{12} \\ C_{13} \\ C_F \\ C_{12} \end{bmatrix} = \begin{bmatrix} 1 & & & & & & & & & & & & & & & & & & 0 \\ & \alpha & & & & & & & & & & & & & & & & & & 0 \\ & & \alpha & & & & & & & & & & & & & & & & & 0 \\ & & & (1-\beta)\alpha & & & & & & & & & & & & & & & & 0 \\ & & & & \alpha\beta & & & & & & & & & & & & & & & 0 \\ & & & & & 0 & & & & & & & & & & & & & & 1 \\ & & & & & & (1-\rho)(1-\alpha) & & & & & & & & & & & & & 0 \\ & & & & & & & (1-\rho)(1-\alpha) + \alpha\beta & & & & & & & & & & & & 1 \\ & & & & & & & & (1-\gamma)[(1-\rho)(1-\alpha) + \alpha\beta] & & & & & & & & & & & 1-\gamma \\ & & & & & & & & & \gamma[(1-\rho)(1-\alpha) + \alpha\beta] & & & & & & & & & & \gamma \\ & & & & & & & & & & (1-f)\rho(1-\alpha) + \gamma[(1-\rho)(1-\alpha) + \alpha\beta] & & & & & & & & & \gamma \\ & & & & & & & & & & & (1-f)\rho(1-\alpha) & & & & & & & & & 0 \\ & & & & & & & & & & & & (1-a) & & & & & & & & 0 \\ & & & & & & & & & & & & & f\rho(1-\alpha) & & & & & & & 0 \\ & & & & & & & & & & & & & & \rho(1-\alpha) & & & & & & 0 \end{bmatrix} u, \quad (4.9)$$

onde C'_{12} é o caudal que permanece em circulação, após o ponto de fuga, e C_F é o caudal de fuga.

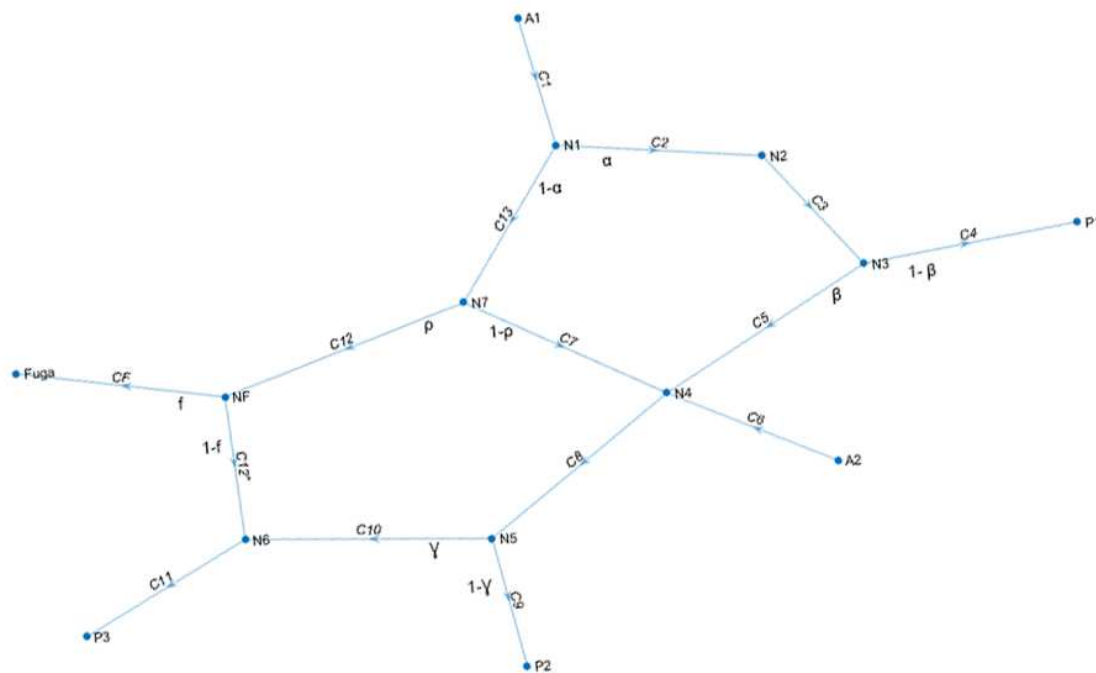


Figura 4.2: Rede com fuga em C_{12}

4.5 Localização de Fuga

A abordagem usada para a localização de fugas teve como inspiração o algoritmo de agrupamento (*clustering*) *K-means*. Os algoritmos usados no treino do sistema e na localização da fuga podem ser consultados nos Algoritmos 2 e 3.

Primeiramente calculou-se o caudal em cada um dos ramos na situação normal – na ausência de perturbações na rede, $C^{(0)}$ – tendo-se usado como entrada da rede o consumo de água do IST num determinado período de tempo. Definiu-se que a entrada *A1* seria responsável pelo abastecimento de 60% do total de água consumida e que a entrada *A2* seria responsável pelo abastecimento dos restantes 40%.

Em seguida, e para o mesmo período de tempo, simularam-se fugas em todos os ramos possíveis e para diferentes níveis de intensidade, tendo-se registado o caudal em cada um dos ramos da rede. Estes caudais foram normalizados em relação ao seu valor nominal, criando-se assim uma assinatura de cada fuga².

Assumiu-se que as medições de caudal estariam disponíveis nos ramos de consumo, C_4 , C_9 e C_{11} , e apenas em alguns dos restantes ramos: C_2 , C_5 , C_7 e C_{12} . Pode então definir-se o vector de caudais observados

$$C^* = [C_2 \ C_4 \ C_5 \ C_7 \ C_9 \ C_{11} \ C_{12}]^T. \quad (4.10)$$

Para cada uma das localizações possíveis de fuga foi calculado o valor médio das medições, $K \in \mathbb{R}^{|C^*| \times |F|}$, onde $|C^*|$ é a cardinalidade do conjunto de ramos cujo caudal é medido e $|F|$ é a cardinalidade do conjunto F , que especifica os ramos que se considera que podem sofrer fugas. Para a rede tratada nesta dissertação o conjunto F é

$$F = \{C_2, C_3, C_5, C_7, C_8, C_{10}, C_{12}, C_{13},\} \quad (4.11)$$

Seja $C_{i,j}^{(n)}(t)$ o caudal no ramo i na presença de uma fuga no ramo j , de intensidade n , no instante t . Seja N o conjunto de intensidades de fuga consideradas. Cada entrada da matriz K é dada por

$$[K]_{i,j} = \frac{1}{|N|} \frac{1}{T} \sum_{n \in N} \sum_{t=1}^T C_{i,j}^{(n)}(t), \quad i \perp C^* \quad j \perp F. \quad (4.12)$$

