

• C •

FCTUC FACULDADE DE CIÊNCIAS  
E TECNOLOGIA  
UNIVERSIDADE DE COIMBRA

DEPARTAMENTO DE  
ENGENHARIA MECÂNICA

# **Otimização da Localização de Sensores para a Detecção de Contaminantes em Redes de Abastecimento de Água**

Dissertação apresentada para a obtenção do grau de Mestre em Engenharia e  
Gestão Industrial

**Autor**

**Margarida Isabel Rocha Dolores**

**Orientador**

**Carlos Alberto Henggeler de Carvalho Antunes**

**Júri**

<b>Presidente</b>	<b>Professor Doutor Cristóvão Silva</b> <b>Professor Auxiliar da Universidade de Coimbra</b>
<b>Vogais</b>	<b>Professor Doutor Álvaro Filipe Peixoto Cardoso de Oliveira Gomes</b> <b>Professor Auxiliar da Universidade de Coimbra</b>
<b>Orientador</b>	<b>Professor Doutor Carlos Alberto Henggeler de Carvalho Antunes</b> <b>Professor Catedrático da Universidade de Coimbra</b>

---

**Coimbra, Setembro de 2014**



## **Agradecimentos**

Agradeço a todos os que, direta ou indiretamente, me ajudaram na elaboração desta dissertação, com particular ênfase ao meu orientador, Professor Doutor Carlos Henggeler Antunes, e à minha família por todo o apoio e paciência.



## Resumo

A existência de água potável em quantidade e com elevada qualidade é fundamental para a sociedade atual. A água é utilizada para cozinhar, beber, tomar banho, mas também para fazer funcionar hospitais, restaurantes e fábricas ou ainda para combater incêndios. A segurança das infraestruturas de abastecimento de água tem preocupado governos, entidades reguladoras e entidades gestoras, levando à conceção e implementação de sistemas de alerta de contaminação que visam minimizar as consequências de incidentes acidentais ou propositados.

Esta dissertação tem como objetivo determinar a localização ótima de sensores numa rede de distribuição de água, para a deteção atempada de eventos de contaminação. Para este fim, devido à natureza combinatoria do problema, foram utilizadas abordagens metaheurísticas com vista à determinação de soluções que satisfaçam um potencial decisor, seguindo as definições decorrentes do evento designado *The Battle of the Water Sensor Networks* (Ostfeld *et al.*, 2008). Este evento serviu de referência relativamente às redes de teste e às quatro funções objetivo dos modelos de otimização usados nesta dissertação (o tempo de deteção,  $Z_1$ , a população afetada,  $Z_2$ , o consumo de água contaminada,  $Z_3$  e a probabilidade de deteção de eventos,  $Z_4$ ).

Os resultados obtidos mostraram que a metaheurística *simulated annealing* permitiu obter maior abrangência da pesquisa do espaço de soluções não dominadas nos modelos multiobjetivo estudados. Verificou-se também que dada a correlação existente entre as funções objetivo  $Z_2$  e  $Z_3$ , a otimização com três funções objetivo produz resultados idênticos à otimização com quatro funções objetivo, reduzindo o tempo computacional requerido.

Nesta dissertação foram ainda discutidas formas de apoiar um decisor na escolha da solução que mais se adequa às suas preocupações, exemplificando abordagens possíveis das quais resultaram propostas de soluções para o problema colocado, comparáveis às identificadas na literatura.

**Palavras-chave:** Metaheurísticas; Redes de Água; Localização de sensores



## Abstract

The existence of safe drinking water in quantity and quality is fundamental to modern society. Water is used for cooking, drinking, bathing, but also to operate hospitals, restaurants and manufacturing plants, or even for fighting fires. The security of drinking water infrastructures concerns governments, regulatory authorities and utilities, and has led to the design and deployment of contamination warning systems that aim at minimizing the consequences of accidental or intentional contamination events.

The objective of this dissertation is to find the optimal location of sensors to install in a water distribution network, in order to timely detect contamination events. Due to the combinatorial nature of the problem, metaheuristics approaches were used to determine solutions that may satisfy a potential decision maker, following the definitions originated for the event called *The Battle of the Water Sensor Networks* (Ostfeld *et al.*, 2008). These definitions include the two test networks and the four objective functions used in this dissertation (the expected time of detection,  $Z_1$ , the expected population affected prior to detection,  $Z_2$ , the expected consumption of contaminated water prior to detection,  $Z_3$ , and the detection likelihood,  $Z_4$ ).

The results obtained show that the metaheuristic simulated annealing allowed a wider coverage of the search space for the non-dominated solutions in the multiobjective models under analysis. Another conclusion was that the existing correlation between objective functions  $Z_2$  and  $Z_3$  led to similar results when optimizing for three and four objective functions, with a significant reduction in computational time requirements for the former option.

This dissertation also discussed ways of helping a decision maker choosing the solution that better matches his/her concerns, exemplifying possible approaches to identify potential solutions, which are comparable with the ones found in the literature for the same test networks.

**Keywords** Metaheuristics; Water distribution networks; sensor placement



# Índice

1.	Introdução .....	1
2.	Revisão bibliográfica .....	3
2.1.	Estudos de otimização em redes de abastecimento de água .....	3
2.2.	Otimização para segurança das redes de abastecimento de água .....	3
2.2.1.	Métodos de otimização para a localização de sensores em RAA .....	6
2.3.	<i>The Battle of the Water Sensor Networks</i> .....	8
2.4.	Processos de otimização adequados a problemas combinatórios .....	10
2.4.1.	<i>Simulated Annealing</i> .....	11
2.4.2.	Algoritmos Genéticos .....	12
2.4.3.	Extensão a problemas multiobjetivo .....	13
3.	Proposta metodológica .....	15
3.1.	Análise do processo de contaminação .....	15
3.1.1.	Simulação de eventos de contaminação em redes de abastecimento de água com o EPANET .....	15
3.1.2.	Ferramenta de suporte à implementação .....	17
3.2.	Funções objetivo a otimizar .....	17
3.2.1.	Definição .....	17
3.2.2.	Abordagem do processo de cálculo das funções objetivo .....	19
3.2.3.	Codificação e avaliação de uma rede de sensores .....	22
3.2.4.	Pesquisa do espaço de soluções .....	24
3.3.	Abordagem de otimização .....	25
3.3.1.	<i>Simulated Annealing</i> .....	26
3.3.2.	Algoritmo Genético .....	31
4.	Estudos de caso .....	37
4.1.	Descrição das redes em estudo .....	37
4.2.	Aplicação do <i>simulated annealing</i> para a determinação da rede de sensores .....	38
4.2.1.	Otimização mono-objetivo com a implementação base .....	38
4.2.2.	Abordagem multiobjetivo .....	42
4.2.3.	Análise de resultados com SA .....	51
4.3.	Aplicação dos algoritmos genéticos para a determinação da rede de sensores .....	55
4.3.1.	Otimização mono-objetivo com a implementação base, incluindo elitismo .....	55
4.3.2.	Abordagem multiobjetivo .....	57
4.3.3.	Análise de resultados com AG .....	63
4.4.	Determinação de soluções para alguns dos restantes casos propostos na BWSN .....	63
4.4.1.	Rede 1 com 20 sensores .....	63
4.4.2.	Rede 2 com 20 sensores .....	67
5.	Conclusões .....	75
5.1.	Trabalho desenvolvido .....	75
5.2.	Pistas de trabalho futuro .....	78
	Referências bibliográficas .....	81
	Apêndice A - Pseudocódigos .....	87
	Apêndice B - Correspondência entre a posição do nó com a etiqueta do nó no Epanet .....	91
	Apêndice C - Resultados dos procedimentos de otimização .....	95



# ÍNDICE DE FIGURAS

Figura 1. Representação de estação de monitorização de contaminantes (Murray <i>et al.</i> , 2010).....	5
Figura 2. Resposta de sensor de cloro (linha preta) a diversos contaminantes (Murray <i>et al.</i> , 2010) .....	5
Figura 3. Representação das soluções apresentadas no caso R1A5 da BWSN, para as funções objetivo $Z_1$ , $Z_2$ e $Z_4$ .....	10
Figura 4. Ferramenta disponibilizada para cálculo das matrizes de avaliação. ....	20
Figura 5. Exemplo de recombinação com um ponto de corte. ....	31
Figura 6. Ilustração do processo de mutação implementado. ....	31
Figura 7. Interpolação linear entre limites. $p_c$ - probabilidade de crossver, $p_m$ – probabilidade de mutação (Vasconcelos <i>et al.</i> , 2001). ....	33
Figura 8. Representação da classificação em frentes de Pareto dos elementos da população.....	34
Figura 9. Procedimento do NSGA-II (Deb <i>et al.</i> , 2002). ....	35
Figura 10. Distribuição espacial da Rede 1 no EPANET.....	37
Figura 11. Distribuição espacial da Rede 2 no EPANET.....	38
Figura 12. . Evolução da solução inicial em cada iteração e da melhor solução, em função da temperatura e das iterações. ....	41
Figura 13. Representação das soluções não dominadas resultantes da otimização biobjetivo.....	44
Figura 14. Frente de Pareto na otimização biobjetivo após consolidação com 11 corridas. ....	45
Figura 15. Frente de Pareto para a otimização triobjetivo ( $Z_1$ , $Z_2$ , $Z_4$ ). ....	48
Figura 16. Localização na frente de Pareto das soluções com menor distância à solução ideal. ...	48
Figura 17. Evolução do número de soluções não dominadas em função da temperatura. ....	49
Figura 18. Visualização tridimensional dos resultados da otimização tetra-objetivo.....	50
Figura 19. Localização das soluções de menor distância à solução ideal na frente de Pareto com 4 objetivos.....	51
Figura 20. Frente de Pareto com otimização tetra-objetivo, reduzida a soluções com probabilidade de deteção superior a 75%. ....	52
Figura 21. Representação da solução proposta na frente de Pareto e comparação com soluções apresentadas na BWSN.....	53
Figura 22. Localização da solução proposta na frente de Pareto da otimização triobjetivo.....	53
Figura 23. Proposta de solução na frente de Pareto biobjetivo $Z_1$ - $Z_4$ .....	54
Figura 24. Evolução do melhor elemento e da média da população com o número de gerações. .	57
Figura 25. Representação das soluções não dominadas resultantes da otimização biobjetivo com NSGA-II.....	59
Figura 26. Localização na frente de Pareto das soluções com menor distância à solução ideal, na otimização com NSGA-II. ....	61
Figura 27. Visualização tridimensional dos resultados obtidos na otimização tetra-objetivo com NSGA-II.....	62

Figura 28. Frente de Pareto para a otimização triobjetivo ( $Z_1, Z_2, Z_4$ ), para o caso R1A20, com SA.....	65
Figura 29. Localização na frente de Pareto das soluções com menor distância à solução ideal, para o caso R1A20, com SA. ....	65
Figura 30. Frente de Pareto reduzida a soluções com probabilidade de detecção superior a 85%, para o caso R1A20, com SA.....	66
Figura 31. Representação da solução proposta na frente de Pareto e comparação com soluções apresentadas na BWSN para o caso R1A20.....	67
Figura 32. Sobreposição das frentes de Pareto obtidas para o caso R2A20.....	70
Figura 33. Análise da Frente de Pareto resultante da otimização triobjetivo, para o caso R2A20.	71
Figura 34. Histograma do número de detecções em cada nó para o caso R2A20. ....	72
Figura 35. Representação de soluções propostas na frente de Pareto para a otimização triobjetivo ( $Z_1, Z_2, Z_4$ ), para o caso R2A20, com SA, desprezando nós com menos de 500 detecções. ....	74
Figura 36. Comparação da frente de Pareto da otimização triobjetivo ( $Z_1, Z_2, Z_4$ ) com SA, para o caso R2A20, desprezando nós com menos de 500 detecções, com as apresentadas em (Ostfeld <i>et al.</i> , 2008).....	74
Figura 37. Representação das soluções propostas para a rede 1 com 5 e 20 sensores. ....	77

# ÍNDICE DE QUADROS

Quadro 1 Sistematização de alguns exemplos de otimização da localização de sensores. ....	7
Quadro 2. Síntese das abordagens utilizadas na BWSN (Ostfeld <i>et al.</i> , 2008). ....	9
Quadro 3. Determinação de possíveis valores de temperatura inicial ( $T_0$ ). ....	39
Quadro 4. Resumo do conjunto de parâmetros usados no <i>Simulated Annealing</i> . ....	39
Quadro 5. Soluções que otimizam cada função objetivo. ....	40
Quadro 6. Resumo dos parâmetros que permitiram obter a melhor solução. ....	40
Quadro 7. Conjunto de parâmetros utilizados nas simulações biobjetivo. ....	42
Quadro 8. Conjuntos de parâmetros de SA que proporcionam maior cobertura do espaço de soluções para a otimização biobjetivo das FO. ....	43
Quadro 9. Soluções com menor distância Euclideana à solução ideal. ....	46
Quadro 10. Soluções com menor distância de Chebyshev à solução ideal na otimização biobjetivo. ....	46
Quadro 11. Resumo dos melhores resultados obtidos na otimização triobjetivo em função dos parâmetros. ....	47
Quadro 12. Soluções com menor distância à solução ideal na otimização triobjetivo. ....	48
Quadro 13. Resumo da caracterização dos resultados obtidos na otimização tetra-objetivo em função dos parâmetros ....	49
Quadro 14. Soluções com menor distância à solução ideal na otimização tetra-objetivo. ....	50
Quadro 15. Solução proposta R1A5, com SA. ....	52
Quadro 16. Resumo dos parâmetros testados para o Algoritmo Genético. ....	55
Quadro 17. Soluções que otimizam cada função objetivo. ....	55
Quadro 18. Resumo dos parâmetros que permitiram obter a melhor solução. ....	56
Quadro 19. Conjunto de parâmetros utilizados nas simulações biobjetivo na otimização com NSGA-II. ....	58
Quadro 20. Conjuntos de parâmetros do NSGA-II que proporcionam maior cobertura do espaço de soluções para otimização biobjetivo das FO. ....	58
Quadro 21. Soluções com menor distância Euclideana à solução ideal na otimização biobjetivo com NSGA-II. ....	60
Quadro 22. Soluções com menor distância de Chebyshev à solução ideal na otimização biobjetivo com NSGA-II. ....	60
Quadro 23. Resumo dos melhores resultados obtidos em cada FO, na otimização triobjetivo em função dos parâmetros, na otimização com NSGA-II. ....	61
Quadro 24. Soluções com menor distância à solução ideal na otimização triobjetivo, na otimização com NSGA-II ....	61
Quadro 25. Resumo dos melhores resultados das diferentes FO, obtidos na otimização tetra-objetivo em função dos parâmetros, com NSGA-II. ....	62

Quadro 26. Soluções com menor distância à solução ideal na otimização tetra-objetivo, com NSGA-II. ....	62
Quadro 27. Resumo dos parâmetros testados na otimização mono-objetivo com SA, para o caso R1A20. ....	63
Quadro 28. Soluções para cada função objetivo no caso R1A20, na otimização mono-objetivo com SA. ....	64
Quadro 29. Soluções com menor distância à solução ideal na otimização triobjetivo (caso R1A20), com SA. ....	64
Quadro 30. Solução proposta R1A20. ....	66
Quadro 31. Soluções para cada função objetivo para o caso R2A20, na otimização mono-objetivo com SA. ....	68
Quadro 32. Conjunto de parâmetros utilizados na otimização mono-objetivo, para o caso R2A20 com SA. ....	68
Quadro 33. Resumo dos parâmetros usados para o caso R2A20 com AG. ....	69
Quadro 34. Soluções para cada função objetivo para o caso R2A20, na otimização mono-objetivo com AG. ....	69
Quadro 35. Soluções para o caso R2A20, com a menor distância à solução ideal na otimização triobjetivo, com SA. ....	70
Quadro 36. Solução proposta R2A20. ....	71
Quadro 37. Soluções para cada função objetivo para o caso R2A20, na otimização mono-objetivo, desprezando nós com menos de 500 detecções, com SA. ....	73
Quadro 38. Soluções para o caso R2A20, com a menor distância à solução ideal na otimização triobjetivo, com SA, desprezando nós com menos de 500 detecções. ....	74
Quadro 39. Solução proposta para o caso R2A20, na otimização triobjetivo com SA. ....	74

## ÍNDICE DE ALGORITMOS

Algoritmo 1. Pseudocódigo da função de avaliação.....	23
Algoritmo 2. Pseudocódigo para registo dos nós adjacentes.....	24
Algoritmo 3. Pseudocódigo da determinação da próxima solução a avaliar. ....	25
Algoritmo 4.Pseudocódigo para o <i>Simulated Annealing</i> .....	27
Algoritmo 5. Pseudocódigo que define a estrutura de vizinhança.....	27
Algoritmo 6. Excerto do pseudocódigo para o <i>simulated annealing</i> mono-objetivo com registo da melhor solução. ....	28
Algoritmo 7. Pseudocódigo da pesquisa local em torno da solução final do <i>simulated annealing</i> . .....	28
Algoritmo 8. Pseudocódigo do SMOSA. ....	30
Algoritmo 9. Pseudocódigo do AG com elitismo (1ª versão). ....	32
Algoritmo 10. Pseudocódigo para determinação de rede equivalente por eliminação de nós com pouca deteção.....	73
Algoritmo 11. Pseudocódigo da função de atualização da lista de soluções não dominadas, para a SMOSA.....	89
Algoritmo 12. Pseudocódigo do Algoritmo Genético base. ....	89
Algoritmo 13. Pseudocódigo para a seleção por torneio binário.....	90
Algoritmo 14. Pseudocódigo para produzir nova geração. ....	90



## ABREVIATURAS

BWSN – *The Battle of the Water Sensor Networks*

CAD – *Computer Aided Design*

SAC – Sistema de aviso de contaminação

EPA – United States of America Environmental Protection Agency

FCTUC – Faculdade de Ciências e Tecnologia da Universidade de Coimbra

FO – Função objetivo

AG – Algoritmo genético

GRASP – *Greedy randomized adaptive search procedure*

NSGA-II – *Nondominated sorted genetic algorithm - II*

OMS – Organização Mundial de Saúde

RAA – Rede de abastecimento de água

RNF – Reservatório de nível fixo

RNV – Reservatório de nível variável

SA – *Simulated Annealing*

SAA – Sistema de abastecimento de água

SLOTS – *Sensors local optimal transformation system*



# 1. INTRODUÇÃO

A água é essencial à vida e um abastecimento satisfatório (adequado, seguro e acessível) deve ser disponibilizado de forma universal. A Organização Mundial de Saúde, estabelece como objetivo primordial a necessidade de efetuar todos os esforços para disponibilizar água potável de qualidade e com o máximo de segurança (WHO, 2008).