A classificação é feita através da minimização da distância euclidiana entre o conjunto de medições observado em cada instante, $C^*(t)$, e cada um dos centros. Assim, dir-se-á que a fuga está presente no ramo j que minimiza

$$d(j, t) = [C^*(t) - K_{*,j}] [C^*(t) - K_{*,j}]^T, \quad \forall j \perp F. \quad (4.13)$$

²Uma fuga deve ser entendida como um par (localização, intensidade)

4. Distribuição em Rede MIMO

Algoritmo 2 Detecção de Fugas em Rede – Treino

```
1:  $C_0(t) \leftarrow Ru(t)$  ▷ Caudal Nominal
2: for  $t = 1 : T$  do ▷ todos os instantes de tempo
3:   for  $j \in J$  do ▷ todas as localizações de fuga
4:     for  $n \in N$  do ▷ todas as intensidades de fuga
5:        $MC(t) = 0$ 
6:       s.t.
7:       'caudais de entrada'
8:       'distribuição de caudal na rede'
9:        $C_{Fuga}(:, t, j, n) \leftarrow C(t) \cdot \setminus C_0(t)$ 
10:    end for
11:  end for
12: end for
13:  $C_{Fuga} \leftarrow \text{mean}(C_{Fuga}, 2)$  ▷ Média ao longo da segunda dimensão do vector – tempo
14:  $C_{Fuga} \leftarrow \text{mean}(C_{Fuga}, 3)$  ▷ Média ao longo das intensidades de fuga
15:  $K \leftarrow C_{Fuga}$ 
```

Algoritmo 3 Detecção de Fugas em Rede – Detecção e Localização

```
1: if  $\sum_i C_i^* = \sum_j C_j^*$  then ▷ Detecção
2:    $F(T) \leftarrow 0$ 
3: else
4:    $F(T) \leftarrow 1$ 
5:
6: end if ▷ Localização
7:  $C^* \leftarrow C^* \cdot \setminus C^{(0)}$ 
8: minimize  $d(j, t)$ 
9: return  $j$ 
```

4.6 Resultados

De modo a avaliar o desempenho do mecanismo desenvolvido para detecção e localização de fugas, simularam-se fugas para o conjunto dos ramos F com intensidades entre 5% e 60% do caudal nominal de cada ramo, em intervalos de 5%, num período de cinco dias, e registou-se o caudal instantâneo em cada ramo da rede. Os resultados obtidos foram usados para calcular os centros K , usando o Algoritmo 2, bem como para validar o método de localização.

De forma a dar noção da dimensão de cada fuga, na tabela 4.1 apresenta-se o caudal de fuga médio aproximado – $L/15min$ – para cada localização e intensidade da fuga. Refira-se que, tipicamente, uma torneira consome entre 100 a 150 $L/15min$.

		Localização							
		C_2	C_3	C_5	C_7	C_8	C_{10}	C_{12}	C_{13}
Intensidade	5%	45	45	40	21	109	72	7	28
	10%	90	90	80	42	218	144	14	56
	15%	135	135	120	630	237	216	210	84
	20%	180	180	160	84	436	288	28	112
	25%	225	225	200	105	545	360	35	140
	30%	270	270	240	126	654	432	420	168
	35%	315	315	280	147	763	504	49	196
	40%	360	360	320	168	872	576	56	224
	45%	405	405	360	189	981	684	63	252
	50%	450	450	400	210	1090	720	70	280
	55%	495	495	440	231	1199	792	77	308
60%	540	540	480	252	1308	864	64	336	

Tabela 4.1: Tabela de Caudais de Fuga – $L/15min$ – por localização e intensidade

Na tabela 4.2 estão patentes os resultados, sucesso e insucesso, de localização de cada fuga. Foi possível verificar que a classificação foi constante ao longo do tempo, pelo que se optou por representar apenas o resultado – sucesso ou insucesso – da localização.

É possível notar que no caso de a fuga estar localizada no ramo C_3 , esta nunca é localizada correctamente. De facto, para um caudal de fuga igual ou superior a 200 $L/15min$ – correspondente a 25% do caudal nominal do ramo em questão – o resultado da localização é C_2 , o que não surpreende visto que não há divisão de caudal no nó que une estes dois ramos e, conseqüentemente, este é igual em C_2 e em C_3 e os efeitos de uma fuga em qualquer um destes ramos são os mesmos nos restantes ramos da rede.

Uma vez que nem todos os ramos veiculam o mesmo volume de água, o custo da falha na localização não é igual para todos os ramos. Na tabela 4.3 apresentam-se as taxas de erro de localização das fugas simuladas anteriormente, divididas em fugas de "baixa intensidade" – fugas para as quais o caudal não excede 200 $L/15min$ – e de "alta intensidade" – para caudais superiores a 200 $L/15min$. Na última coluna da tabela estão os resultados se se ignorar os erros resultantes de localizar uma fuga em C_2 quando esta ocorre em C_3 .

4. Distribuição em Rede MIMO

		Localização							
		C_2	C_3	C_5	C_7	C_8	C_{10}	C_{12}	C_{13}
Intensidade	5%	×	×	×	×		×		×
	10%	×	×	×	×		×		×
	15%	×	×	×	×		×		×
	20%		×		×				
	25%		×						
	30%		×						
	35%		×						
	40%		×						
	45%		×						
	50%		×		×				
	55%		×		×				
	60%		×		×				

Tabela 4.2: Sucessos e Insucessos (×) na localização de Fugas Abruptas no conjunto de treino

	Total	Ignorando C3
$C_F \leq 200$	19.79%	17.71%
$C_F > 200$	13.54%	5.21%
<i>Total</i>	33.33%	25%

Tabela 4.3: Taxa de Erro na Localização sobre o grupo de treino

É possível notar que para fugas de "baixa intensidade" a taxa de erro é praticamente a mesma, independentemente de se ignorar ou não os erros mencionados. Já para fugas de "alta intensidade" a taxa de erro é praticamente 3 vezes mais baixa quando se ignoram os referidos erros. Esta redução deve-se à eliminação de um elevado número de erros, que tinham um grande peso sobre a média.

Apesar de a taxa de erro global ser relativamente alta, 25%, a maioria dos erros surgem na localização de fugas de baixa intensidade cuja localização é, naturalmente, de maior dificuldade, já que não têm uma "assinatura" tão distinta, e que, por outro lado, são menos relevantes já que o seu impacte ambiental e económico é mais reduzido.

Posteriormente testou-se a capacidade de extrapolação do método simulando falhas de intensidades não compreendidas nas simulações aquando do cálculo dos centros, mas dentro dos limites das falhas simuladas anteriormente. Na tabela 4.5 estão novamente indicados os sucessos e insucessos de localização de fugas para cada nível de intensidade.

Novamente é óbvio que o ramo C_3 é crítico, visto que em nenhum caso a localização da fuga foi feita correctamente, pelos motivos já apresentados. Na tabela 4.6 estão registadas as taxas de erro de localização das fugas simuladas para intensidades não compreendidas no grupo de treino – o grupo de teste à do conjunto de treino. Pode-se ver que para fugas de "baixa intensidade" a taxa de erro é praticamente a mesma em ambos os conjuntos no entanto, para fugas de caudal superior a 200 L/15min a taxa de erro no conjunto de teste é claramente superior, o que era esperado. Novamente, e pelos mesmos motivos, para fugas de alta intensidade, ao

		Localização							
		C_2	C_3	C_5	C_7	C_8	C_{10}	C_{12}	C_{13}
Intensidade	8%		×	×		×			
	12%	×	×	×	×	×	×		×
	17%	×	×	×	×	×	×		×
	23%	×	×	×	×	×	×		×
	28%		×		×				
	33%		×		×				
	37%		×						
	43%		×						
	48%		×						
	54%		×						
	58%		×						

Tabela 4.4: Sucessos e Insucessos (×) na localização de Fugas Abruptas no conjunto de teste

ignorar-se as confusões feitas na localização entre os ramos C_2 e C_3 , a taxa de erro diminui significativamente

Pode-se ainda verificar que em praticamente 90% das ocorrências as fugas de alta intensidade são correctamente localizadas. Assim, conclui-se que o método proposto poderá ser viável para intensidades de fuga fora do âmbito das simulações usadas no treino, se estas forem de "alta intensidade".

	Total	Ignorando C3
$C_F \leq 200 L \setminus 15min$	18.18%	18.18%
$C_F > 200 L \setminus 15min$	18.18%	9.09%
<i>Total</i>	36.36%	27.27%

Tabela 4.5: Taxa de Erro na Localização sobre o grupo de teste

Os resultados conjuntos das duas situações – conjunto de treino e conjunto de teste – são apresentados na tabela 4.6. Embora a taxa de erro global seja superior a dois terços, para fugas de "alta intensidade" e ignorando os erros em C_3 segundo o critério anteriormente explicado, as fugas são localizadas com sucesso em quase 95% das ocorrências.

	Com C3	Ignorando C3
$C_F \leq 200$	19.02%	19.02%
$C_F > 200$	15.76%	7.07%
<i>Total</i>	34.78%	26.09%

Tabela 4.6: Taxa de Erro Global na Localização

Por fim, calculou-se o custo médio por hora de fuga localizada erradamente. Como se pode ver na tabela 4.7, apesar de gozarem de uma menor taxa de erro na sua localização, as fugas de "alta intensidade" têm um custo médio por hora muito superior ao das fugas de "baixa intensidade", o que se deve apenas à enorme diferença do caudal de fuga. A metade inferior da tabela mostra o custo quando se ignoram os erros supra mencionados.

4. Distribuição em Rede MIMO

		Treino	Teste	Combinado
inc C_3	$C_F \geq 200$	3.4807	2.3844	2.8758
	$C_F < 200$	0.5063	0.6972	0.1475
	<i>Todos</i>	1.7147	1.5408	1.6207
ign C_3	$C_F \geq 200$	5.2609	2.5122	3.5694
	$C_F < 200$	0.5063	0.6972	0.5936
	<i>Todos</i>	1.4969	1.3022	1.3996

Tabela 4.7: Custo médio – €/h – de fugas não localizadas

O custo médio ao ignorar-se C_3 nas condições anteriores, para fugas de "alta intensidade", é significativamente mais elevado, já que este ramo contribuía com um elevado número de erros de baixo custo para o cálculo da média.

5

Conclusões

Contents

5.1	Discussão e Conclusões	55
5.2	Trabalho Futuro	56

5. Conclusões

Neste capítulo apresentam-se as principais conclusões que é possível retirar do trabalho desenvolvido e enumeram-se alguns pontos que podem vir a ser melhorados ou desenvolvidos em trabalhos futuros sobre este assunto.

5.1 Discussão e Conclusões

Esta dissertação tem como objectivo principal a detecção de fugas em redes de distribuição de água. O trabalho acabou por ser dividido em duas partes: na primeira considerou-se uma situação de *pipeline*, onde apenas há uma entrada e saída de água e uma medição de caudal; na segunda parte considerou-se uma rede com múltiplas entradas e saídas de água interligadas por um conjunto de ramos e de nós.

Ficou mostrado que os algoritmos desenvolvidos na primeira parte do trabalho são capazes de detectar os dois tipos de fuga considerados com elevado grau de sucesso e em tempo útil. O sucesso na detecção de fugas é fruto de um trabalho prévio de modelação e previsão do consumo que foi totalmente conseguido, como ficou patente ao longo desta dissertação.

Apesar do sucesso global, houve algumas limitações que condicionaram o desempenho dos métodos em alguns casos particulares.

A suposição de que o consumo é aproximadamente periódico com periodicidade diária não se verificou em alguns períodos do ano lectivo, pelo que a predição do consumo não é suficientemente exacta para que as fugas sejam detectadas nesses períodos usando um limiar de decisão constante no tempo. Os períodos em causa são geralmente os de duração mais curta, como o de "regresso à aulas".

Outra suposição que havia sido feita foi a de que o consumo médio é constante por troços, ou seja, constante para cada período em que se dividiu o ano, o que se verificou na generalidade das semanas em todos os períodos.

A maior limitação dos algoritmos apresentados prende-se com a difícil previsão do consumo em dias feriado e ao fim-de-semana, seja pela inexistência de dados suficientes para o estabelecimento de um modelo para o consumo – caso dos feriados – seja pelo facto de o consumo não seguir um padrão evidente que possa ser previsto – caso dos fins-de-semana, em particular aos Sábados.

De forma a testar o desempenho dos algoritmos sem as condicionantes anteriormente descritas, que são impossíveis de controlar, optou-se por fazer esta avaliação sobre o período de aulas do primeiro semestre, tendo sido concatenados os dias úteis. Para fugas de caudal superior às "perdas aceitáveis", estas foram detectadas para a maioria das intensidades em mais de 80% dos casos e até a 10 horas após o seu início.

5. Conclusões

Na segunda parte do trabalho mostrou-se que é possível proceder à detecção de fugas numa rede MIMO sem ser necessário recorrer à sua simulação hidráulica, o que tem sido a prática corrente nos métodos propostos recentemente.

Os resultados da localização de fugas em rede foram muito positivos, já que foi possível proceder à sua localização correctamente em 95% dos casos considerados de "alta intensidade".

Apesar dos resultados serem globalmente satisfatórios, a localização de fugas de baixa intensidade ficou aquém do inicialmente pretendido. A sua fraca intensidade traduz-se numa "assinatura" menos característica sobre os restantes ramos da rede, tornando-se mais difícil a sua identificação.