A existência de água potável em quantidade e com elevada qualidade é fundamental para a sociedade atual. A água é utilizada para cozinhar, beber, tomar banho, mas também para hospitais, restaurantes e fábricas ou ainda para combater incêndios. Assim, a contaminação das infraestruturas de água potável pode afetar severamente a saúde pública e a economia de uma comunidade (Murray *et al.*, 2009). Consequentemente, a segurança das infraestruturas de abastecimento de água tem preocupado diversas entidades com responsabilidade sobre a sua gestão, nomeadamente governos, entidades reguladoras e entidades gestoras.

Um exemplo da preocupação com este assunto é a Diretiva 98/83/EC que tem por objetivo a proteção da saúde humana dos efeitos nocivos resultantes de qualquer contaminação da água destinada ao consumo humano, assegurando a sua salubridade e limpeza (Conselho Europeu, 1998). Em Portugal, existem várias entidades gestoras de água que têm vindo a implementar voluntariamente a metodologia “Plano de Segurança da Água”, estruturada de acordo com as recomendações da Organização Mundial de Saúde e da Associação Internacional da Água (Hilaco, 2012). Um Plano de Segurança da Água para consumo humano pode definir-se como um documento que identifica e prioriza riscos plausíveis que podem verificar-se num sistema de abastecimento, estabelece um sistema de controlo para os reduzir e eliminar, e estabelece processos para verificar a eficiência da gestão dos sistemas de controlo e a qualidade da água produzida. Os seus objetivos incluem a prevenção de contaminação durante o armazenamento e a distribuição (Vieira e Morais, 2005).

A análise das vulnerabilidades dos sistemas de abastecimento de água potável permitiu identificar a rede de distribuição como a parte mais vulnerável, devido ao grande número de pontos de acesso e facilidade de acesso físico, e à incapacidade de detetar contaminações em tempo útil sem a existência de sistemas integrados e fiáveis de monitorização e vigilância. Um relatório preparado pela Agência de Proteção Ambiental Americana (EPA) identificou um conjunto de contaminantes que se introduzidos num sistema de distribuição de água potável poderia causar até 10 000 mortes (US EPA, 2005). O mesmo relatório concluiu que na ausência de um sistema de alerta de contaminação (SAC), muitos destes eventos de contaminação podem não ser detetados durante semanas antes de aparecerem os primeiros casos de doença, numa altura em que já pode ser tarde para detetar a origem e a abrangência do incidente.

Esta dissertação tem como objetivo determinar a localização ótima de sensores numa rede de distribuição de água, para a detecção atempada de eventos de contaminação, procurando assim minimizar as suas potenciais consequências. Para este fim pretende-se aplicar metaheurísticas habitualmente usadas para problemas de natureza combinatória, com vista à determinação de soluções que satisfaçam um potencial decisor. Pretende-se obter soluções que otimizem simultaneamente quatro objetivos considerados fundamentais na literatura sobre o tema, procurando assim soluções de compromisso de boa qualidade que possam ser obtidas em tempo útil.

O texto encontra-se organizado em 5 capítulos. Neste primeiro capítulo foi descrita a motivação e o enquadramento, assim como os objetivos do trabalho.

No capítulo seguinte faz-se uma revisão bibliográfica sobre estudos de otimização em redes de abastecimento de água, com particular ênfase nos estudos de localização ótima de sensores para eventos de contaminação, nomeadamente os que decorreram do evento designado por *The Battle of the Water Sensor Networks* (Ostfeld *et al.*, 2008), que serve de referência para dados e pressupostos. Faz-se ainda uma revisão sobre processos de otimização adequados a problemas combinatórios, nomeadamente as metaheurísticas *simulated annealing* e algoritmos genéticos.

No capítulo 3 descreve-se a proposta metodológica com a qual se pretendeu abordar o problema, incluindo a forma como os eventos de contaminação são simulados e como essa informação é incorporada nos algoritmos de otimização, descrevendo também alguns detalhes da implementação das metaheurísticas escolhidas em vertentes mono e multiobjetivo.

No capítulo 4 descrevem-se os estudos de caso, constituídos por duas redes de diferente dimensão e para as quais foi proposta a localização de dois conjuntos de sensores com um número pré-definido, apresentando os resultados de otimização com base em diferentes abordagens. As diferentes análises realizadas destinaram-se, por um lado, a caracterizar exaustivamente o espaço de soluções, e por outro, a determinar os conjuntos de parâmetros mais adequados. As diferentes abordagens foram comparadas considerando duas estratégias para a identificação de soluções entre os conjuntos de soluções não dominadas, uma sem consideração de informação de preferências dos decisores face às diferentes funções objetivo, e outra por aplicação de um conjunto de pressupostos sobre a possível postura de um decisor, obtidos num processo interativo.

O capítulo 5 apresenta as conclusões e enuncia algumas pistas de trabalho futuro.

## **2. REVISÃO BIBLIOGRÁFICA**

### **2.1. Estudos de otimização em redes de abastecimento de água**

O principal objetivo de um sistema de abastecimento de água (SAA) é satisfazer as necessidades dos consumidores finais através de um abastecimento contínuo de elevada qualidade com uma pressão adequada (Kang e Lansey, 2012). A conceção de um sistema deste tipo coloca necessariamente um conjunto de desafios como por exemplo a seleção e dimensionamento de componentes tais como tubagens, válvulas, reservatórios, que constituem um campo privilegiado para a aplicação de algoritmos de otimização. De igual modo, as decisões de utilização de componentes operacionais tais como bombas e válvulas, de forma a satisfazer os requisitos garantindo a minimização dos custos de investimento e manutenção são também alvos preferenciais para este tipo de procedimentos (Alperovits e Shamir, 1977).

Os estudos de otimização nesta área têm-se dedicado a áreas tão diversas como projeto, nomeadamente através do dimensionamento dos diâmetros das tubagens (Alperovits e Shamir, 1977; Kang e Lansey, 2012; Savic e Walters, 1997), a otimização do escalonamento de bombas (Marques e Sousa, 2008; Mohammed e Abdulrahman, 2009), reabilitação de tubagens e deteção de perdas de água (Halhal e Walters, 1997; Mutikanga *et al.*, 2013; Wu *et al.*, 2010), controlo da qualidade da água, nomeadamente através da estimação da dosagem ótima de cloro e da localização de estações de cloragem (Gibbs *et al.*, 2010; Huang e McBean, 2008; Munavalli e Kumar, 2003), e também da localização de sensores para a deteção de eventos de contaminação accidental ou intencional (Ailamaki *et al.*, 2003; Berry *et al.*, 2005; Ostfeld *et al.*, 2008).

### **2.2. Otimização para segurança das redes de abastecimento de água**

Um SAA é particularmente vulnerável a eventos de contaminação accidentais ou intencionais devido à sua geografia distribuída e aos múltiplos pontos de acesso. A deteção de contaminantes num sistema desta natureza é particularmente complexo devido à dimensão dos sistemas que podem chegar a milhares de quilómetros de rede, à variabilidade dos perfis de consumo e ao facto de serem preferencialmente sistemas em malha, facilitando a mistura e diluição dos contaminantes (Hart e Murray, 2010).

Os sistemas de aviso de contaminação (SAC) têm vindo a ser propostos como formas viáveis de mitigar os riscos associados (Murray *et al.*, 2009; US Army, 2009; US EPA, 2005).

Um SAC integra um sistema de monitorização de dados em tempo real, ligado a sensores, assim como outras estratégias de deteção, tais como, sistemas de vigilância de saúde pública, monitorização da segurança física das instalações, vigilância das queixas dos consumidores e recolhas periódicas de amostras de água, permitindo uma rápida tomada de decisão e uma resposta efetiva a eventos de contaminação (US EPA, 2005).

Para que os SAC sejam viáveis e eficazes, é necessário ultrapassar vários desafios técnicos. Um dos aspetos essenciais reside na colocação estratégica dos sensores na RAA. Para um dado número de sensores, limitado para que o sistema não se torne impraticável, a sua localização tem que minimizar os riscos potenciais de um evento de contaminação para a saúde pública. Assim, vários tipos de estratégia têm sido usados (Hart e Murray, 2010):

- Opinião de especialistas: Nestes casos a localização dos sensores é apenas decidida pelo julgamento humano, sem recorrer a modelos computacionais.
- Métodos de ordenação: Neste caso, a informação dos especialistas é usada para ordenar potenciais localizações de sensores, que posteriormente são escolhidas com base num modelo de preferências sobre as localizações desejáveis, tais como a proximidade a instalações críticas. O processo de escolha pode ainda usar sistemas de informação geográfica e modelos da RAA para garantir uma boa cobertura.
- Otimização: A localização dos sensores pode ser decidida através de métodos de otimização que determinem qual a conceção da rede de sensores que minimizem os riscos de contaminação. Estes tipos de métodos usam modelos computacionais para estimar o desempenho da rede de sensores, através de modelos da rede e *software* de análise hidráulica e da qualidade da água.

A implementação dos SAC requer a existência de sensores capazes de identificar contaminantes específicos (sensores diretos) ou que detetem alterações significativas na qualidade da água que possam indicar incidentes de contaminação (sensores indiretos), inseridos num sistema de supervisão (Figura 1). Os sensores indiretos são por exemplo os sensores para determinar o PH, o cloro, a condutividade elétrica, o potencial de oxidação-redução e o conteúdo total de carbono orgânico (Murray *et al.*, 2010; Rosen e Bartrand, 2013). Estes parâmetros típicos da qualidade da água tendem a variar significativamente com as alterações normais no funcionamento de reservatórios, bombas ou válvulas, assim como com as variações diárias e sazonais na captação e nos consumos. Assim, é necessário que os sistemas de deteção sejam capazes de distinguir entre o que são as variações normais e as decorrentes de um evento de contaminação. O exemplo de um sistema capaz de fazer essa deteção é o *software* CANARY da Sandia (Murray *et al.*, 2010).

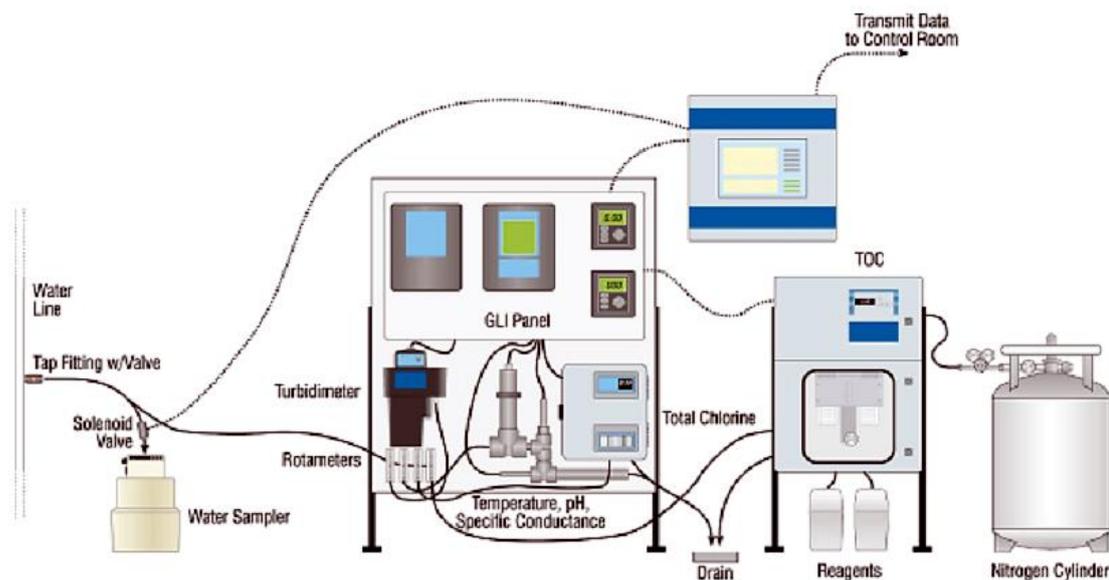


Figura 1. Representação de estação de monitorização de contaminantes (Murray *et al.*, 2010)

A capacidade de detetar diferentes contaminantes com um sensor de cloro está ilustrada na Figura 2. Sobre este assunto, é também referido na literatura que, com exceção de contaminantes específicos para “guerra química” e toxinas de origem vegetal classificadas como agentes bioquímicos, as restantes 10 classes principais de contaminantes são detetáveis indiretamente através de 3 parâmetros de medição comum: cloro residual, condutividade e conteúdo total de carbono orgânico. Alguns dos dois tipos restantes podem ser determinados através de analisadores de compostos orgânicos voláteis. Outros tipos de sensores podem determinar a contaminação potencial através da resposta de pequenos organismos, nomeadamente peixes, mexilhões, dâfnias e algas (Philadelphia Water Department e CH2M HILL, 2013).

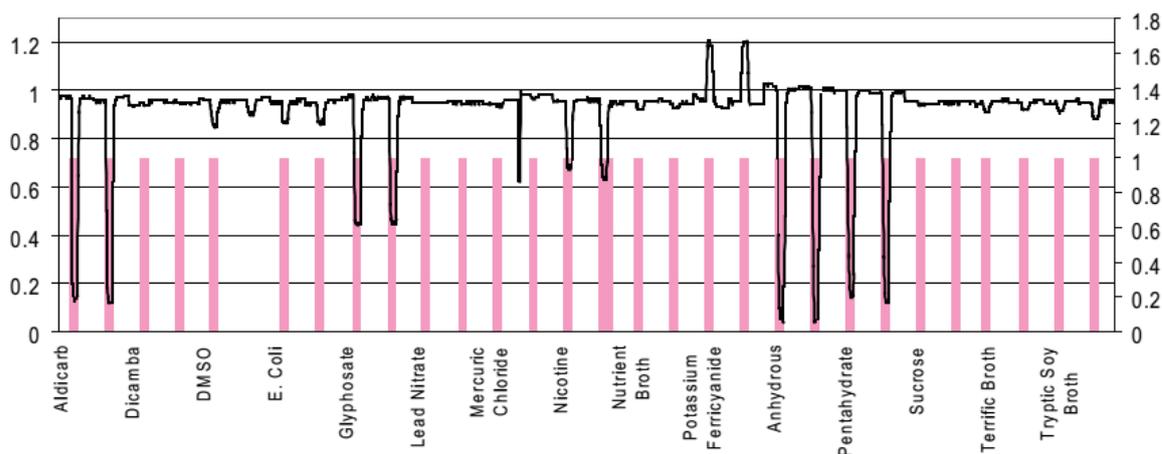


Figura 2. Resposta de sensor de cloro (linha preta) a diversos contaminantes (Murray *et al.*, 2010)

### **2.2.1. Métodos de otimização para a localização de sensores em RAA**

O interesse crescente na instalação de redes de sensores, para detetar eventos de contaminação accidental ou propositada em redes de abastecimento de água, tem levado ao desenvolvimento de diversos modelos para a otimização da localização de um conjunto de sensores que permitam à entidade gestora da rede detetar atempadamente e minimizar o impacto desses eventos de contaminação.

Hart e Murray (2010) fazem a revisão de aproximadamente 90 trabalhos relacionados com a localização ótima de sensores em RAA. Segundo estes autores as estratégias de otimização podem ser diferenciadas pela forma como usam os modelos da rede para avaliação do risco. Os estudos iniciais usavam simulações simplificadas do transporte de contaminantes, nomeadamente simulações em regime permanente, ou simulações de caudal médio. Estudos mais recentes usam simulações dinâmicas mais complexas, através de ferramentas como o EPANET (Rossman, 2000), estimando os caudais e pressões nos SAA, assim como o transporte de contaminantes e interações químicas simples. Este tipo de simulação permite ter em conta efeitos transitórios e resulta em estimativas mais precisas do risco. No entanto, necessitam de dados atuais e calibrados para a modelação das redes.

Eliades e Polycapou (2006) referem que a primeira abordagem a este problema foi feita por Lee e Deininger (1992), que procuravam colocar os sensores de forma a maximizar a cobertura dos nós de maior consumo, usando modelos de programação inteira. Desde então, vários autores prosseguiram com abordagens baseadas em programação inteira ou programação inteira mista, mas mais recentemente as abordagens baseadas em metaheurísticas têm vindo a ganhar importância por serem a única forma de lidar com redes de grande dimensão, incorporando análises mais complexas e com múltiplos objetivos. O Quadro 1 resume as principais abordagens encontradas na literatura.

Embora alguns dos métodos indicados no Quadro 1 tenham como intenção uma otimização multiobjetivo, nenhum especificou qualquer tipo de interação, sistema de apoio ou mapeamento das preferências dos decisores. Os que referem a forma como estabeleceram compromissos entre objetivos descrevem que optaram por considerar pesos iguais para todas as funções objetivo normalizadas numa escala comum. No entanto, sem aferir e explicitar as preferências do decisor face às diferentes funções objetivo, não é possível avaliar a qualidade das soluções propostas, sendo este um aspeto claramente menosprezado nesses trabalhos.

Quadro 1 Sistematização de alguns exemplos de otimização da localização de sensores.

Referência	Abordagem	Objetivos de otimização	Observações
Berry <i>et al.</i> (2005)	Programação inteira mista	Minimiza a população em risco	
Eliades e Polycarpou (2006)	Algoritmo de pesquisa: "Iterative deepening of Pareto solutions"	<ol style="list-style-type: none"> <li>1. Tempo de detecção</li> <li>2. População contaminada</li> <li>3. Consumo de água contaminada</li> <li>4. Probabilidade de detecção</li> <li>5. Cobertura dos nós de consumo</li> </ol>	Trabalho no âmbito da BWSN (ver 2.3). Segue estritamente as regras da competição para os objetivos 1 a 4.
Huang <i>et al.</i> (2006)	Algoritmo genético	<ol style="list-style-type: none"> <li>1. Tempo de detecção</li> <li>2. População contaminada</li> <li>4. Probabilidade de detecção</li> </ol>	Trabalho no âmbito da BWSN (ver 2.3).
Krause <i>et al.</i> (2006)	Algoritmo "Greedy", programação inteira mista e <i>simulated annealing</i>	<ol style="list-style-type: none"> <li>1. Tempo de detecção</li> <li>2. População contaminada</li> <li>3. Consumo de água contaminada</li> <li>4. Probabilidade de detecção</li> </ol>	Trabalho no âmbito da BWSN (ver 2.3), embora não tenha seguido as definições de forma estrita.
Xu <i>et al.</i> (2008)	Teoria de grafos	<ul style="list-style-type: none"> <li>• Centralidade dos nós na rede (medida da pertença do nó ao caminho mais curto entre os pares dos restantes nós)</li> <li>• "Receivability" (mede a possibilidade de um determinado nó receber água dos restantes)</li> </ul>	
Dorini <i>et al.</i> (2010)	Algoritmo baseado em pesquisa local: SLOTS	Idêntico a (Krause <i>et al.</i> , 2006)	Formulação mono-objetivo e formulação multiobjetivo
Watson <i>et al.</i> (2009)	Programação inteira mista e algoritmo "GRASP"	<ul style="list-style-type: none"> <li>• População contaminada</li> </ul>	Procura minimizar o caso mais grave, ou seja, a maior quantidade de população afetada.
Eliades <i>et al.</i> (2014)	Vários, incluindo algoritmos genéticos	<ul style="list-style-type: none"> <li>• Número de sensores</li> <li>• Volume médio de água contaminada</li> <li>• Volume de água contaminada no caso mais grave</li> </ul>	Plataforma aberta configurável

### 2.3. *The Battle of the Water Sensor Networks*<sup>1</sup>

O acréscimo significativo de trabalhos sobre a otimização da localização de sensores para detecção de contaminações, nomeadamente após os atentados de 11 de setembro de 2001, originou abordagens algorítmicas diferentes, o que levou alguns investigadores a propor um conjunto de regras comuns para permitir a sua comparação. Essas regras foram formalizadas como *The Battle of the Water Sensor Networks* (BWSN), um desafio lançado paralelamente ao *8th Annual Water Distribution Systems Analysis Symposium*, que decorreu em Cincinnati, Ohio, EUA, em 2006, tendo como objetivo comparar o desempenho de soluções para redes de sensores, em duas redes de abastecimento de água, cujos dados foram disponibilizados para o efeito.

A este desafio responderam quinze grupos de pesquisa, os quais podiam desenvolver as suas próprias metodologias, tendo no entanto que respeitar o conjunto de regras pré-definidas que incluíam as características dos eventos de contaminação e os objetivos a considerar na avaliação das soluções, incluindo as suas especificações de cálculo. Estes objetivos são: minimização do tempo esperado de deteção ( $Z_1$ ), do valor esperado de população afetada antes da deteção ( $Z_2$ ), e do valor esperado de consumo de água contaminada antes da deteção ( $Z_3$ ), assim como a maximização da probabilidade de deteção ( $Z_4$ ).

As abordagens utilizadas pelos grupos de pesquisa participantes encontram-se listadas no Quadro 2. Constata-se que na sua maioria recorreram a abordagens baseadas em metaheurísticas, de modo a tentar vencer as dificuldades associadas à natureza combinatória do problema.

A metodologia usada para avaliar as soluções propostas pelos concorrentes baseou-se numa ferramenta específica<sup>2</sup>, sendo composta por duas partes. A primeira, cria uma matriz de eventos de contaminação, gerados aleatoriamente ou de forma sistemática, nos diferentes nós das redes de teste, e a partir dela gera matrizes de avaliação dos quatro objetivos na perspetiva de cada um dos nós da rede a estudar. A segunda, usa as matrizes geradas para avaliar uma dada rede de sensores, de acordo com os objetivos especificados para a BWSN. A geração das matrizes de avaliação é efetuada através do *software* EPANET 2.0<sup>3</sup> que efetua a simulação hidráulica e da consequente evolução da contaminação química da rede.

No decorrer dos trabalhos foi verificado pelos vários grupos que as funções objetivo (FO)  $Z_i$  ( $i=1, 2, 3$ ) competiam com  $Z_4$ , o que o tornava num problema claramente multiobjectivo, i.e. não existiam soluções que otimizassem simultaneamente todas as funções objetivo. As soluções listadas em (Ostfeld *et al.*, 2008) para cada um dos casos apresentados são na sua maioria não comparáveis entre si, sendo umas melhores numas funções objetivo e outras noutras, correspondendo assim ao conceito de solução não dominada. Uma solução não dominada é uma

<sup>1</sup> Esta secção baseia-se na referência (Ostfeld *et al.*, 2008)

<sup>2</sup> <http://www.water-simulation.com/wsp/bwsn>

<sup>3</sup> <http://www.epa.gov/nrmrl/wswrd/dw/epanet.html>

solução admissível para o problema que se caracteriza por não existir outra solução admissível que melhore simultaneamente todas as funções objetivo, ou seja, para conseguir melhorar o valor de uma função objetivo é necessário aceitar degradar pelo menos o valor de uma outra função (Antunes *et al.*, 2012). Algumas das soluções apresentadas à BWSN são soluções dominadas, como se pode observar na Figura 3.

**Quadro 2. Síntese das abordagens utilizadas na BWSN (Ostfeld *et al.*, 2008).**

<b>Participantes</b>	<b>Metodologias</b>
Alzamora e Ayala	Algoritmos baseados na topologia.
Berry <i>et al.</i>	Formulação baseada em teoria da localização discreta complementada com metaheurística.
Dorini <i>et al.</i>	Método de otimização multiobjectivo com restrições designado “ <i>noisy cross entropy sensor locator (nCESL)</i> ”.
Eliades e Polycarpou	Método multiobjectivo com aprofundamento iterativo das soluções de Pareto.
Ghimire <i>et al.</i>	Heurística baseada na procura, localizando os sensores na proximidade dos nós de maior consumo de água ou então na maior quantidade de massa libertada.
Guan <i>et al.</i>	Baseado em algoritmo genético, com agregação dos quatro objetivos num único objetivo.
Gueli	Otimização multiobjectivo através do modelo predador-presa, baseado nos processos de evolução.
Huang <i>et al.</i>	Algoritmo genético multiobjectivo e <i>data mining</i> .
Krause <i>et al.</i>	Algoritmo <i>greedy</i> complementado com uma estratégia de <i>simulated annealing</i> .
Ostfeld e Salomons	<i>Nondominated sorted genetic algorithm II (NSGA-II)</i>
Preis e Ostfeld	NSGA-II
Propato e Piller	Programação linear inteira – mista.
Trachtman	Abordagem de engenharia selecionando locais com base em fatores racionais tais como, a distribuição da população, a pressão do sistema, a circulação de caudais, a localização de consumidores críticos, etc.
Wu e Walski	Otimização multiobjectivo com algoritmo genético, sendo os eventos de contaminação gerados pelo método de Monte Carlo.

No desafio BWSN concluiu-se que não existem regras de localização que possam ser definidas de uma forma genérica, e que os critérios de engenharia e a intuição não chegam para localizar de forma eficaz os sensores, sendo fundamental efetuar uma análise quantitativa baseada em métodos de otimização. A análise dos resultados demonstrou que os sensores não necessitam de ser agrupados e que a localização de sensores em instalações verticais (captações, reservatórios, bombas) não é um requisito.

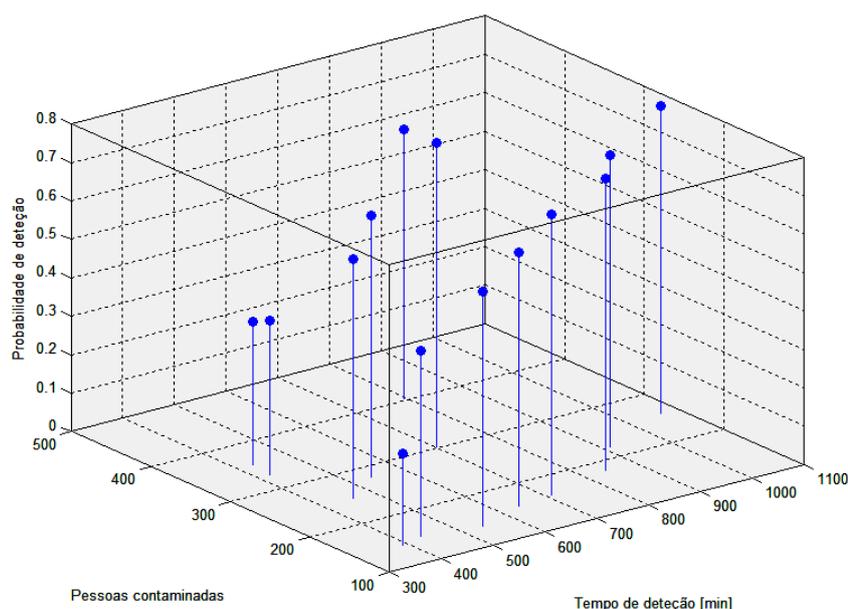


Figura 3. Representação das soluções apresentadas no caso R1A5 da BWSN, para as funções objetivo  $Z_1$ ,  $Z_2$  e  $Z_4$ .

## 2.4. Processos de otimização adequados a problemas combinatórios

A Otimização é o campo de conhecimentos cujas técnicas visam determinar os extremos (máximos e mínimos) de funções, em determinados domínios (Takahashi e Cunha, 2012). Dentro dos processos de otimização, estes autores distinguem: algoritmo, procedimento que garante a obtenção do ótimo exato da função objetivo (FO); aproximação, procedimento que permite obter um valor muito próximo do ótimo exato, não divergindo deste mais do que uma quantidade teoricamente assegurada; heurística, procedimento de busca de “boas soluções” adaptado para o problema, mas sem garantias da qualidade relativa da solução encontrada, em comparação com o ótimo exato, com a vantagem de o fazer num tempo computacional aceitável.

No caso concreto da Otimização Combinatória, o número normalmente muito elevado de soluções discretas possíveis torna a resolução do problema através de algoritmos, e mesmo de aproximações, muito morosa e mesmo impraticável do ponto de vista do esforço computacional. O objetivo principal passa a ser assim, a resolução do problema obtendo soluções de “boa qualidade” mas sem a garantia de obter a solução ótima, num tempo computacional “aceitável”, o que pode ser atingido através de uma heurística.

Para além das heurísticas mais clássicas, desenvolvidas para classes de problemas bastante particulares, tem vindo a ser desenvolvida uma classe de técnicas heurísticas estocásticas, de âmbito geral, para problemas de otimização combinatória, sendo designadas de metaheurísticas. Estas técnicas fundamentam-se em conceitos relativamente genéricos, tais como,

vizinhanças de pontos no espaço de soluções discretas, trajetórias que ligam dois pontos ou organização da informação adquirida a respeito de regiões do espaço de soluções, permitindo a construção de algoritmos de características eficientes tanto de pesquisa local como de pesquisa global (Takahashi e Cunha, 2012).

Existe uma grande diversidade de metaheurísticas com origens bastante distintas, tendo aparecido nas últimas décadas algoritmos tão diferentes como otimização por colônia de formigas, pesquisa tabu, algoritmos genéticos, *simulated annealing*, *particle swarm*, entre outros. Devido às funcionalidades específicas de cada metaheurística, a sua comparação é de várias formas mais difícil do que a comparação de outros algoritmos (Silberholz e Golden, 2010).

Blum e Roli (2003) sugerem várias formas de classificar metaheurísticas, dependendo das características selecionadas para as diferenciar, sendo cada uma delas o resultado de um ponto de vista específico, nomeadamente:

- Serem inspiradas na natureza ou não;
- Serem baseadas em população vs. em métodos de trajetória;
- Possuírem funções objetivo estáticas ou dinâmicas;
- Possuírem uma ou mais estruturas de vizinhança alternativas;
- Serem ou não baseadas em memória.

A grande diversidade de métodos e a ausência de resultados conclusivos sobre quais podem ser mais adequados para cada tipo de problema, conjugado com a limitação de tempo que impede o ensaio exaustivo de alternativas, levou à escolha para esta dissertação de duas metaheurísticas, procurando explorar a diferença entre um método de trajetória – *simulated annealing*, e um método de população – algoritmo genético, executando assim uma comparação entre uma pesquisa em torno de uma solução e o desempenho de um processo com pesquisa paralela.

#### **2.4.1. *Simulated Annealing***

O conceito de *Simulated Annealing* (SA) foi introduzido por Kirkpatrick *et al.* (1983) e é uma metaheurística tipicamente usada para problemas de otimização combinatória com variáveis discretas (Glover e Kochenberger, 2003), sendo baseada na analogia termodinâmica de arrefecimento de metais. SA é uma técnica que permite obter boas soluções para problemas de otimização mono e multiobjectivo, com tempo de computação razoavelmente reduzido (Suman e Kumar, 2005). Bem parametrizado, este método permite obter uma solução de muito “boa qualidade” para um problema mono-objectivo ou conjuntos de soluções não dominadas para problemas de otimização multiobjectivo.

Esta técnica baseia-se no conceito de pesquisa local, explorando um subconjunto de soluções admissíveis através de movimentos da solução atual para soluções vizinhas. Porém, enquanto as heurísticas de pesquisa local apenas levam a solução a movimentar-se no sentido do melhoramento, a metaheurística SA permite movimentos que pioram o valor da FO, evitando assim o bloqueio em ótimos locais. Esses movimentos ocorrem de acordo com uma probabilidade decrescente, função de um parâmetro de controlo que é designado por temperatura em referência à analogia termodinâmica em que se baseia o método (Alves, 2001). À medida que a temperatura tende para zero, os movimentos de degradação da FO diminuem e a distribuição de soluções converge para um conjunto em que a probabilidade de se encontrar a solução ótima aumenta (Glover e Kochenberger, 2003).

### **2.4.2. Algoritmos Genéticos**

O algoritmo genético (AG) é um método de otimização e pesquisa inspirado no processo de evolução natural e sobrevivência dos indivíduos mais aptos. As operações básicas envolvidas na otimização por AG incluem tipicamente quatro operadores: avaliação de desempenho (*fitness*), seleção, recombinação (*crossover*) e mutação (Zou e Lung, 2004).

Uma população inicial de soluções é normalmente gerada aleatoriamente. Cada elemento da população é representado por um conjunto de parâmetros que descreve a solução, codificado de uma forma que se identifica como cromossoma. A analogia com a natureza é estabelecida pela evolução da população de soluções através da sua substituição total ou parcial em cada iteração, designada por geração, por soluções obtidas a partir da geração anterior, através de recombinação ou mutação dos seus elementos, até se atingir um número limite de gerações. Em cada geração, é feita a avaliação de desempenho da solução através do valor de *fitness*. Este valor de *fitness* pode ser apenas o valor de FO ou incorporar elementos de avaliação adicionais como, por exemplo, uma componente de penalização de soluções não admissíveis. Os elementos com melhor desempenho são selecionados e dão origem aos seus descendentes por recombinação e por mutação, de acordo com uma dada probabilidade.