Não obstante as limitações referidas, o balanço final do trabalho é muito positivo. Foi possível desenvolver um método original de detecção e de localização fugas que poderá ser aplicado com sucesso a situações reais, em particular ao caso do Instituto Superior Técnico. A sua aplicação tanto poderá ocorrer no imediato, através da aplicação directa da primeira parte do trabalho, como futuramente, quando a situação de monitorização da distribuição de água não se limitar ao fornecimento da EPAL e passar a contemplar toda a rede, o que permitiria localizar e intervir eventuais fugas num prazo de tempo muito mais curto relativamente ao que actualmente se verifica. Isto permitiria não só diminuir a pegada ecológica mas também a factura da água, permitindo canalizar mais investimento para áreas de investigação.

5.2 Trabalho Futuro

No que respeita a distribuição em rede SISO o principal motivo de interesse em trabalhos futuros deverá ser a aplicação dos métodos desenvolvidos para medições de pressão em vez de caudal, uma vez que estes sensores são de custo muito mais baixo e de muito mais fácil manutenção. Os outros aspectos a melhorar não são inerentes aos métodos desenvolvidos, mas sim aos dados disponíveis. Os resultados obtidos através da aplicação dos mesmos métodos poderão ser significativamente melhores se se tiver por base um conjunto de dados mais abrangente.

Em relação ao caso de distribuição em rede MIMO há vários motivos de interesse para investigação futura. Primeiramente, e tal como no caso anterior, é de grande importância a aplicação destes métodos tendo por base sensores de pressão. Outro aspecto que deverá ser alvo de estudo é a incerteza na caracterização da rede e determinar até que ponto os métodos apresentados podem ser aplicados a uma rede cujos pesos dos ramos não são conhecidos com exactidão.

Há ainda duas modificações possíveis aos métodos apresentados que poderão melhorar os resultados obtidos na localização. Em primeiro lugar, a integração do trabalho desenvolvido para a rede SISO poderá eliminar alguns erros na localização de fugas, particularmente as de baixa

intensidade. Ao fazer-se a previsão do consumo de água em cada ramo e comparando com os dados obtidos através do sensor correspondente, elimina-se grande parte da incerteza na localização de fugas que ocorram em ramos monitorizados. Em segundo lugar a utilização de outros métodos de aprendizagem das fugas além dos de agrupamento, como por exemplo redes neuronais, poderá ter um impacto muito significativo na localização de fugas de fraca intensidade, podendo eventualmente reduzir a taxa de erro para próximo da taxa exibida pelo método desenvolvido para redes SISO.

5. Conclusões

Bibliografia

- [1] C. Arsene, D. Al-Dabass, and J. Hartley. Using neural networks to monitor piping systems. Process Safety Progress, pages 119–127, 2003.
- [2] C. Arsene, D. Al-Dabass, and J. Hartley. Decision support system for water distribution systems based on neural networks and graphs. volume 39 (18), pages 315–323, 2014.
- [3] D. Atherton, K. Morton, and B. Megelas. Detecting breaks in prestressing pipe wire. Journal of American Water Works Association, 92:50–56, 2009.
- [4] A. Candelieri, D. Soldi, D. Contia, and F. Archetti. Analytical leakages localization in water distribution networks through spectral clustering and support vector machines. the icewater approach. 2014.
- [5] A. F. Colombo, P. Lee, and B. W. Karney. A selective literature review of transient-based leak detection methods. Journal of Hydro-environment Research, 2 (4):212–227, 2009.
- [6] A. Donnelly. Water optimization for network efficiency; reducing the environmental water footprint of a major water utility. 13th european forum eco-innovation, 2012.
- [7] M. Eiswirth and L. S. Burn. New methods for defect diagnosis of water pipelines. pages 137–150, 2001.
- [8] M. Farley and R. Liemberger. Developing a non-revenue water reduction strategy: planning and implementing the strategy. Water Science and Technology: Water Supply, 5 (1):41–50, 2005.
- [9] J. Gertler. Fault Detection and Diagnosis in Engineering Systems. CRC Press, New York, 1998.
- [10] O. Hunaidi and W.T. Chu. Leak location using blind system identification in water distribution pipelines. Construction Technology Updates, 40 (1-2):134–148, 2008.
- [11] S. M. Kay. Fundamentals of Statistical Singal Processing: Estimatin Theory. Prentice Hall PTR, Upper Saddle River, NJ, 1993.
- [12] J. M. Lemos. Controlo por computador. Lecture Slides, 1.

Bibliografia

- [13] S. R. Mounce, A. J. Day, A.S. Wood, A. Khan, P. D. Widdop, and J. Machell. A neural network approach to burst detection. Water Science & Technology, 45 (4-5):237–46, 2002.
- [14] C.V. Palau, F.J. Arregui, and M. Carlos. Burst detection in water networks using principal component analysis. Journal of Water Resources Planning and Management, 138:47–54, 2012.
- [15] E. Piatyszeka, P. Voigniera, and D. Graillotb. Fault detection on a sewer network by a combination of a kalman filter and a binary sequential probability ratio test. Journal of Hydrology, 230 (3-4):258–268, 2000.
- [16] S. L. Prescott and B. Ulanicki. Time series analysis of leakage in water distribution networks, volume 2. Research Studies Press Ltd, Baldlock, Hertfordshire, Englnd, 2001.
- [17] R. Pérez, V. Puig, J. Pascual, A. Peralta, E. Landeros, and Ll. Jordanas. Pressure sensor distribution for leak detection in barcelona water distribution network. Water Science & Technology Water Supply, 6:715–721, 2009.
- [18] R. Pérez, V. Puig, J. Pascual, A. Peralta, J. Quevedo, and E. Landeros. Pressure sensor distribution for leak detection in barcelona water distribution network. Control Engineering Practice, 19 (10):1157–1167, 2011.
- [19] R. Pérez, V. Puig, J. Quevedo, G. Sanz, M. A. Cugueró Escofet, F. Nejjari, J. Meseguer, G. Cembrano, J M. Mirats Tur, and R. Sarrate. Leak localization in water networks: A model-based methodology using pressure sensors applied to a real network in barcelona. IEEE Control Systems, 34 (4):24–36, 2014.
- [20] R. S. Pudar and J. A. Liggett. Leaks in pipe networks. Journal of Hydraulic Engineering, 118:1031–1046, 1992.
- [21] J. Quevedo, V. Puig, G. Cembrano, J. Blanch, J. Aguilar, D. Saporta, G. Benito, M. Hedó, and A. Molina. Validation and reconstruction of flow meter data in the barcelona water distribution network. Control Engineering Practice, 18 (6):640–651, 2010.
- [22] J. Ragot and D. Maquin. Fault measurement detection in an urban water supply network. Journal of Process Control, 16 (9):887–902, 2006.
- [23] A. Sarrate, J. Blesa, P.D. Widdop, F. Nejjari, and J. Machell. Clustering techniques applied to sensor placement for leak detection and location in water distribution networks. Mediterranean Conference on Control and Automation, 2014.
- [24] R. Sarrate, J. Blesa, and F. Nejjari. Sensor placement for leak monitoring in drinking water networks combining clustering techniques and a semi-exhaustive search. 2016.