Em cada etapa da evolução, a probabilidade de que um cromossoma da população anterior seja reproduzido na geração seguinte está dependente do valor da sua *fitness*. Os elementos com melhor desempenho são os que têm uma maior probabilidade de serem selecionados. Assim, a nova população terá uma maior probabilidade de conter as melhores soluções.

O AG assenta num processo de pesquisa paralela do espaço de potenciais soluções. Este princípio permite que estruturas de soluções com uma elevada *fitness* originem um maior

número de descendentes nas gerações seguintes, idealmente gerando melhores soluções (Savic e Walters, 1995).

### **2.4.3. Extensão a problemas multiobjectivo**

Os problemas reais são intrinsecamente multidimensionais, sendo expectável que a consideração explícita de múltiplos eixos de avaliação do mérito das soluções potenciais concorra para a obtenção de soluções mais adequadas (Antunes *et al.*, 2012; Gaspar Cunha *et al.*, 2012). O tratamento deste tipo de problemas será, na maioria das circunstâncias, redutor se o circunscrevermos à procura da solução ótima para um desses objetivos ou para algum objetivo “agregado”, como seja a eficiência do ponto de vista económico medida por um qualquer indicador.

Ao contrário da otimização mono-objetivo, uma solução de um problema multiobjetivo não pode ser determinada sem a intervenção de um decisor. O conceito de ótimo de Pareto, definido originalmente no contexto da economia política, estabelece que a alocação ótima de recursos de uma sociedade não é alcançada enquanto for possível melhorar a satisfação de um indivíduo mantendo inalterada a satisfação dos outros (Marler e Arora 2004). Uma solução ótima de Pareto, também designada solução não dominada ou eficiente caracteriza-se assim, por não existir uma outra solução admissível que melhore simultaneamente todos os objetivos. O conceito de eficiência refere-se ao espaço das variáveis de decisão, enquanto o conceito de não dominância se refere ao espaço das FO, sendo uma solução não dominada a imagem de uma solução eficiente (Antunes *et al.*, 2012). A obtenção de uma solução para um problema envolve duas etapas: a determinação das soluções eficientes ou ótimas de Pareto e a etapa de tomada de decisão, para seleção de uma recomendação final de entre o conjunto das soluções eficientes (Takahashi *et al.*, 2012).

Assim, num problema multiobjetivo é também necessário incluir informação sobre as preferências do decisor no processo de apoio à tomada de decisão, permitindo discriminar entre as soluções não dominadas. A estrutura de preferência do decisor representa um conjunto de opiniões, valores, convicções e perspetivas próprias da realidade em causa. Não é, assim, possível classificar uma solução como boa ou má, sem ter em consideração os compromissos que o decisor pode estabelecer relativamente aos graus de satisfação dos diferentes objetivos, e que podem evoluir ao longo do processo de apoio à tomada de decisão. Este processo deve assim ser constituído por processos de geração de ações potenciais, avaliação, interpretação de informação, alterações de valores, aprendizagem e adaptação de preferências (Antunes *et al.*, 2012).



## **3. PROPOSTA METODOLÓGICA**

### **3.1. Análise do processo de contaminação**

Nas definições da BWSN foi especificado o EPANET 2.00.10 como ferramenta a usar para simular o funcionamento da rede (Ostfeld *et al.*, 2005).

O conceito essencial que está na base da proposta BWSN é a simulação de eventos de contaminação que podem ocorrer em qualquer nó da rede e a medição da capacidade de uma rede de sensores para detetar atempadamente essa contaminação e prevenir a ocorrência de danos maiores, minimizando a população afetada e o consumo de água contaminada. O transporte da água contaminada ao longo da rede resulta da circulação hidráulica que nela ocorre, e para a qual o EPANET oferece um modelo de cálculo robusto e testado, sendo apenas necessário fornecer dados característicos da rede, quer estáticos relativos aos elementos que compõem a rede, quer dinâmicos em termos de padrões de consumo.

O procedimento de teste preconizado pela BWSN compreende quatro casos, a seguir descritos:

Caso A – O evento contaminante ocorre com uma injeção de 125 l/h de contaminante com uma concentração de 230 000 mg/l durante duas horas. Cada evento de contaminante ocorre num único nó da RAA, e pode começar em qualquer instante com igual probabilidade. Cada solução deve ser testada com um conjunto de eventos e caracterizada com uma estatística apropriada. É assumido que os sensores detetam qualquer concentração de contaminante e que as ações corretivas são todas realizadas de imediato.

Caso B – Apenas difere do caso A por considerar que a injeção de contaminante ocorre durante 10 horas.

Caso C – Difere do caso A, ao considerar que as ações de correção têm um atraso de três horas, o que afetará o cálculo da população afetada e do consumo de água contaminada.

Caso D – Difere do caso A, por considerar que o evento de contaminação ocorre em simultâneo em dois nós da RAA.

#### **3.1.1. Simulação de eventos de contaminação em redes de abastecimento de água com o EPANET**

O EPANET é um programa de computador que permite executar simulações estáticas e dinâmicas do comportamento hidráulico e de qualidade da água em redes de abastecimento de água em pressão. Este simulador foi desenvolvido pela EPA (Loureiro e Coelho, 2002; Rossman, 2000).

Com este programa de simulação é possível obter o caudal em cada tubagem, a pressão em cada nó, a altura de água em cada reservatório de nível variável, e a concentração de compostos químicos através da rede, durante o período de simulação, subdividido em múltiplos passos de cálculo. O EPANET permite ainda determinar a idade e a origem da água em qualquer ponto da rede, quando esta tem mais do que uma fonte de abastecimento.

O EPANET tem sido utilizado para situações de projeto, como, por exemplo, a expansão de redes, assim como para definir estratégias de gestão para melhorar a qualidade de água, por exemplo através da análise do decaimento do cloro residual, ou para gestão de consumos. As estratégias comuns de análise e gestão da qualidade de água incluem:

- Alteração da origem num sistema com várias origens;
- Alteração dos tempos/horários dos sistemas de bombagem e no enchimento /esvaziamento dos reservatórios de nível variável;
- Estudo sobre necessidades de tratamento adicional, nomeadamente análise do decaimento do cloro residual e possível localização de estações de recloração;
- Seleção de tubagens para limpeza e substituição.

A simulação hidráulica com o EPANET requer a modelação hidráulica de RAA, descrevendo-a através de um conjunto de componentes. Uma RAA é necessariamente constituída por tubagens e nós (do inglês *junctions*) que vão unir as tubagens, e ainda por válvulas. Uma RAA tem associado pelo menos um reservatório, o qual pode ter capacidade ilimitada (RNF) ou altura variável (RNV). A distribuição da água é feita sobre pressão, o que pode implicar o uso de bombas em alguns pontos da rede. A representação da rede pode ser efetuada diretamente no programa ou pode ser importada a partir de programas de CAD.

Relativamente à simulação relativa à qualidade da água, o programa permite:

- Modelação do transporte de um constituinte não reativo (ex<sup>o</sup>: traçador) na rede, ao longo do tempo;
- Modelação do transporte, mistura e transformação de um constituinte reativo, à medida que este sofre decaimento (ex<sup>o</sup>: cloro residual) ou crescimento (ex<sup>o</sup>: um subproduto da desinfecção) com o tempo, incluindo reações no seio do escoamento e nas paredes das tubagens;
- Modelação da idade da água ao longo da rede (tempo de percurso);
- Determinação da percentagem do caudal com origem num determinado nó, que atinge qualquer outro nó ao longo do tempo (ex<sup>o</sup>: cálculo da importância relativa de duas origens de água diferentes);
- Aplicação de coeficientes de reação globais, mas que podem ser modificados individualmente, tubagem a tubagem;

- Relacionamento do coeficiente de reação na parede com a rugosidade da tubagem;
- Variação temporal da concentração ou a entrada de massa em qualquer nó da rede;
- Modelação de reservatórios de armazenamento de nível variável como reatores de mistura completa, de escoamento em êmbolo ou ainda de mistura com dois compartimentos.

### 3.1.2. Ferramenta de suporte à implementação

Sendo o objetivo da dissertação a aplicação de metaheurísticas para a localização ótima de sensores que permitam detetar eventos de contaminação, tornou-se necessário usar uma ferramenta adequada para implementação desses métodos, tendo-se optado pelo Matlab. Porém, é necessária interação com o EPANET em alguns pontos da implementação, o que é facilitado pela existência do *EPANET Programmer's Toolkit*<sup>4</sup>, uma biblioteca dinâmica de funções que implementa quase todas as funcionalidades do EPANET e que pode ser ligada com programas desenvolvidos em qualquer linguagem, incluindo o Matlab. Para este último caso, foi ainda útil o uso do *EPANET Matlab Toolkit* (Elíades 2009) que implementa um conjunto de funções Matlab para simplificar a interação com o *EPANET Programmer's Toolkit*.

## 3.2. Funções objetivo a otimizar

### 3.2.1. Definição

As quatro funções objetivo consideradas pela BWSN para a avaliação das redes de sensores foram os a seguir definidos (Ostfeld *et al.* 2008):

#### 1. Tempo de deteção esperado ( $Z_1$ )

Representa o tempo em cada nó que antecede a deteção do contaminante, ou seja, o momento em que a poluição atingiu o nó depois da injeção.

Para uma situação de contaminação particular, o tempo de deteção de um sensor é o tempo que decorre entre o início da contaminação e a primeira identificação da presença do contaminante. O tempo da primeira deteção,  $t_j$ , refere-se à localização do sensor  $j$ . O tempo de deteção de um evento de contaminação por um sensor numa RAA,  $t_d$ , é o menor tempo registado em todos os sensores da RAA:

$$t_d = \min t_j, \quad j = 1, \dots, n^\circ \text{ de sensores}$$

A função objetivo a ser minimizada é o valor associado à probabilidade de distribuição do evento de contaminação

$$Z_1 = E(t_d)$$

<sup>4</sup> <http://www.epa.gov/nrmrl/wswrd/dw/epanet.html>

em que  $E(t_d)$  representa a esperança matemática do tempo de deteção mínimo  $t_d$ . Todos os eventos de contaminação não detetados não são considerados.

## 2. Valor esperado da população afetada antes da deteção da contaminação ( $Z_2$ )

Num contexto específico de contaminação, a população afetada é função da ingestão da massa contaminante. Por sua vez, a quantidade de massa contaminante ingerida depende do tempo de deteção do sensor. São considerados dois pressupostos: 1 - depois de detetada, não é ingerida massa contaminante; 2 - toda a massa ingerida em eventos não detetados não é considerada. Assim, a massa consumida/ingerida, antes da deteção, por um indivíduo, num determinado nó  $i$  da rede é dada por

$$M_i = \varphi \Delta t \sum_{k=1}^N c_{ik} \rho_{ik}$$

em que  $\varphi$  = quantidade média de água consumida por pessoa por dia (l/dia/pessoa) (capitação);  $\Delta t$  = passo temporal da avaliação (dias);  $c_{ik}$  = concentração de contaminante no nó  $i$ , no intervalo de tempo  $k$  (mg/l);  $\rho_{ik}$  = multiplicador da dose para o nó  $i$ , no intervalo de tempo  $k$  (adimensional);  $N$  = número de passos temporais antes da deteção, i.e., o maior inteiro tal que  $N\Delta t \leq t_d$ . A série  $\rho_{ik}$ ,  $k=1, \dots, N$  tem valor médio de 1. Assim,  $\varphi$  é a taxa volumétrica de ingestão e pretende modelar a variação da taxa de ingestão ao longo do dia. Assume-se que a taxa de ingestão varia com a quantidade de água consumida no respetivo nó.

$$\rho_{ik} = \frac{q_{ik}}{\bar{q}_i} \quad \forall k \in N$$

em que  $q_{ik}$  = consumo de água no intervalo de tempo  $k$ , no nó  $i$ , e  $\bar{q}_i$  = consumo médio de água no nó  $i$ .

O modelo dose-resposta dá-nos a probabilidade de uma pessoa que ingere uma massa ( $M_i$ ) ser afetada (i.e., fique doente ou com sintomas)

$$R_i = \Phi \left\{ \beta \log_{10} \left[ \frac{(M_i/W)}{D_{50}} \right] \right\}$$

$R_i$  é a probabilidade [0, 1] de uma pessoa que ingere uma massa contaminante  $M_i$  venha a ficar doente ou com sintomas;  $\beta$  é o parâmetro de declive de Probit (adimensional);  $W$  é a massa corporal média (kg/pessoa);  $D_{50}$  é a dose que resultaria numa probabilidade de 50% de ficar doente ou com sintomas (mg/kg);  $\Phi$  é a função de distribuição normal cumulativa.

A população afetada,  $P_a$ , para uma contaminação particular é calculada como

$$P_a = \sum_{i=1}^V R_i P_i$$

Em que  $P_i$  = população afeta ao nó  $i$ ;  $V$  = número total de nós. A função objetivo a ser minimizada é o valor esperado associado à probabilidade de ocorrência do evento de contaminação

$$Z_2 = E(P_a)$$

Os valores a considerar para o cálculo são:  $\varphi=2$  l/dia;  $\beta=0,34$ ;  $D_{50}=41$  mg/kg;  $W=70$ kg. A capitação considerada é de 300 l/dia/pessoa.

### 3. Consumo esperado de água contaminada antes da detecção ( $Z_3$ )

A função a minimizar é o volume esperado de água contaminada consumida antes da detecção da contaminação no nó:

$$Z_3 = E(V_d)$$

$V_d$  é o somatório da água consumida com a concentração de contaminante superior a um limiar definido  $C$ ;  $E(V_d)$  é o valor esperado de  $V_d$ . Assim, como para a população afetada esperada, pressupõe-se que não é consumida água após a detecção da contaminação e que a água contaminada não detetada não é considerada. O limiar de concentração de contaminante considerado é de 0,3 mg/l.

### 4. Probabilidade de detecção ( $Z_4$ )

Para um determinado conjunto de sensores na RAA (i.e., quantidade e localização dos sensores), a probabilidade de detecção a maximizar é dada por:

$$Z_4 = \frac{1}{S} \sum_{r=1}^S d_r$$

em que  $d_r = 1$  se o evento  $r$  for detetado por um dos sensores, e 0 caso contrário;  $S$  representa a quantidade total de situações de contaminação consideradas.

#### 3.2.2. Abordagem do processo de cálculo das funções objetivo

A implementação da avaliação na BWSN foi realizada em duas etapas.

A primeira, ilustrada na Figura 4, calcula matrizes de avaliação, uma por cada objetivo, para um conjunto de eventos de contaminação. A ferramenta de apoio disponibilizada, e que se refere apenas ao caso A, segue uma de três possíveis abordagens:

- a) Geração de injeção de contaminantes em todos os nós em todos os instantes de tempo, de acordo com o intervalo de tempo usado como “passo” para a simulação de qualidade (5 min.), resultando assim, em 288 eventos para cada nó;

- b) Geração de duas injeções de contaminante em cada nó, em instantes de tempo aleatórios;
- c) Geração aleatória de um número determinado de eventos (ex<sup>o</sup>: 3000) em nós determinados aleatoriamente.

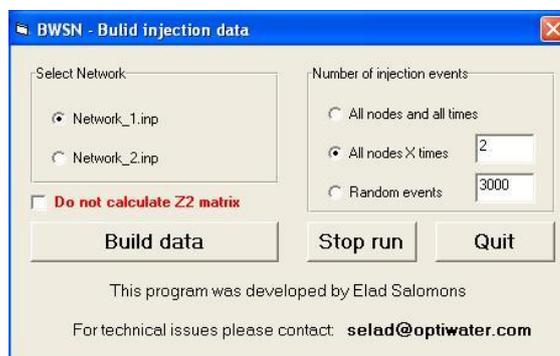


Figura 4. Ferramenta disponibilizada para cálculo das matrizes de avaliação<sup>5</sup>.

As matrizes geradas são guardadas sob a forma de ficheiros de texto para posterior leitura pela ferramenta de avaliação das soluções, e correspondem a cada uma das funções objetivo,  $Z_1$ ,  $Z_2$ ,  $Z_3$  e  $Z_4$ , anteriormente definidos. Cada uma das matrizes possui um número de linhas igual ao número de eventos testado, e um número de colunas igual ao número de nós da rede, correspondendo cada valor numa dada linha à função objetivo em causa. Os valores indicados foram medidos como se o nó correspondente fosse o único a conter um sensor que permitisse interromper o abastecimento de água instantaneamente em caso de deteção.