- [25] C. Sivapragasam, R. Maheswaran, and V. Venkatesh. Ann-based model for aiding leak detection in water distribution networks. Process Safety Progress, 5:111–114, 2008.
- [26] K. Takahashi and M. Sato. Estimation of a buried pipe location by borehole radar. IEEE International Proceedings, Geoscience and Remote Sensing Symposium, 1:304–307, 2005.
- [27] K. M. Tsang. Sensor data validation using gray models. ISA Transactions, 42 (1):9–17, 2003.
- [28] Li Xia and Li Guo-jin. Leak detection of municipal water supply network based on the cluster-analysis and fuzzy pattern recognition. In 2010 International Conference on E-Product E-Service and E-Entertainment (ICEEE), pages 1–5, 2010.
- [29] J. Yang, Y. Wen, and P. Li. Detecting leaks in water-distribution pipes. Journal of Sound and Vibration, 310 (1-2):134–148, 2000.
- [30] J. Yang, Y. Wen, and P. Li. Detecting leaks in water-distribution pipes. Third International Conference on Network and System Security, pages 534–539, 2000.



Resultados das Simulações de Monte Carlo para fuga abrupta

A. Resultados das Simulações de Monte Carlo para fuga abrupta

F [%]	<i>Atraso</i>	TP [%]	FP [%]	TN [%]	FN [%]
0.05	2.87 [h]	7.08	43.31	46.69	2.92
0.10	48.10 [min]	7.99	43.27	46.73	2.01
0.15	17.23 [min]	8.16	43.21	46.79	1.84
0.20	5.82 [min]	8.23	43.25	46.75	1.77
0.25	1.96 [min]	8.26	43.20	46.80	1.74
0.30	0.70 [min]	8.26	43.20	46.80	1.74
0.35	0.30 [min]	8.27	43.36	46.64	1.73
0.40	0.16 [min]	8.28	43.24	46.76	1.72
0.45	0.05 [min]	8.28	43.33	46.67	1.72
0.50	0.04 [min]	8.28	43.38	46.62	1.72
0.55	0.01 [min]	8.26	43.30	46.70	1.74
0.60	0.00 [min]	8.29	43.10	46.90	1.71

Tabela A.1: Resultados das Simulações de Monte Carlo para $\epsilon = 0.010$

F [%]	<i>Atraso</i>	TP [%]	FP [%]	TN [%]	FN [%]
0.05	10.58 [h]	3.45	7.94	82.06	6.55
0.10	3.23 [h]	5.93	8.00	82.00	4.07
0.15	1.23 [h]	6.54	7.98	82.02	3.46
0.20	30.55 [min]	6.77	8.00	82.00	3.23
0.25	12.89 [min]	6.86	7.90	82.10	3.14
0.30	5.22 [min]	6.91	7.94	82.06	3.09
0.35	2.29 [min]	6.93	7.95	82.05	3.07
0.40	0.89 [min]	6.95	7.93	82.07	3.05
0.45	0.59 [min]	6.96	7.93	82.07	3.04
0.50	0.35 [min]	6.97	7.95	82.05	3.03
0.55	0.15 [min]	6.97	7.99	82.01	3.03
0.60	0.08 [min]	6.98	7.94	82.06	3.02

Tabela A.2: Resultados das Simulações de Monte Carlo para $\epsilon = 0.012$

F [%]	<i>Atraso</i>	TP [%]	FP [%]	TN [%]	FN [%]
0.05	16.41 [h]	1.29	0.54	89.46	8.71
0.10	6.38 [h]	4.74	0.53	89.47	5.26
0.15	2.68 [h]	5.86	0.53	89.47	4.14
0.20	1.29 [h]	6.26	0.54	89.46	3.74
0.25	37.71 [min]	6.46	0.52	89.48	3.54
0.30	17.50 [min]	6.55	0.52	89.48	3.45
0.35	8.69 [min]	6.61	0.54	89.46	3.39
0.40	4.02 [min]	6.64	0.54	89.46	3.36
0.45	1.97 [min]	6.66	0.54	89.46	3.34
0.50	1.10 [min]	6.67	0.55	89.45	3.33
0.55	0.53 [min]	6.68	0.54	89.46	3.32
0.60	0.32 [min]	6.68	0.53	89.47	3.32

Tabela A.3: Resultados das Simulações de Monte Carlo para $\epsilon = 0.014$

A. Resultados das Simulações de Monte Carlo para fuga abrupta

F [%]	<i>Atraso</i>	TP [%]	FP [%]	TN [%]	FN [%]
0.05	18.46 [h]	0.38	0.02	89.98	9.62
0.10	9.28 [h]	3.81	0.01	89.99	6.19
0.15	4.29 [h]	5.38	0.02	89.98	4.62
0.20	2.19 [h]	6.00	0.01	89.99	4.00
0.25	1.18 [h]	6.28	0.02	89.98	3.72
0.30	37.66 [min]	6.43	0.01	89.99	3.57
0.35	19.76 [min]	6.52	0.02	89.98	3.48
0.40	9.85 [min]	6.57	0.01	89.99	3.43
0.45	5.35 [min]	6.60	0.02	89.98	3.40
0.50	2.68 [min]	6.62	0.02	89.98	3.38
0.55	1.66 [min]	6.64	0.01	89.99	3.36
0.60	0.82 [min]	6.65	0.02	89.98	3.35

Tabela A.4: Resultados das Simulações de Monte Carlo para $\epsilon = 0.016$

F [%]	<i>Atraso</i>	TP [%]	FP [%]	TN [%]	FN [%]
0.05	19.75 [h]	0.10	0.00	90.00	9.90
0.10	11.67 [h]	3.01	0.00	90.00	6.99
0.15	5.72 [h]	4.96	0.00	90.00	5.04
0.20	3.06 [h]	5.75	0.00	90.00	4.25
0.25	1.76 [h]	6.12	0.00	90.00	3.88
0.30	1.03 [h]	6.32	0.00	90.00	3.68
0.35	35.32 [min]	6.44	0.00	90.00	3.56
0.40	19.77 [min]	6.52	0.00	90.00	3.48
0.45	11.22 [min]	6.56	0.00	90.00	3.44
0.50	5.85 [min]	6.60	0.00	90.00	3.40
0.55	3.32 [min]	6.62	0.00	90.00	3.38
0.60	1.90 [min]	6.63	0.00	90.00	3.37

Tabela A.5: Resultados das Simulações de Monte Carlo para $\epsilon = 0.018$

F [%]	<i>Atraso</i>	TP [%]	FP [%]	TN [%]	FN [%]
0.05	21.13 [h]	0.02	0.00	90.00	9.98
0.10	13.72 [h]	2.28	0.00	90.00	7.72
0.15	7.14 [h]	4.55	0.00	90.00	5.45
0.20	3.96 [h]	5.50	0.00	90.00	4.50
0.25	2.34 [h]	5.96	0.00	90.00	4.04
0.30	1.48 [h]	6.20	0.00	90.00	3.80
0.35	52.09 [min]	6.36	0.00	90.00	3.64
0.40	30.54 [min]	6.46	0.00	90.00	3.54
0.45	18.42 [min]	6.52	0.00	90.00	3.48
0.50	10.89 [min]	6.56	0.00	90.00	3.44
0.55	6.50 [min]	6.59	0.00	90.00	3.41
0.60	3.40 [min]	6.61	0.00	90.00	3.39

Tabela A.6: Resultados das Simulações de Monte Carlo para $\epsilon = 0.020$

F [%]	<i>Atraso</i>	TP [%]	FP [%]	TN [%]	FN [%]
0.05	21.92 [h]	0.00	0.00	90.00	10.00
0.10	15.33 [h]	1.66	0.00	90.00	8.34
0.15	8.50 [h]	4.12	0.00	90.00	5.88
0.20	4.83 [h]	5.24	0.00	90.00	4.76
0.25	2.90 [h]	5.80	0.00	90.00	4.20
0.30	1.87 [h]	6.09	0.00	90.00	3.91
0.35	1.18 [h]	6.28	0.00	90.00	3.72
0.40	43.81 [min]	6.40	0.00	90.00	3.60
0.45	27.51 [min]	6.48	0.00	90.00	3.52
0.50	16.59 [min]	6.53	0.00	90.00	3.47
0.55	9.84 [min]	6.57	0.00	90.00	3.43
0.60	6.03 [min]	6.59	0.00	90.00	3.41

Tabela A.7: Resultados das Simulações de Monte Carlo para $\epsilon = 0.022$

F [%]	<i>Atraso</i>	TP [%]	FP [%]	TN [%]	FN [%]
0.05	21.73 [h]	0.00	0.00	90.00	10.00
0.10	16.46 [h]	1.25	0.00	90.00	8.75
0.15	9.72 [h]	3.74	0.00	90.00	6.26
0.20	5.57 [h]	5.03	0.00	90.00	4.97
0.25	3.47 [h]	5.63	0.00	90.00	4.37
0.30	2.26 [h]	5.98	0.00	90.00	4.02
0.35	1.49 [h]	6.19	0.00	90.00	3.81
0.40	59.64 [min]	6.33	0.00	90.00	3.67
0.45	38.64 [min]	6.43	0.00	90.00	3.57
0.50	24.41 [min]	6.49	0.00	90.00	3.51
0.55	14.43 [min]	6.54	0.00	90.00	3.46
0.60	8.93 [min]	6.57	0.00	90.00	3.43

Tabela A.8: Resultados das Simulações de Monte Carlo para $\epsilon = 0.024$

A. Resultados das Simulações de Monte Carlo para fuga abrupta

B

Resultado das Simulações de Monte Carlo para Fuga Progressiva

B. Resultado das Simulações de Monte Carlo para Fuga Progressiva

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga} [m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	32.89	3.55	11.11	20.75	45.92	92.00
1.0	31.03	5.67	12.84	20.85	45.82	92.00
1.5	26.98	5.89	15.93	20.66	46.01	92.00
2.0	23.47	5.60	19.89	20.81	45.86	92.00
2.5	20.18	4.78	24.09	20.81	45.86	92.00
3.0	17.72	4.16	27.51	20.75	45.91	92.00
3.5	16.42	3.96	29.47	20.74	45.92	92.00
4.0	15.09	3.68	30.37	20.92	45.74	92.00

Tabela B.1: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.052$

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga} [m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	45.57	5.99	8.29	15.74	50.93	92.00
1.0	40.92	9.10	10.13	15.56	51.11	92.00
1.5	36.32	9.81	12.89	15.64	51.03	92.00
2.0	31.09	8.54	17.03	15.65	51.01	92.00
2.5	26.10	7.00	21.70	15.77	50.90	92.00
3.0	22.65	5.89	25.87	15.88	50.79	92.00
3.5	20.63	5.38	28.46	15.66	51.01	92.00
4.0	18.18	4.71	29.79	15.74	50.92	92.00

Tabela B.2: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.053$

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga} [m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	57.49	9.62	6.17	11.39	55.28	92.00
1.0	53.83	15.20	7.55	11.28	55.38	92.00
1.5	47.80	15.57	10.20	11.35	55.32	92.00
2.0	40.06	13.01	14.19	11.35	55.32	92.00
2.5	33.04	10.00	19.21	11.30	55.36	92.00
3.0	28.52	8.29	23.98	11.26	55.41	92.00
3.5	24.95	7.10	27.29	11.41	55.25	92.00
4.0	22.22	6.27	29.08	11.44	55.22	92.00

Tabela B.3: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.054$

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga} [m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	71.02	14.67	4.36	8.01	58.66	92.00
1.0	66.79	24.09	5.47	7.87	58.79	92.00
1.5	60.62	24.35	7.70	7.93	58.73	92.00
2.0	51.26	20.08	11.52	7.90	58.76	92.00
2.5	41.92	15.15	16.42	7.95	58.71	92.00
3.0	34.43	11.16	21.81	7.88	58.78	92.00
3.5	29.56	9.11	25.98	7.94	58.72	92.00
4.0	26.41	8.02	28.26	7.99	58.67	92.00

Tabela B.4: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.055$

B. Resultado das Simulações de Monte Carlo para Fuga Progressiva

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga} [m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	81.76	21.01	2.97	5.33	61.34	92.00
1.0	79.29	35.21	3.80	5.41	61.25	92.00
1.5	72.03	35.65	5.76	5.33	61.34	92.00
2.0	62.00	28.89	9.05	5.29	61.38	92.00
2.5	50.51	21.00	13.83	5.39	61.27	92.00
3.0	41.24	15.02	19.53	5.36	61.31	92.00
3.5	34.50	11.66	24.47	5.31	61.35	92.00
4.0	30.23	9.85	27.51	5.32	61.34	92.00

Tabela B.5: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.056$

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga} [m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	91.17	27.56	1.96	3.43	63.23	92.00
1.0	89.85	47.43	2.57	3.40	63.27	92.00
1.5	85.21	52.25	4.05	3.46	63.20	92.00
2.0	75.49	43.83	6.88	3.42	63.25	92.00
2.5	60.33	29.17	11.43	3.44	63.22	92.00
3.0	47.89	19.93	17.17	3.42	63.25	92.00
3.5	39.98	14.89	22.66	3.47	63.20	92.00
4.0	33.87	11.68	26.46	3.44	63.22	92.00