Não obstante a existência da ferramenta acima referida, procurou-se implementar rotinas de cálculo das matrizes em causa que permitissem a sua aplicação noutros casos ou noutras variantes. A implementação dessas rotinas possui, assim, os seguintes passos:

1. Determinação de um evento de contaminação ou conjunto de eventos, a avaliar em simultâneo;
2. Alteração dos dados de entrada do EPANET de forma a contemplar os eventos de contaminação;
3. Execução da simulação hidráulica e de qualidade através do EPANET e obtenção dos registos de concentração de contaminante em cada intervalo de tempo e dos caudais de procura em cada nó;
4. Cálculo do tempo de deteção ( $Z_1$ ) para cada nó como a diferença entre o instante em que é detetado o primeiro vestígio de contaminação nesse nó, e o instante em que ocorreu a contaminação em si (a ferramenta sinaliza um caso não detetado, atribuindo ao nó o dobro do tempo de simulação, em minutos);

<sup>5</sup> <http://www.water-simulation.com/wsp/bwsn>

- 4.1. Com base no tempo de deteção em cada nó, determina a população afetada antes da deteção da contaminação ( $Z_2$ ), com base nas doses de contaminação verificadas em todos os nós nos instantes que antecederam a deteção;
- 4.2. Com base no tempo de deteção em cada nó, calcula o consumo de água contaminada ( $Z_3$ ), pela soma dos consumos de água nos instantes anteriores à deteção em todos os nós em que a concentração de contaminante é maior ou igual a 0.3 mg/l;
5. Os valores de  $Z_4$  para o evento serão 1 se o nó detetou contaminação ou 0 caso contrário.

O conjunto de passos é repetido para todos os eventos que são simulados.

No EPANET, os perfis ou padrões temporais, de consumo e de injeção de contaminantes, correspondem a vetores de valores entre 0 e 1, também designados fatores multiplicativos. A dimensão dos vetores corresponde ao tempo total de simulação a dividir pelo passo de simulação, que é normalmente de 30 minutos. Os vetores correspondentes aos perfis são relativos a um valor máximo de escala, que no caso da simulação de qualidade da água depende do tipo de injeção considerado, podendo corresponder a:

- Um valor fixo de concentração (*Concentration*), que será atribuído a todos os caudais que entram na rede no nó em causa tais como o caudal que entra na rede a partir de um reservatório ou um caudal que seja definido como uma procura negativa;
- Um acréscimo fixo de massa (*Mass booster*) ao caudal que entra no nó a partir de outros pontos da rede;
- Um acréscimo fixo de concentração (*Flow paced booster*) à resultante da mistura de todos os caudais que entram no nó;
- Um valor fixo de concentração (*Setpoint booster*) que será atribuído ao caudal que sai do nó, caso a concentração resultante seja inferior ao valor estabelecido.

Assim, define-se um perfil de injeção a aplicar a um nó constituído por zeros, exceto nos intervalos correspondentes ao período de injeção determinado. A parametrização da injeção será então efetuada pela aplicação desse perfil no nó selecionado, assim como da quantidade de massa de contaminante injetada e da especificação do tipo de injeção, como sendo um acréscimo fixo de massa.

Para o caso concreto, a injeção de massa será dada por:

$$\frac{125 \text{ l/h} \times 230\,000 \text{ mg/l}}{60 \text{ min/h}} = 479\,167 \text{ mg/min}$$

A ferramenta de apoio disponibilizada pela BWSN define eventos de contaminação em cada intervalo de 5 minutos, correspondendo ao “passo” de simulação da qualidade que está definida por omissão no EPANET. No entanto, todas as definições hidráulicas das redes de teste estão definidas com um passo de simulação de 30 minutos e a própria especificação da BWSN

define que o contaminante é estável, não se verificando degradação, pelo que só ocorre transporte. Logo, não parece justificar-se a simulação de um número tão elevado de eventos. Por outro lado, o EPANET apenas possui uma definição de espaço temporal para os perfis; como todos os perfis de consumo estão com um passo de 30 minutos, não seria possível simular desta forma eventos de contaminação com uma variação de 5 minutos. De qualquer forma, uma vez que foram disponibilizadas pela Universidade de Exeter<sup>6</sup> as matrizes de *benchmarking* para as redes de teste nos quatro casos, optou-se por utilizá-las por razões de comparabilidade, não obstante terem-se obtido resultados coerentes com a rotina implementada.

### 3.2.3. Codificação e avaliação de uma rede de sensores

Uma solução é composta por um conjunto de localizações para sensores, podendo ser representada ou codificada por um vetor de números inteiros. A avaliação da solução é efetuada pela leitura das quatro matrizes anteriormente referidas, da seguinte forma: para o evento correspondente à linha  $i$  das matrizes, o tempo de deteção é dado pelo menor dos valores  $t_{ij}$ , sendo  $j$  o índice das colunas correspondentes aos sensores da solução na matriz  $Z_1$ . Se  $k$  for o índice do nó correspondente ao menor  $t_{ij}$ , a população afetada e o consumo de água contaminada são determinados diretamente a partir da posição correspondente ( $ik$ ) nas matrizes  $Z_2$  e  $Z_3$ . Caso algum dos valores da linha de  $Z_4$ , nas colunas correspondentes à solução, sejam diferentes de zero, o tempo de deteção, a população afetada e o consumo de água contaminada determinados, são contabilizados para a estatística descritiva a usar, tendo-se optado nesta dissertação pela média de todos os valores considerados. A descrição mais detalhada está representada no Algoritmo 1.

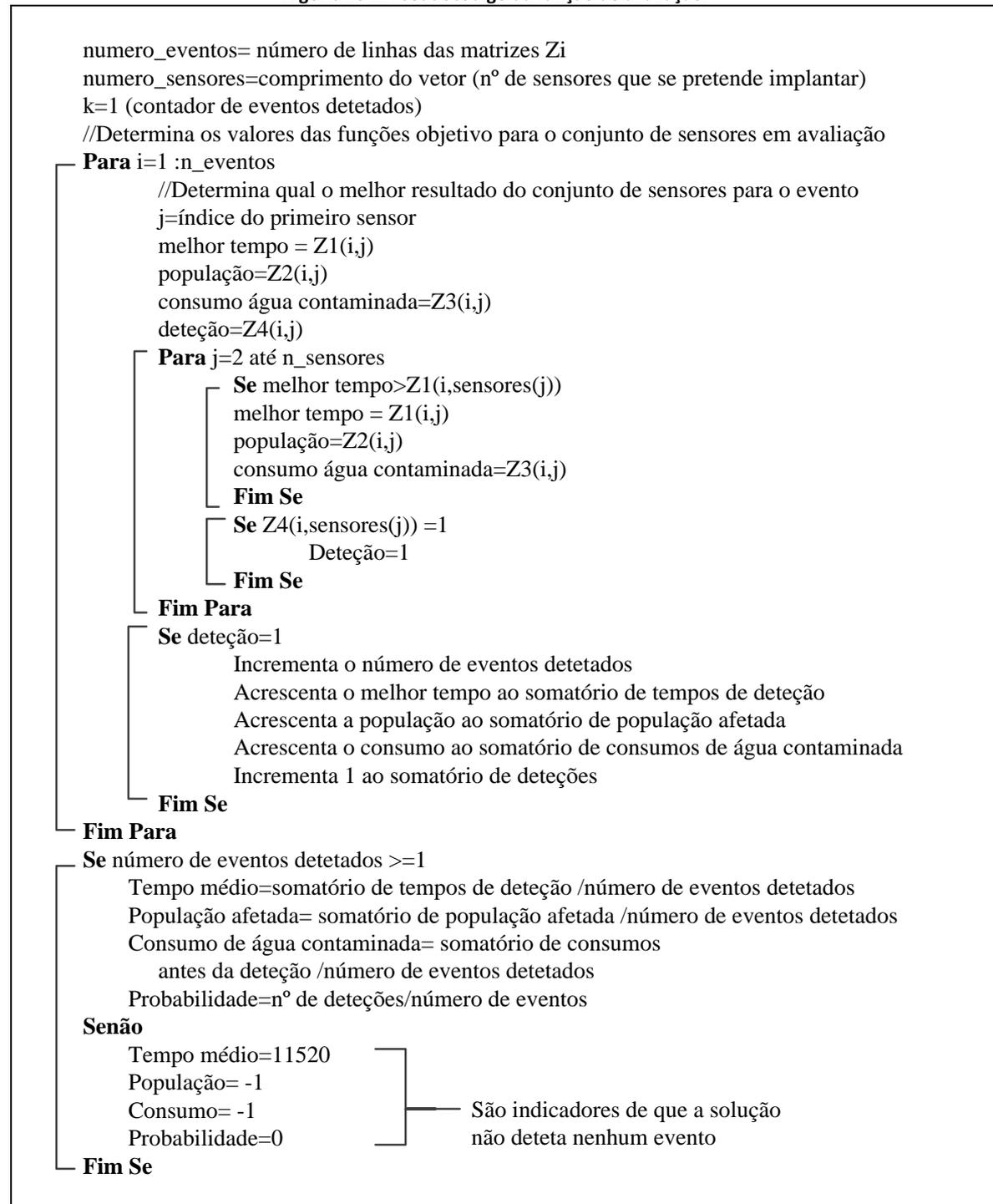
A opção de usar apenas os eventos detetados é justificada por (Ostfeld *et al.*, 2008) como forma de evitar que os sensores sejam essencialmente colocados longe do centro da rede de forma a maximizar deteções, não necessariamente minimizando o tempo de deteção. Para ilustrar este facto, recorde-se que numa RAA com um tamanho significativo, protegida por um número relativamente reduzido de sensores, a não passagem de água contaminada por um dos nós em que os sensores estão instalados tem uma probabilidade elevada, tendo ainda em conta a duração limitada das simulações efetuadas. Vários pontos de injeção de contaminante na periferia da RAA podem ocasionar difusões limitadas dos contaminantes, não atravessando a RAA e por isso nunca sendo detetadas. A inclusão desses efeitos, nos cálculos das funções objetivo, pode levar a comportamentos não intuitivos (Ostfeld *et al.*, 2008). Se a otimização for feita com o intuito de minimizar o tempo de deteção e forem considerados os eventos não detetados (para os quais o tempo de deteção corresponde a uma penalização), serão evitadas todas as soluções que possam levar a não deteções, resultando assim numa aproximação às soluções que otimizam a probabilidade de deteção ( $Z_4$ ), e ocorrerá um afastamento de soluções que levem a tempos de

---

<sup>6</sup> <http://emps.exeter.ac.uk/engineering/research/resources/benchmarks/>

deteção curtos. Relativamente aos objetivos  $Z_2$  e  $Z_3$  esse problema não é tão significativo. No entanto, os promotores da BWSN entenderam manter a lógica coerente com as diversas funções objetivo, sendo essa a razão da introdução da função objetivo adicional que explicita a maximização da probabilidade de deteção, reforçando o carácter multiobjectivo do problema.

**Algoritmo 1. Pseudocódigo da função de avaliação.**



De acordo com (Ostfeld *et al.*, 2005), a avaliação das soluções propostas na BWSN deveria ter sido efetuada com base numa amostragem dos eventos de contaminação, ou seja, com base num conjunto limitado e escolhido aleatoriamente de linhas das matrizes  $Z_n$  ( $n=1,2,3,4$ ). Os dados que caracterizam a amostragem efetuada e a estatística descritiva usada não foram documentados, tendo sido disponibilizada uma ferramenta de avaliação, em conjunto com a ferramenta de apoio apresentada anteriormente (Figura 4), para permitir comparar as soluções de forma equitativa. No entanto, verificou-se através da rotina implementada nesta dissertação, que os resultados listados em (Ostfeld *et al.*, 2008) foram obtidos com a média de todos os eventos simulados, tendo-se desta forma validado a rotina implementada.

### 3.2.4. Pesquisa do espaço de soluções

Relativamente à pesquisa do espaço de soluções, optou-se por respeitar a estrutura da RAA, ou seja, determinando a próxima solução a avaliar, no seguimento de uma pesquisa local através da troca de um nó por outro adjacente. Assim, tornou-se necessário criar previamente uma estrutura de dados que registasse os nós adjacentes a cada nó da rede, de acordo com o pseudocódigo descrito no Algoritmo 2.

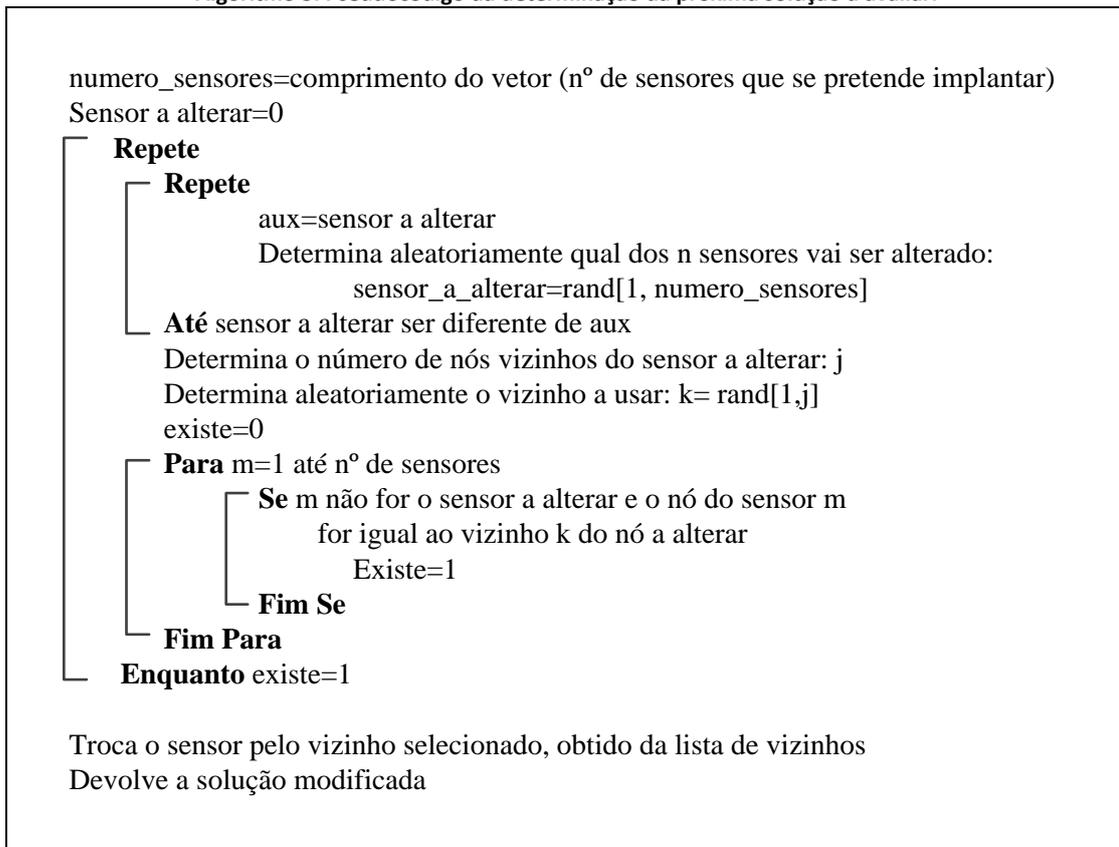
**Algoritmo 2. Pseudocódigo para registo dos nós adjacentes.**

```
Obtém do Epanet o número de tubos e o número de nós
//Inicializa a lista de nós vizinhos
Para o nó i=1 até ao nº total de nós
    Atribui zero vizinhos a todos os nós
Fim para

Para cada tubo i=1 até ao número total de tubos
    Obtém do Epanet o nó de origem e o nó de destino
    origem=nó de origem
    destino=nó de destino
    Acrescenta o nó de fim como nó vizinho do nó de origem
    Acrescenta o nó de origem como nó vizinho do nó de fim
Fim para

Sai do Epanet
```

Conhecendo os nós adjacentes de cada nó da rede, pode proceder-se então à determinação de uma solução próxima, escolhendo aleatoriamente um dos sensores a ser modificado, e depois determinando aleatoriamente qual dos nós adjacentes ao nó onde o sensor estava localizado será usado, garantindo apenas que não se escolhe um nó que já pertença à solução. O pseudocódigo do processo encontra-se descrito no Algoritmo 3.

**Algoritmo 3. Pseudocódigo da determinação da próxima solução a avaliar.**

### 3.3. Abordagem de otimização

O problema em análise, corresponde a um tipo de problemas de muito difícil resolução através de métodos de otimização matemáticos exatos, dada a natureza combinatoria das soluções possíveis, e as inúmeras relações não lineares entre variáveis e os diferentes parâmetros da rede. As propostas dos concorrentes à BWSN, citadas em (Ostfeld *et al.*, 2008), refletem este efeito, como se pode verificar pelo Quadro 2 (ver 2.3), onde dez das quinze propostas foram baseadas em metaheurísticas, consideradas procedimentos computacionais capazes de resolver o problema de determinar soluções suficientemente boas em tempo razoável (Gaspar Cunha *et al.*, 2012). De facto, para uma rede com  $n$  nós, onde se pretenda colocar  $m$  sensores, existem  ${}^nC_m$  soluções possíveis, o que resulta num número muito elevado mesmo para uma pequena RAA, como é o caso da Rede 1 de teste com 129 nós, em que a colocação de cinco sensores representa aproximadamente 275 000 000 combinações possíveis.

Tal como explicado na secção 2.4, nesta dissertação optou-se por usar como abordagens as metaheurísticas *Simulated Annealing* e Algoritmos Genéticos.

### 3.3.1. *Simulated Annealing*

O algoritmo Simulated Annealing (SA) baseia-se numa analogia termodinâmica com base no processo de cristalização dos metais, sendo um processo de pesquisa aleatória em que iterativamente se procura atingir o ótimo, mas permitindo esporadicamente aceitar soluções piores que a melhor a solução até então obtida de forma a evitar convergência prematura em ótimos locais. Estas ocorrências obedecem a uma probabilidade que é função de um parâmetro de controlo decrescente com o tempo, de forma análoga ao arrefecimento dos metais, tal como se mostra no Algoritmo 4.

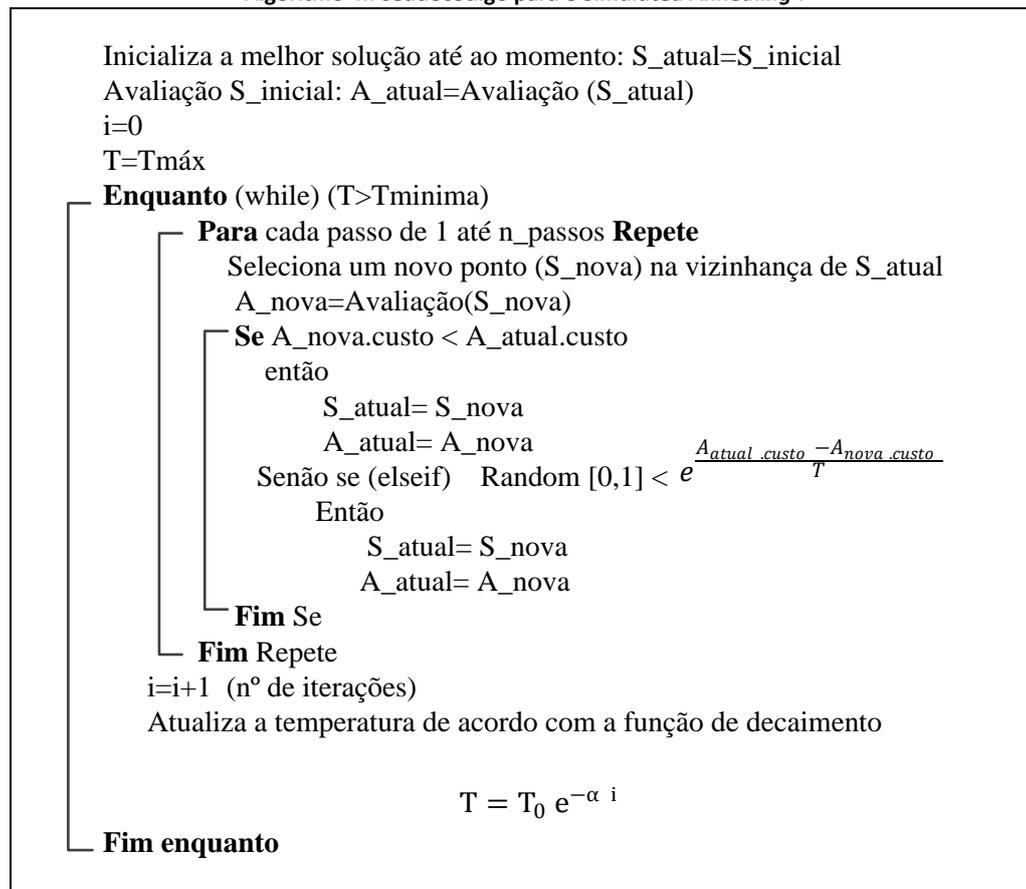
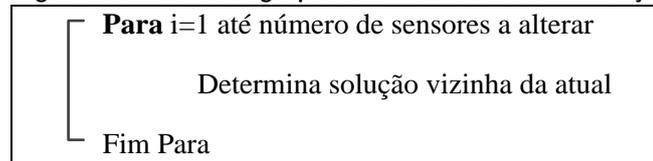
A definição de uma função de decaimento da temperatura (*cooling schedule*) possui também diferentes opções, existindo várias variantes. Kirkpatrick (1983) utiliza um arrefecimento proporcional, mas por exemplo Suman e Kumar (2005) referem outras hipóteses, nomeadamente logarítmica, *Cauchy*, e exponencial, sendo esta última considerada a mais rápida a levar à convergência para o ótimo global. Após uma experimentação inicial entre a forma proporcional e a exponencial, optou-se por esta última, por produzir melhores resultados. Assim, a expressão utilizada nesta dissertação foi a seguinte:

$$T = T_0 e^{-\alpha i} \quad (1)$$

Em que  $T_0$  é a temperatura inicial,  $\alpha$  é o fator de decaimento e  $i$  é a iteração corrente.

Os elementos essenciais que precisam de ser especificados para a aplicação em causa são a função de avaliação de cada solução e o procedimento que permite explorar o espaço de soluções através da determinação de soluções vizinhas, ou seja, a estrutura de vizinhança.

Neste caso, a função de avaliação da solução teve como base a implementação genérica referida no ponto 3.2.3. O mesmo aconteceu com a função que define a estrutura de vizinhança para a pesquisa no espaço de soluções, que se baseou na definição efetuada no ponto 3.3, mas aqui admitindo a possibilidade de trocar um ou dois sensores conforme representado no Algoritmo 5

**Algoritmo 4. Pseudocódigo para o Simulated Annealing .****Algoritmo 5. Pseudocódigo que define a estrutura de vizinhança.**

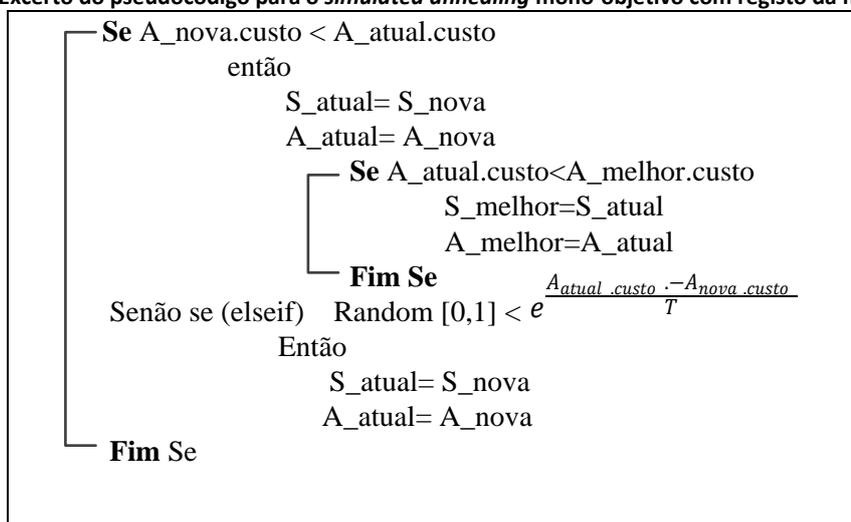
### 3.3.1.1. Abordagem mono-objetivo

#### **Implementação mono-objetivo base**

Para operacionalizar a otimização de cada uma das funções objetivo (FO), isoladamente, foi definida uma variável global que permite selecionar a FO de interesse. Assim, a função de avaliação devolve um valor de custo que em função da FO selecionada, será  $Z_1$ ,  $Z_2$ ,  $Z_3$  ou  $(1-Z_4)$  como valor a minimizar. Independentemente do custo, a função devolve também os valores individuais das FO, para poder caracterizar as soluções obtidas. Foi ainda introduzido o registo da melhor solução global, independentemente da variação da solução corrente em cada iteração do método, para evitar que uma “boa solução” possa ser perdida em definitivo ao aceitar

degradar a solução corrente como forma de evitar os ótimos locais, como se pode verificar no excerto do pseudocódigo ilustrado no Algoritmo 6.

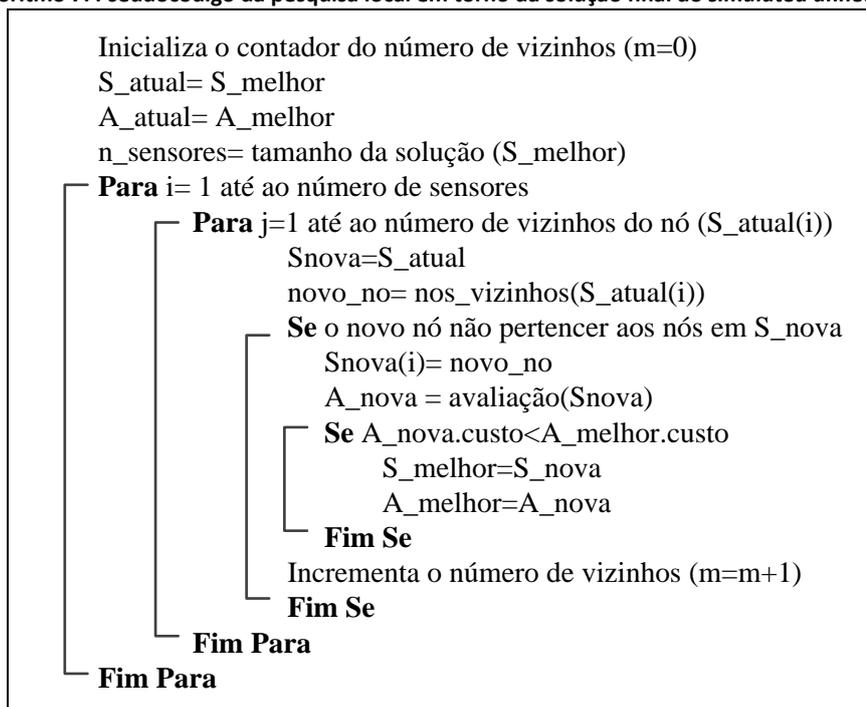
**Algoritmo 6. Excerto do pseudocódigo para o *simulated annealing* mono-objetivo com registo da melhor solução.**



#### **Implementação mono-objetivo base com pesquisa local**

A fim de garantir que a solução final não fica acidentalmente ao lado de uma solução melhor que não chegou a ser detetada, foi acrescentada uma fase de pesquisa local, destinada a analisar sistematicamente todas as soluções vizinhas dessa solução, de acordo com a estrutura de vizinhança anteriormente definida, mas em que apenas se considera a troca de um sensor, como se mostra no pseudocódigo do Algoritmo 7.

**Algoritmo 7. Pseudocódigo da pesquisa local em torno da solução final do *simulated annealing*.**



### 3.3.1.2. Abordagem multiobjetivo

A adaptação de *simulated annealing* para otimização multiobjetivo decorreu da sua simplicidade e da capacidade de produzir um conjunto de soluções ótimas de Pareto com pouco esforço computacional, sendo pouco suscetível à forma da frente de Pareto, dois problemas comuns para as técnicas de programação matemática (Suman e Kumar, 2005).

De acordo com estes autores, diversas variantes do SA para multiobjetivo têm vindo a ser propostas por vários investigadores, sendo referidas as abordagens SMOSA (Suppapitnarm e Parks, 1999), UMOSA (Ulungu *et al.*, 1998), PSA (Czyzak e Jaszkievicz, 1997), e ainda a WMOSA (Suman, 2002) e PDMOSA (Suman, 2003). Uma outra abordagem, designada AMOSA pode ser encontrada em (Bandyopadhyay *et al.*, 2008).

Nesta dissertação optou-se pela abordagem SMOSA, consistindo as principais modificações ao algoritmo SA nos seguintes aspetos:

- O processo de pesquisa é baseado no conceito de não dominância, em vez de ser uma comparação simples entre alternativas;
- As soluções não dominadas vão sendo arquivadas, sendo a solução inicial em cada iteração escolhida aleatoriamente de entre estas;
- Em cada passo, a probabilidade de aceitação de uma solução dominada é calculada com base na agregação das diferenças relativamente a cada objetivo e num valor de temperatura específico.

$$P = \min \left( 1, \prod_{i=1}^N \exp \left\{ \frac{-\Delta S_i}{T_i} \right\} \right) \quad (2)$$

em que  $N$  é o número de funções objetivo,  $\Delta S_i$  é a diferença para a  $FO_i$  entre a solução nova e a solução atual,  $T_i$  é o valor da temperatura para a  $FO_i$ .

Um dos aspetos que mais varia nas diferentes abordagens de SA multiobjetivo consiste na função de probabilidade de aceitação. Antunes *et al.* (2011) definem outras alternativas para esta função, nomeadamente:

Curva logística:

$$P = \frac{2}{1 + e^{\frac{\Delta}{T_k}}} \quad (3)$$

Linear escalar:

$$P = \min \left( 1, e^{\frac{-\Delta}{T_k}} \right) \quad (4)$$

Chebyshev (forte):

$$P = \min \left[ 1, \min_j \left( 1, e^{\frac{-\omega_j \delta_j}{T_k}} \right) \right] \quad (5)$$

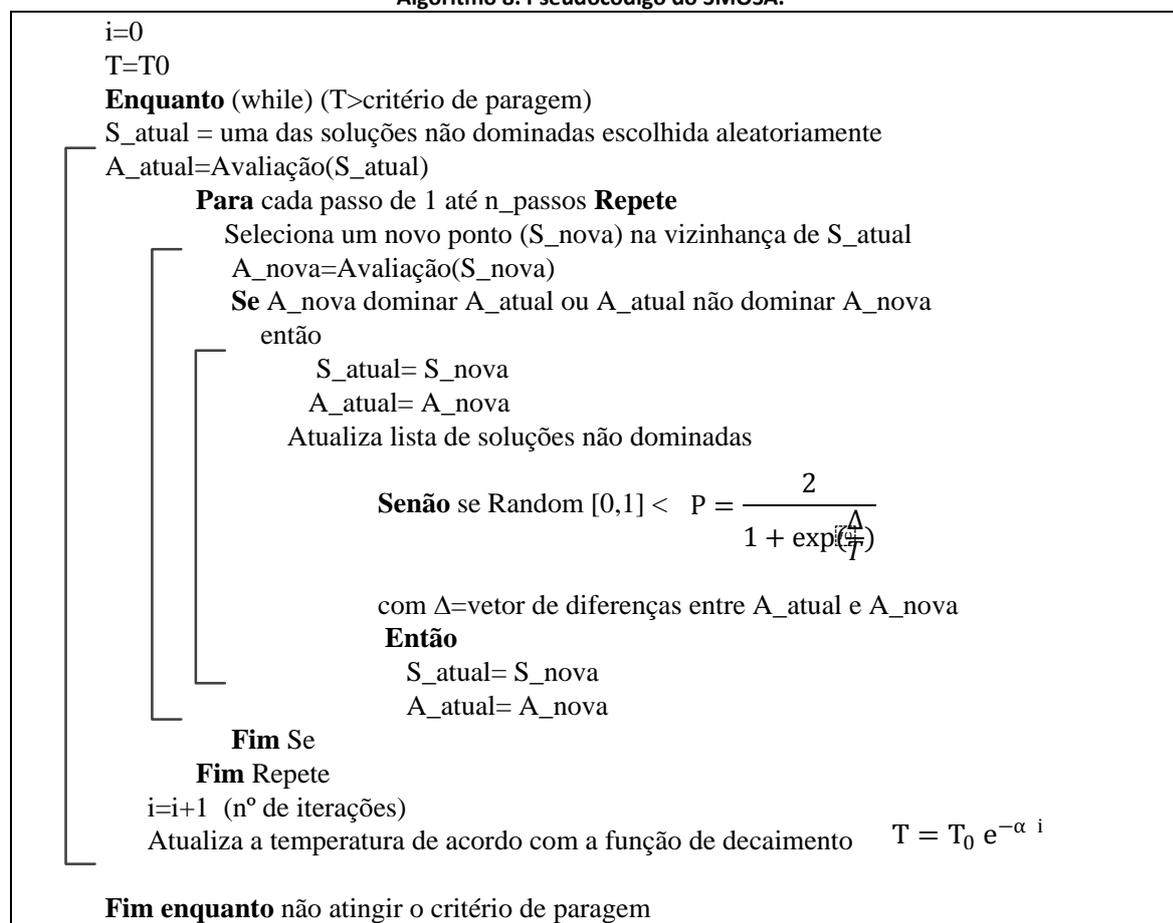
Fraco:

$$P = \min \left[ 1, \max_j \left( e^{\frac{-\omega_j \delta_j}{T_k}} \right) \right] \quad (6)$$

sendo  $T_k$  a temperatura na iteração  $k$ , e  $\Delta$  a soma ponderada das diferenças entre FO.

A atualização da lista de soluções não dominadas, que é essencial ao funcionamento do algoritmo, corresponde a comparar a nova solução com todas as soluções anteriormente arquivadas, verificando se já pertence à lista e verificando a relação de dominância. Caso a nova solução não exista ainda e domine alguma das existentes, a solução dominada é eliminada e a nova solução é inserida na lista. Caso a nova solução não domine nem seja dominada por nenhuma das outras existentes no arquivo, é também inserida. O pseudocódigo da função de atualização da lista encontra-se no Apêndice A.