Tabela B.6: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.057$

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga} [m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	98.53	35.03	1.23	2.11	64.56	92.00
1.0	99.03	61.40	1.66	2.10	64.56	92.00
1.5	96.35	71.53	2.76	2.14	64.53	92.00
2.0	86.03	60.48	5.15	2.13	64.54	92.00
2.5	71.10	41.17	9.11	2.13	64.54	92.00
3.0	55.27	25.53	14.74	2.13	64.54	92.00
3.5	45.13	18.36	20.69	2.14	64.53	92.00
4.0	38.08	14.32	25.34	2.12	64.54	92.00

Tabela B.7: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.058$

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga} [m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	107.02	41.34	0.74	1.29	65.37	92.00
1.0	107.35	74.85	1.03	1.26	65.41	92.00
1.5	105.13	90.27	1.86	1.31	65.36	92.00
2.0	97.88	81.09	3.62	1.27	65.39	92.00
2.5	83.16	58.42	6.94	1.26	65.41	92.00
3.0	64.26	34.83	12.32	1.31	65.36	92.00
3.5	50.53	22.83	18.66	1.31	65.36	92.00
4.0	42.03	16.94	23.98	1.30	65.37	92.00

Tabela B.8: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.059$

$F[\%/dia]$	$Atraso[h]$	$Q_{fuga}[m^3]$	TP [%]	FP [%]	TN [%]	FN [%]
0.5	110.51	46.34	0.43	0.73	65.94	92.00
1.0	110.36	86.10	0.64	0.74	65.92	92.00
1.5	113.75	109.73	1.19	0.72	65.95	92.00
2.0	107.95	106.05	2.51	0.74	65.92	92.00
2.5	94.02	78.22	5.23	0.75	65.91	92.00
3.0	73.29	47.05	10.18	0.72	65.94	92.00
3.5	56.24	28.34	16.52	0.77	65.90	92.00
4.0	46.25	20.35	22.38	0.72	65.95	92.00

Tabela B.9: Resultados das Simulações de Monte Carlo para $\epsilon' = 0.060$

B. Resultado das Simulações de Monte Carlo para Fuga Progressiva

C

Coeficientes dos Modelos de Consumo para cada período do ano

C. Coeficientes dos Modelos de Consumo para cada período do ano

Período	$A(z)$	$C(z)$	\bar{Q} [$m^3/15min$]
1/SET/2015 a 6/SET/2015	$z^2 - 1.4053z + 0.4208$	$z^2 - 0.7103z - 0.1510$	2.32
7/SET/2015 a 27/SET/2015	$z^2 - 0.8320z - 0.0003$	$z^2 - 0.0243z + 0.0235$	2.83
28/SET/2015 a 20/DEZ/2015	$z^2 - 1.4059z + 0.4409$	$z^2 - 0.8806z + 0.0437$	2.43
21/DEZ/2015 a 3/JAN/2016	$z^2 - 1.2452z + 0.2773$	$z^2 - 0.7265z - 0.0898$	1.74
4/JAN/2016 a 31/JAN/2016	$z^2 - 0.8420z - 0.0341$	$z^2 - 0.2598z - 0.1066$	1.77
1/FEV/2016 a 14/FEV/2016	$z^2 - 0.1433z - 0.7381$	$z^2 + 0.4163z - 0.2648$	0.98
15/FEV/2016 a 8/MAR/2016	$z^2 - 0.9265z - 0.0085$	$z^2 - 0.4548z - 0.2177$	0.93
Sábado	$z^2 - 1.1395z + 0.1840$	$z^2 - 0.5968z - 0.1092$	0.98
Domingo	$z^2 - 0.1977z - 0.7515$	$z^2 + 0.2823z - 0.5363$	0.93

Tabela C.1: Parâmetros dos Modelos do Sistema Σ e consumo médio para cada período do ano

C. Coeficientes dos Modelos de Consumo para cada período do ano

D

Resultados da Validação dos Modelos: Autocorrelação do resíduo e consumos real e obtido através do modelo identificado

D. Resultados da Validação dos Modelos: Autocorrelação do resíduo e consumos real e obtido através do modelo identificado

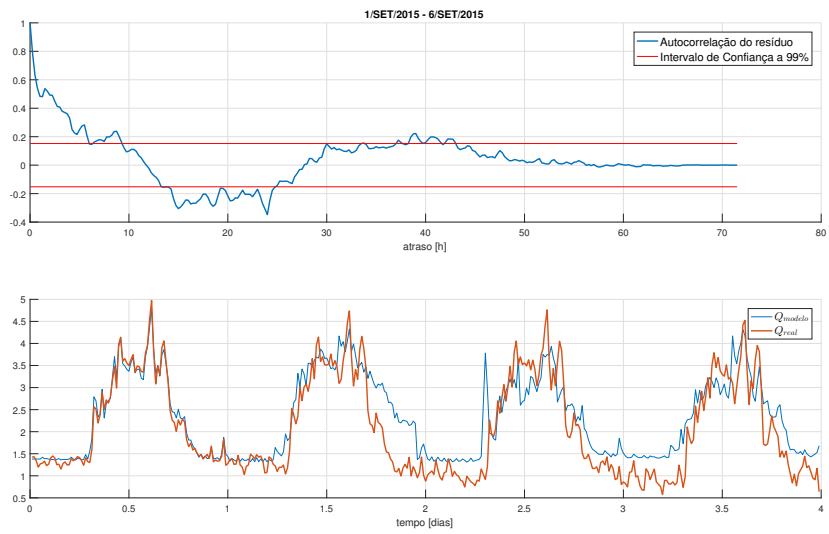


Figura D.1: 1 Setembro 2015 a 6 Setembro 2015

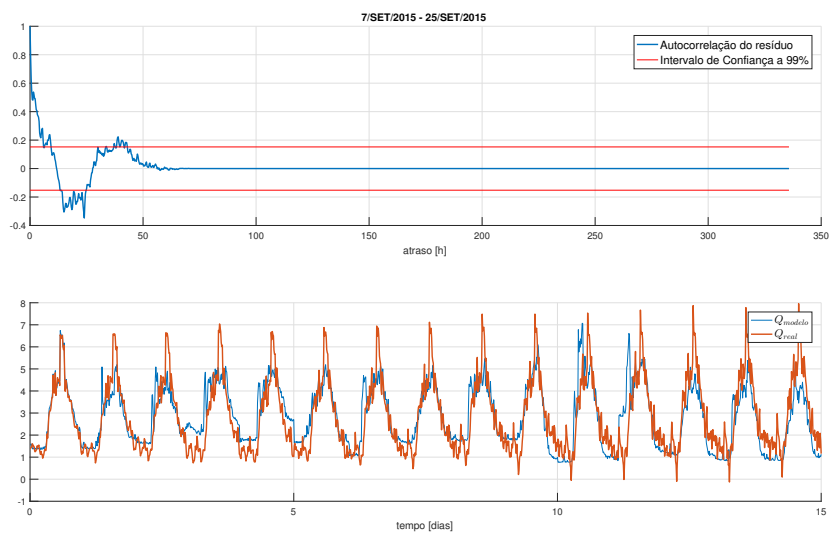


Figura D.2: 7 Setembro 2015 a 27 Setembro 2015

D. Resultados da Validação dos Modelos: Autocorrelação do resíduo e consumos real e obtido através do modelo identificado