**Algoritmo 8. Pseudocódigo do SMOSA.**



### 3.3.2. Algoritmo Genético

O algoritmo é iniciado com a geração de um conjunto de soluções aleatórias (população inicial). Cada elemento da população é avaliado individualmente em termos de desempenho, determinado através da função objetivo que se pretende otimizar. De seguida, é selecionado um conjunto de elementos designados progenitores, por torneio binário entre pares de soluções escolhidas aleatoriamente entre a população. A partir do conjunto de progenitores é então determinada uma nova geração de soluções, através da aplicação de operadores genéticos, tais como, recombinação ou mutação, de acordo com probabilidades especificadas para cada um. Os elementos resultantes formam uma nova geração da população que se espera conterá soluções melhores que a que lhe deu origem. Assim, ao fim de vários ciclos de evolução, os resultados tenderão a incluir a solução ótima. O pseudocódigo deste algoritmo pode ser consultado no Apêndice A.

A codificação e avaliação das soluções é efetuada de acordo com a descrição efetuada na secção 3.2.3. Os restantes aspetos específicos da implementação consistem no processo de recombinação, em que se optou por um ponto de corte fixo, no gene central do cromossoma, tal como ilustrado na Figura 5.

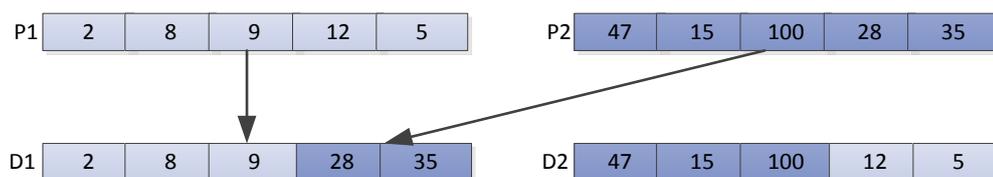


Figura 5. Exemplo de recombinação com um ponto de corte.

O processo de mutação, tal como a função de vizinhança do SA, implementa o processo descrito no Algoritmo 3 e que está ilustrado na Figura 6.

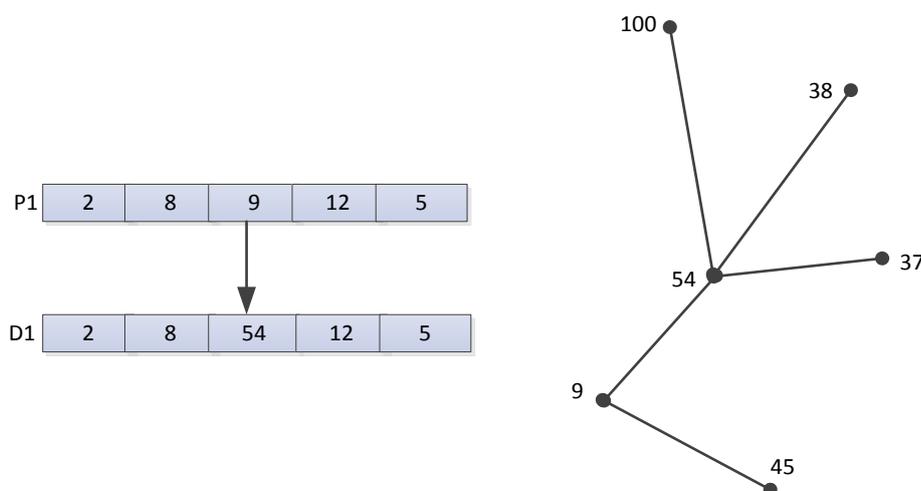


Figura 6. Ilustração do processo de mutação implementado.

### 3.3.2.1. Função *fitness* mono-objetivo

A *fitness* de cada elemento corresponde ao valor da FO escolhida usando uma implementação semelhante à realizada para a função de avaliação em SA (*vide* 3.3.1.1), com uma variável global para definir qual das funções objetivo deve ser considerada. A avaliação da *fitness* é efetuada imediatamente após a criação do elemento, seja na geração aleatória inicial, seja após a aplicação do operador de recombinação e/ou de mutação, ficando o valor da *fitness* guardado junto com a solução.

### 3.3.2.2. Introdução de elitismo

Por vezes os melhores elementos da população podem produzir descendentes de qualidade inferior e, por isso, as melhores soluções podem ser perdidas de uma geração para a seguinte. Frequentemente o AG irá redescobrir estas soluções em gerações subsequentes, não existindo, no entanto, garantia que isso aconteça. Para evitar este problema pode-se usar uma estratégia designada por elitismo, em que se deixam passar intocados para a geração seguinte os melhores elementos da geração anterior, numa percentagem pré-estabelecida. Esta estratégia tem como objetivo acelerar a convergência (Deb *et al.*, 2002).

De forma a poder experimentar diferentes níveis de elitismo implementou-se esta estratégia de um modo que permite definir o tamanho do conjunto elite, ou seja, o número dos melhores elementos da geração anterior que passam à geração seguinte sem sofrerem a atuação dos operadores genéticos, como uma percentagem da população total. Os restantes elementos da nova geração são então gerados como habitualmente, através dos operadores de recombinação e mutação de acordo com as probabilidades definidas, como se ilustra no Algoritmo 9.

**Algoritmo 9. Pseudocódigo do AG com elitismo (1ª versão).**

```

geração =0
Gera um conjunto de soluções aleatórias;
Enquanto a geração < max_geracoes
    Incrementa geração: geracao = geracao + 1
    Determina número de elementos da elite
    Ordena a população pela fitness, para que os primeiros sejam a elite
    Seleciona os progenitores da próxima geração:
        Progenitores = selecao(P, tamanho_pop_elite);
    Gera a próxima geração a partir dos progenitores, mantendo a elite:
        Pnova = união da elite com a nova geração obtida por:
            produz_nova_geracao(Progenitores, Prob_Mutacao, Prob_Crossover);
Fim Enquanto
Determina a melhor solução da geração: solucao = melhor(P);

```

### 3.3.2.3. Compensação de perda de diversidade

Um problema típico dos AG convencionais consiste na possibilidade de convergência prematura (Dai *et al.*, 2006). Uma das causas desse fenómeno reside na perda de diversidade genética na população de geração para geração. Uma forma de manter a diversidade na população, e simultaneamente a capacidade de convergência dos algoritmos genéticos, consiste no uso de probabilidades dinâmicas de recombinação e mutação (Srinivas e Patnaik, 1994).

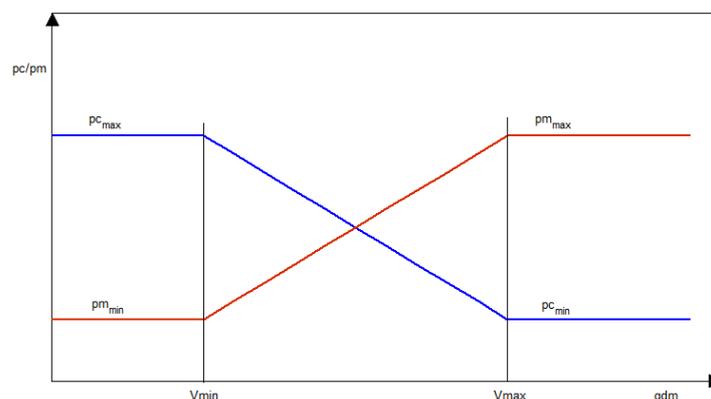


Figura 7. Interpoilação linear entre limites. pc- probabilidade de crossver, pm – probabilidade de mutação (Vasconcelos *et al.*, 2001).

Para permitir a implementação deste tipo de variação dinâmica das probabilidades foi necessário encontrar uma forma de medição da diversidade genética, tendo-se adotado o indicador sugerido por (Vasconcelos *et al.*, 2001), o rácio entre o valor médio e o valor máximo da *fitness* em cada geração. Quando este valor for 1 indica que todos os elementos têm o mesmo código genético, logo a diversidade genética estará no mínimo. Nesta circunstância, a probabilidade de mutação deve ter um valor máximo, que decresce à medida que a diversidade genética aumenta.

### 3.3.2.4. Abordagem multiobjetivo

Os algoritmos genéticos permitem fazer evoluir simultaneamente um conjunto de soluções, a população. São, por isso, particularmente aptos para a resolução de problemas de otimização multiobjetivo, nos quais se pretende, em geral, identificar uma frente de soluções não dominadas, tendo dado origem a várias abordagens para a obtenção numa única execução do algoritmo de uma população de soluções que representam diferentes compromissos entre as funções objetivo do problema (Takahashi *et al.*, 2012).

Na década de 1990, vários tipos de algoritmos genéticos foram sugeridos para resolver problemas de otimização multiobjetivo, nomeadamente o MOGA de Fonseca e Fleming (1993); NSGA de Srinivas e Deb (1994) e o NPGA de Horn *et al.* (1994). Mais recentemente, o NSGA-II de Deb *et al.* (2002) e o SPEA2 de (2001) têm vindo a receber grande destaque, e, segundo Takahashi *et al.* (2012), são em grande parte responsáveis pela rápida popularização da otimização multiobjetivo com algoritmos genéticos.

Para a resolução do problema multiobjetivo em estudo foi decidido usar o NSGA-II por ser uma das abordagens metaheurísticas multiobjetivo, baseadas em algoritmos genéticos, mais usadas e com melhor desempenho computacional (Peng *et al.*, 2009; Takahashi *et al.*, 2012). O NSGA-II implementa uma estratégia elitista, uma métrica capaz de medir a diversidade genética da população e tem uma complexidade não muito elevada.

A principal diferença entre o NSGA-II e as abordagens de algoritmos genéticos mono-objetivo, consiste na seleção feita durante o torneio binário. Num contexto mono-objetivo, a decisão sobre qual dos participantes do torneio binário é o selecionado depende exclusivamente do valor da *fitness*. No NSGA-II proposto por Deb *et al.* (2002), a decisão depende em primeiro lugar da existência de uma relação de dominância que previamente classificou todas as soluções em frentes de Pareto de diferente ordem (*rank*) e, em segundo lugar, no caso de pertencerem à mesma frente, pela “distância à multidão” (*crowding distance*), uma métrica destinada a proporcionar o máximo de diversidade genética da população.

A classificação de todas as soluções em frentes de Pareto de ordem  $n$ , é feito através de um procedimento designado por “*fast nondominated sorting*”, em que numa primeira fase se determinam os conjuntos de soluções dominadas por cada solução da população, e o número de soluções que dominam cada solução. As soluções são depois agrupadas de acordo com este último índice que definirá o número de ordem da frente de Pareto, sendo a frente de ordem 1 composta pelas soluções que não são dominadas por nenhuma outra, como ilustrado na Figura 8.

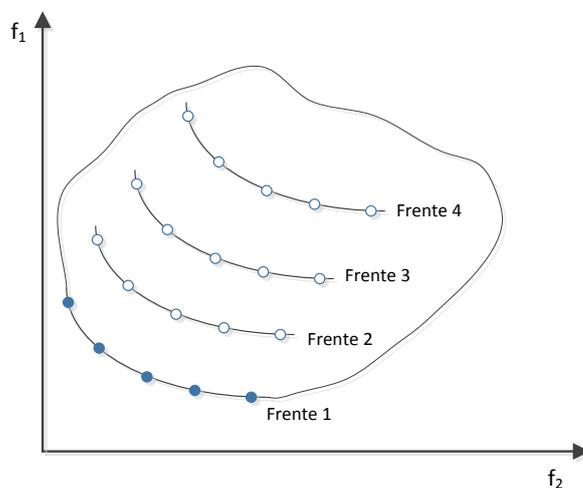


Figura 8. Representação da classificação em frentes de Pareto dos elementos da população.

Conforme o nome do próprio algoritmo indica, uma característica importante no NSGA-II consiste na introdução de elitismo, que neste caso é implementada de uma forma em que a população de onde são extraídos os progenitores e as soluções descendentes, por recombinação ou mutação, competem pela passagem à geração seguinte. Assim, em primeiro lugar a população original é duplicada com a união dos descendentes, gerados por recombinação ou por mutação. A população resultante é depois classificada em frentes de Pareto, através do “*fast nondominated sorting*”, depois ordenada pela ordem crescente das frentes de Pareto e dentro

de cada frente de Pareto, por ordem decrescente do “*crowding distance*”. Finalmente, são selecionados os primeiros  $n$  elementos da população, sendo  $n$  o seu tamanho original. O procedimento está ilustrado na Figura 9.

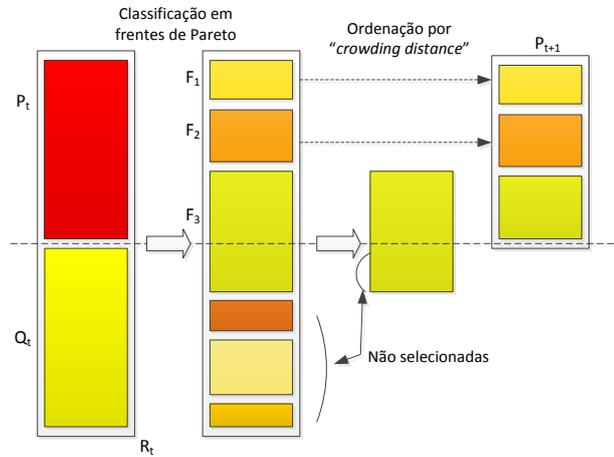


Figura 9. Procedimento do NSGA-II (Deb *et al.*, 2002).

Os pseudocódigos referentes às diferentes partes da implementação NSGA-II encontram-se no Apêndice A.



## 4. ESTUDOS DE CASO

### 4.1. Descrição das redes em estudo

A BWSN forneceu dois modelos de rede para servirem de teste, representando redes reais descaracterizadas. A Rede 1 encontra-se representada na Figura 10, e é composta por 129 nós, uma captação, 2 reservatórios, 168 tubos, 2 sistemas de bombagem, 8 válvulas, sendo sujeita a dois perfis de procura. O sistema foi simulado para um tempo total de 96 horas.

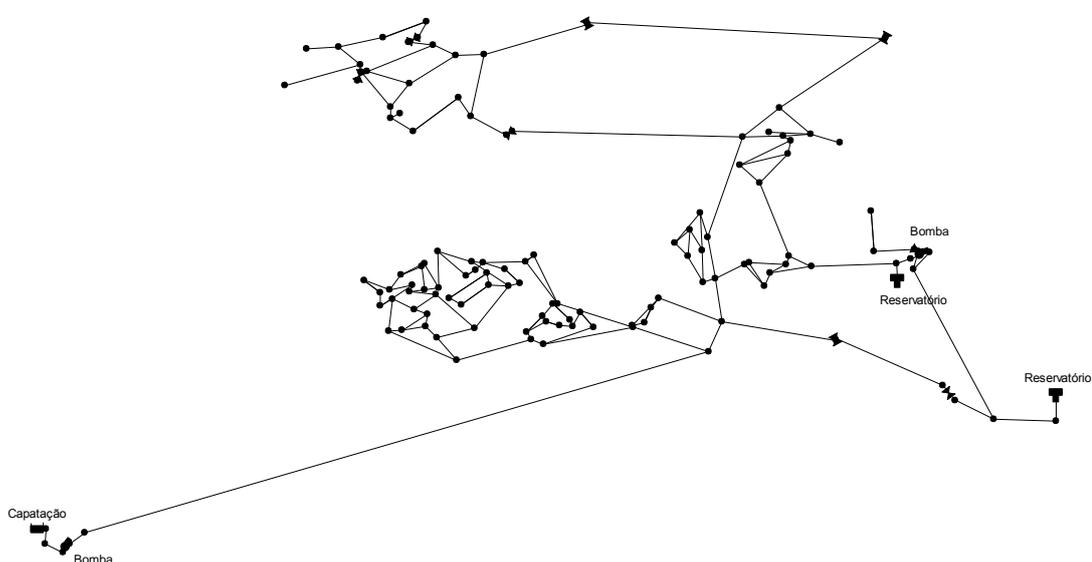


Figura 10. Distribuição espacial da Rede 1 no EPANET.

A Figura 11 mostra a Rede 2. Esta é composta por 12 523 nós, 2 captações, 2 reservatórios, 14 822 tubos, 4 sistemas de bombagem, 5 válvulas sendo sujeito a quatro perfis de procura. O sistema foi simulado para um tempo total de 48 horas.

As redes de sensores a implantar eram compostas por um número fixo de 5 e de 20 sensores.

Sendo os perfis de consumo e bombagem fornecidos com as redes correspondentes a um dia, que se repete, consideram-se apenas eventos de contaminação durante as primeiras 24 horas. No entanto, a simulação é efetuada para um período superior para poder acautelar o efeito do transporte da água ao longo da rede.

O Apêndice B contém as listas de correspondência entre a etiqueta dos nós (ID) que é usada na BWSN para identificar os elementos da solução, e o índice numérico dos nós, que por razões operacionais se optou por usar nesta dissertação.