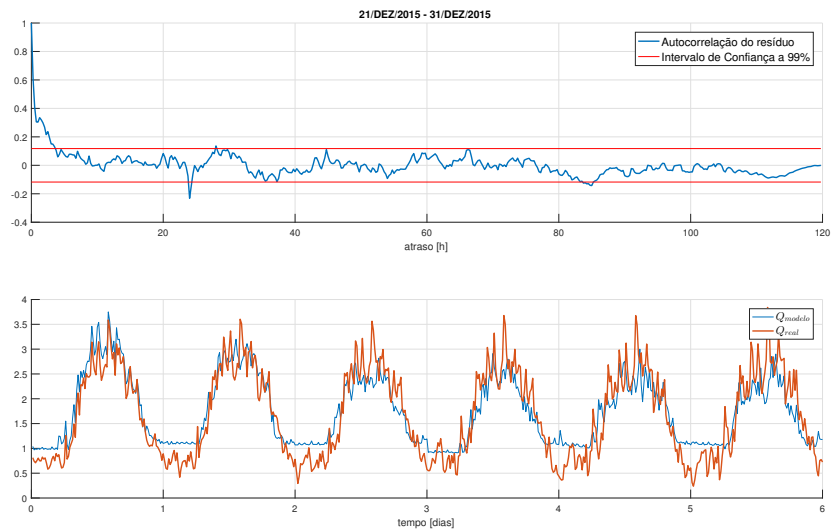


Figura D.3: 21 Dezembro 2015 a 3 Janeiro 2016

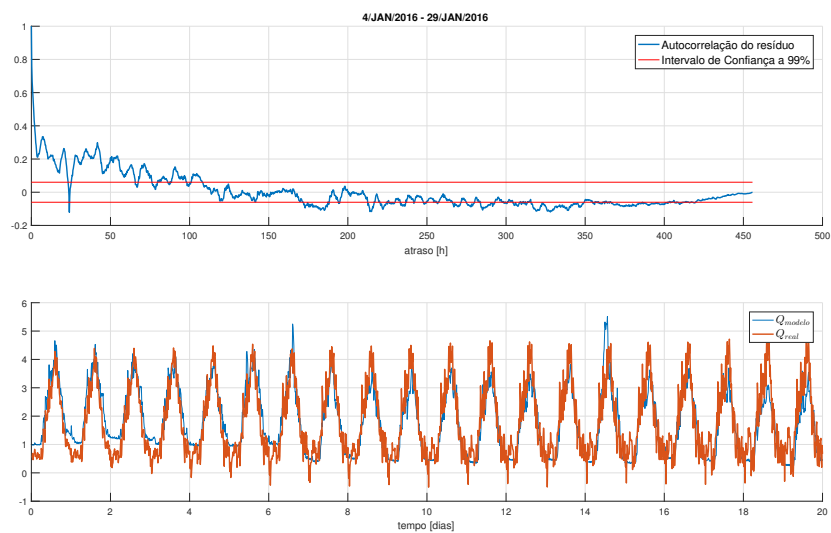


Figura D.4: 4 Janeiro 2016 a 31 Janeiro 2016

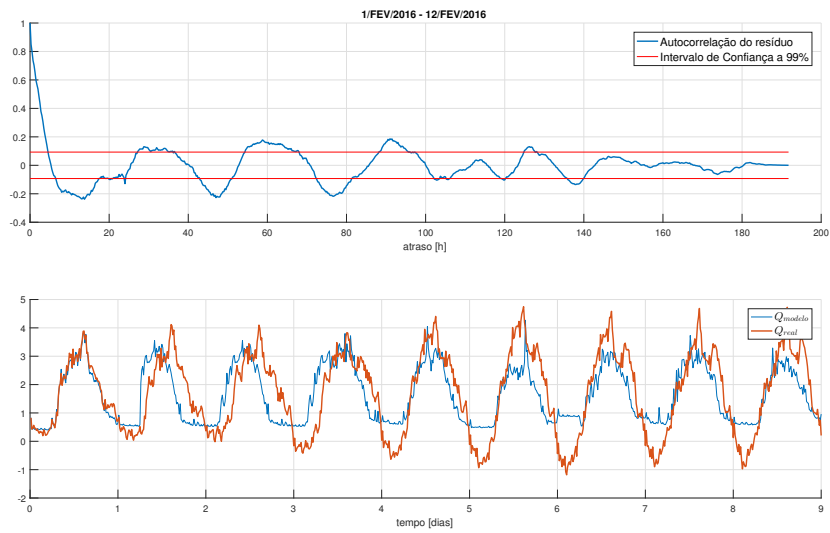


Figura D.5: 1 Fevereiro 2016 a 14 Fevereiro 2015

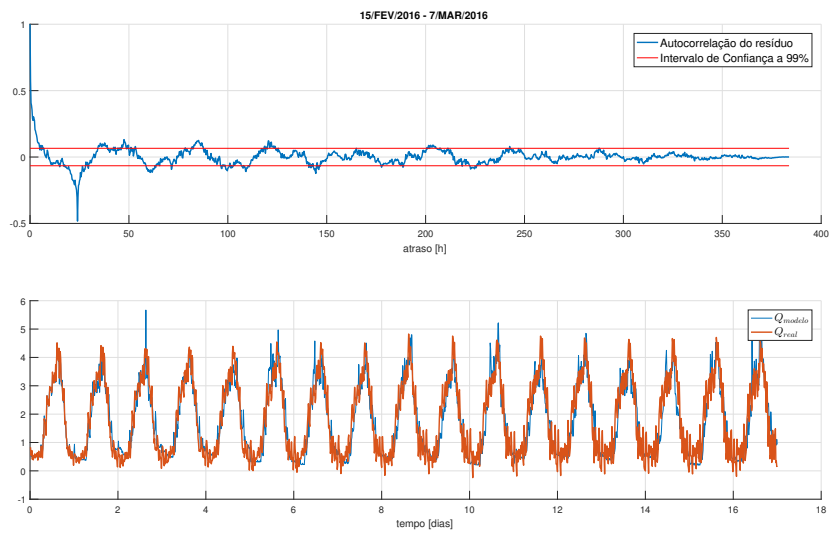


Figura D.6: 15 Fevereiro 2015 a 8 Março 2016

D. Resultados da Validação dos Modelos: Autocorrelação do resíduo e consumos real e obtido através do modelo identificado