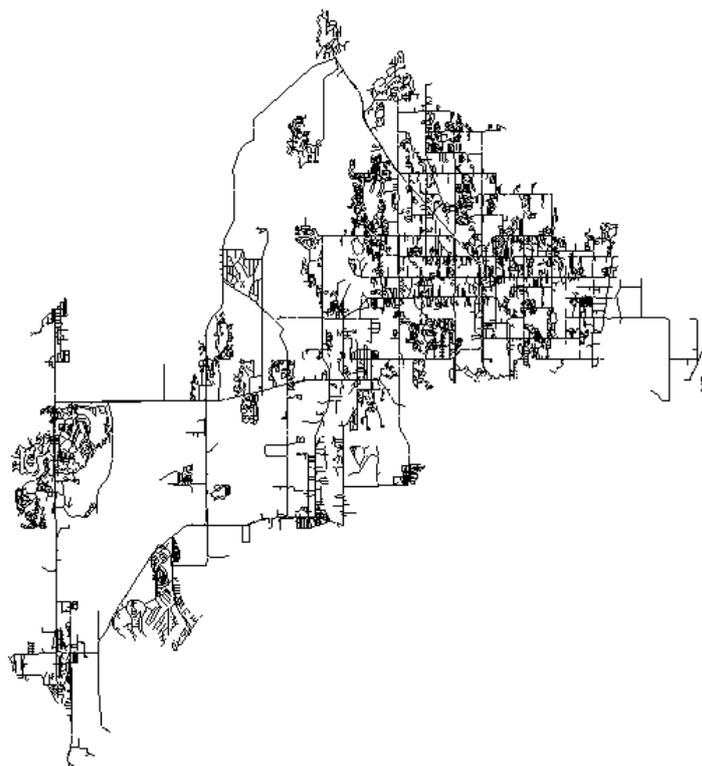


Figura 11. Distribuição espacial da Rede 2 no EPANET

## 4.2. Aplicação do *simulated annealing* para a determinação da rede de sensores

### 4.2.1. Otimização mono-objetivo com a implementação base

De acordo com (Pirlot, 1996), não existem regras de aceitação generalizadas para a determinação dos valores dos parâmetros a utilizar na heurística SA, sendo recomendado uma procura sistemática em torno de uma escolha inicial. Relativamente à temperatura inicial,  $T_0$ , o processo deve começar com uma temperatura suficientemente alta, selecionando  $T_0$  de forma a garantir uma certa probabilidade  $p_0$  de aceitar piores soluções no início. Pirlot (1996) sugere  $p_0 = 0.9$ , mas refere que alguns autores obtiveram melhores resultados com valores de  $p_0$  bastante reduzidos (da ordem de 0.4). Relativamente ao fator de decaimento e ao número de passos, estes fatores estão correlacionados com o tempo de computação. No que diz respeito ao fator de decaimento, não se encontrou nenhuma sugestão de procedimento para a sua determinação, tendo-se encontrado apenas alguns valores típicos, nomeadamente 0.05 para a fórmula de decaimento adotada (*vide* 3.3.1). Quanto ao número de passos, Pirlot (1996) sugere um múltiplo do tamanho médio da vizinhança, citando um autor que refere dezasseis vezes o tamanho da vizinhança.

Para estimar um conjunto inicial de valores a usar executaram-se 100 avaliações de soluções geradas aleatoriamente para determinar a gama de variação de cada função, de acordo com o Quadro 3.

**Quadro 3. Determinação de possíveis valores de temperatura inicial ( $T_0$ ).**

Função	$\Delta f$	Probabilidade de aceitação	
		0.9	0.4
		$T_0$	
$Z_1$	1094	10383.4	1193.9
$Z_2$	872	8276.3	951.7
$Z_3$	83438	791928.5	91060.6
$Z_4$	0.432	4.1	0.5

Em face dos resultados obtidos, adotaram-se as combinações de parâmetros que estão listados no Quadro 4.

Para testar as combinações de parâmetros, executou-se 11 vezes o algoritmo de otimização, para cada uma das estruturas de vizinhança (trocando 1 ou 2 sensores), com vista a poder apresentar, para além dos resultados da otimização, estatísticas relativas à taxa de sucesso, ou seja, o número de vezes que se atingiu a melhor solução, assim como o tempo de execução correspondente.

**Quadro 4. Resumo do conjunto de parâmetros usados no *Simulated Annealing*.**

Função	$T_0$	$\alpha$	n_passos
$Z_1$	500	0.01	50
	1000		100
	10000	0.05	200
$Z_2$	500	0.05	50
	1000	0.01	100
	10000		200
$Z_3$	1000	0.01	50
	10000	0.05	100
	100000		200
$Z_4$	0.1	0.05	50
	0.5	0.01	100
	5		200

Os melhores valores obtidos para as funções objetivo encontram-se listados no Quadro 5, correspondendo cada caso a uma solução única. A análise detalhada à localização dos sensores resultante das soluções permite verificar que os nós 32, 35 e 38 são comuns à melhor solução relativamente a  $Z_1$  e à melhor solução relativamente a  $Z_3$ . Embora não se tenham verificado diferenças significativas, os melhores resultados foram obtidos com a estrutura de vizinhança em que se troca apenas um sensor. De notar que nos casos da otimização de  $Z_3$  e  $Z_4$  foi ainda necessária alguma experimentação adicional para finalmente se optar pelos valores do Quadro 6.

As funções objetivo  $Z_2$  e  $Z_3$  apresentam uma correlação significativa de aproximadamente 0.795, calculada pela expressão (4), em que  $\bar{x}, \bar{y}$  representam a média dos resultados obtidos para cada uma das duas FO.

$$\text{Correl}(X, Y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} \quad (4)$$

Quadro 5. Soluções que otimizam cada função objetivo.

Função otimizada	Valor da solução				Sensores				
	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [galões]	Z <sub>4</sub> [%]					
Z <sub>1</sub>	<b>151.71</b>	108	2422.855	20.33%	32	35	38	41	50
Z <sub>2</sub>	196.57	<b>59</b>	1060.832	22.82%	22	27	31	38	55
Z <sub>3</sub>	192.65	67	<b>619.114</b>	24.86%	21	25	32	35	38
Z <sub>4</sub>	1256.88	670	43041.836	<b>83.92%</b>	11	46	84	101	125

Os ótimos identificados para cada uma das FO evidenciam a necessidade de uma análise multiobjetivo. A solução que otimiza  $Z_1$  corresponde a uma probabilidade de detecção ( $Z_4$ ) muito baixa. Por sua vez os ótimos de  $Z_2$  e  $Z_3$  são obtidos em soluções com resultados semelhantes em  $Z_1$  e  $Z_4$ , mas com diferenças não totalmente desprezáveis em  $Z_3$  e  $Z_2$  respectivamente. O ótimo de  $Z_4$  é obtido numa solução com valores muito maus nos outros três objetivos.

Os resultados estatísticos relativos aos conjuntos de parâmetros que permitiram obter o ótimo em cada função objetivo encontram-se resumidos no Quadro 6. Para cada função é indicado o conjunto de parâmetros que obteve maior taxa de sucesso, ou seja, com o qual mais vezes se atingiu o ótimo da função. As estatísticas relativas às soluções obtidas para a otimização de cada FO podem ser vistas no Apêndice C. Os resultados indicados foram obtidos num computador com um processador Intel Core i7-3939K a 3.2GHz com 32GB de memória RAM.

Quadro 6. Resumo dos parâmetros que permitiram obter a melhor solução.

Função	Conjunto de Parâmetros					
	T <sub>0</sub>	α	n_passos	Critério de paragem	Taxa de sucesso	T execução [s]
Z <sub>1</sub>	1000	0.01	200	10 <sup>-2</sup>	82%	1632
Z <sub>2</sub>	500	0.01	50		82%	392
Z <sub>3</sub>	1000	0.05	200	10 <sup>-3</sup>	100%	393
Z <sub>4</sub>	1	0.01	100		55%	522

A Figura 12 mostra a evolução da melhor solução obtida para cada FO em função da temperatura, ilustrando também a forma como a solução inicial em cada iteração evolui, de acordo com a probabilidade decrescente de aceitar soluções piores.

Os resultados obtidos demonstram a convergência para a melhor solução obtida a partir de uma temperatura de 1 (entre 600 e 700 iterações) para as funções  $Z_1$  e  $Z_2$ , a partir de uma temperatura de 10 na função  $Z_3$  (94 iterações), e de uma temperatura de 0.010 para a função  $Z_4$  (691 iterações). As soluções em cada iteração ainda oscilam em alguns casos para temperaturas inferiores às acima indicadas, mas a melhor solução guardada até ao momento já não sofre alterações. Assim, podem obter-se tempos de execução inferiores aos indicados no Quadro 6 se os critérios de paragem forem revistos adequadamente.

Verificou-se ainda que em 792 execuções do processo de otimização, se obtiveram em 2 casos valores de  $Z_1$ ,  $Z_2$  e  $Z_3$  bastante inferiores às soluções aqui qualificadas como melhores. No entanto, verificou-se que correspondiam a situações sem significado, porque os nós selecionados não possuíam caudal, logo, não ocorria quase nenhuma deteção a não ser de algumas injeções de contaminante nos próprios nós com sensor, resultando em probabilidades de deteção muito baixas (menos de 1%). Entendeu-se assim que estes casos não deveriam ser considerados.

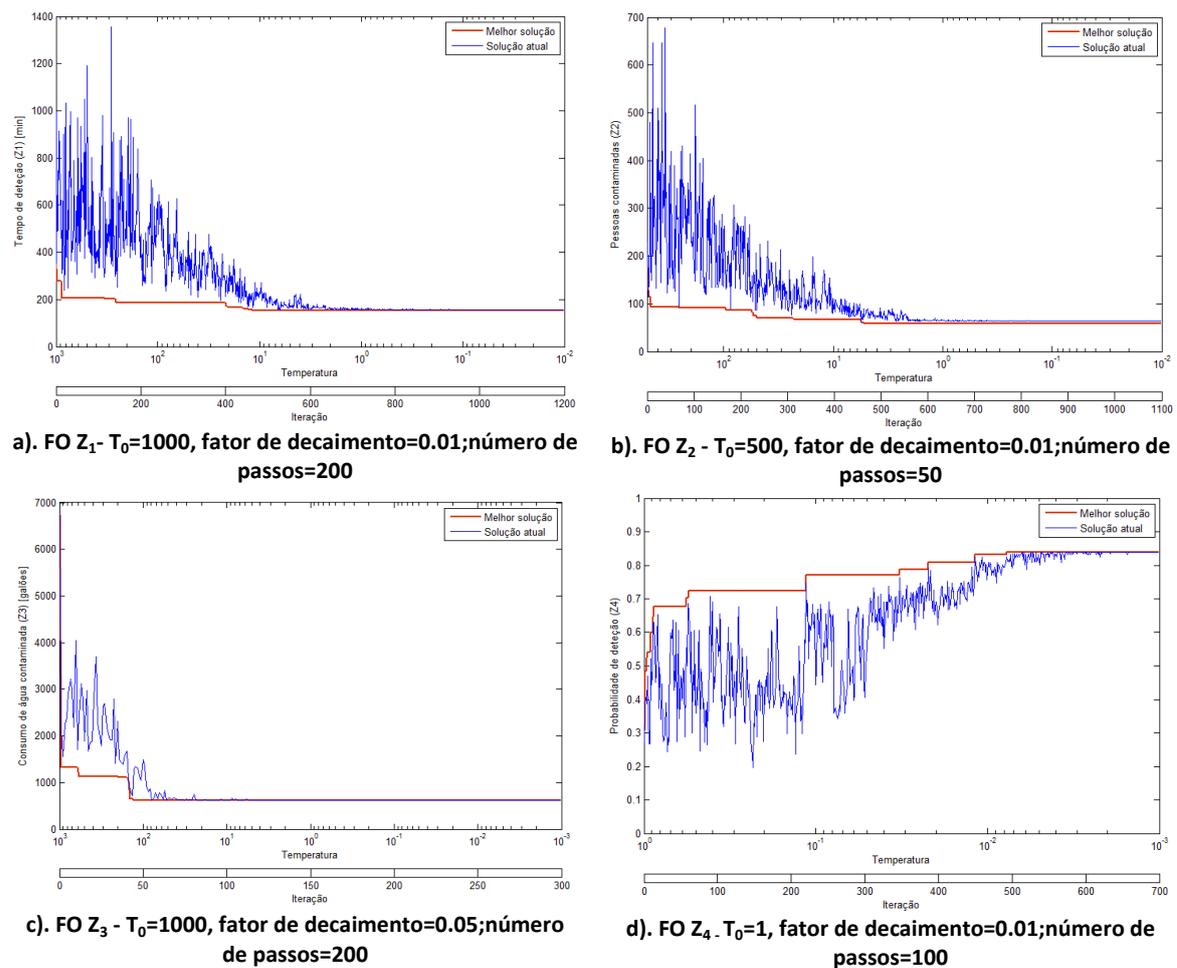


Figura 12. . Evolução da solução inicial em cada iteração e da melhor solução, em função da temperatura e das iterações.

### 4.2.2. Abordagem multiobjectivo

Para a implementação multiobjectivo, começou-se por testar a otimização simultânea de cada par de FO e analisar a extensão das frentes de Pareto obtidas, com o objetivo de testar o conjunto de parâmetros mais adequado. Em seguida procedeu-se a uma otimização triobjectivo, considerando as FO  $Z_1$ ,  $Z_2$  e  $Z_4$ , devido à forte correlação existente entre  $Z_2$  e  $Z_3$ . Finalmente, testou-se a otimização simultânea das quatro FO.

Tal como referido anteriormente, para a otimização multiobjectivo com SA usou-se a implementação SMOSA (Suman e Kumar, 2005), cujo pseudocódigo se encontra no Algoritmo 8, tendo-se optado numa primeira fase pela seguinte função de determinação da probabilidade de aceitação, por modificação da equação (2):

$$P = \min \left( 1, \prod_{i=1}^N \exp \left\{ \frac{-\Delta S_i}{k_i T} \right\} \right) \quad (5)$$

Relativamente à equação (2), o valor de temperatura específico de cada objetivo  $T_i$ , foi substituído pelo produto de  $k_i \times T$ , correspondendo  $k_i$  a valores com a ordem de grandeza das melhores soluções obtidas em cada função:  $k_1=150$ ;  $k_2=60$ ;  $k_3= 600$ ;  $k_4=0.8$ , de modo a que se possa usar um único valor de temperatura independente da FO.

Posteriormente, e para tentar a obtenção de melhores resultados, considerou-se a função de probabilidade “curva logística” tal como definido em (Antunes *et al.*, 2011):

$$P = \frac{2}{1 + \exp\left(\frac{\Delta}{T}\right)} \quad (3)$$

sendo  $\Delta = \sum_{j=1}^p w_j \Delta S_j$ ,  $w_j$  a ponderação atribuída ao objetivo  $Z_j$  e  $T$  a temperatura. A título indicativo atribuiu-se aos pesos  $w_j$  o inverso dos fatores de escala de  $k_j$  anteriormente definidos.

#### 4.2.2.1. Abordagem Biobjectivo

A principal particularidade da implementação da abordagem biobjectivo é a implementação da relação de dominância.

Foram testados os parâmetros listados no Quadro 7, considerando como critério de paragem  $T < 10^{-2}$ , com exceção da otimização para os pares  $Z_i$  ( $i=1,2,3$ ) com  $Z_4$ , em que o critério de paragem utilizado foi  $T < 10^{-3}$ , uma vez que para a otimização de  $Z_4$  se revelou necessário um critério de paragem mais baixo.

**Quadro 7. Conjunto de parâmetros utilizados nas simulações biobjectivo.**

$T_0$	Fator de decaimento	Número de passos
10	0.01	200
100		
1000		

Os resultados obtidos com os testes efetuados permitiram identificar os conjuntos de parâmetros e a estrutura de vizinhança que melhor garantiam a abrangência do espaço de soluções, permitindo atingir os ótimos individuais de cada uma das funções em causa, ou muito aproximados, estando ilustrados na Figura 13.

No processo de otimização das funções  $Z_1-Z_2$  e  $Z_1-Z_3$ , foram obtidos os ótimos individuais de cada uma das FO para os conjuntos de parâmetros com  $T_0=1000$  e  $T_0=100$ . Na otimização de  $Z_2-Z_3$ , também foi possível obter os ótimos individuais das duas FO, mas, neste caso, em todos os conjuntos de parâmetros listados no Quadro 7. Assim, optou-se por realizar as simulações com o conjunto de parâmetros: fator de decaimento=0.01 e 200 passos,  $T_0=100$  para  $Z_1-Z_2$  e  $Z_1-Z_3$ , e  $T_0=10$  para  $Z_2-Z_3$ , por terem um tempo médio de execução inferior. A estrutura de vizinhança considerada correspondeu à troca de um único sensor.

Nos casos relativos à otimização dos pares  $Z_i$  ( $i=1,2,3$ ) com  $Z_4$ , não foi possível garantir a obtenção das soluções ótimas individuais de cada uma das FO. Na otimização do par  $Z_2-Z_4$ , com os vários conjuntos de parâmetros do Quadro 7 apenas foi possível atingir um dos extremos das FO, correspondente ao ótimo individual de  $Z_4$ , para  $T_0=100$ . No entanto, foi para  $T_0=10$  que se obteve o valor mais próximo do ótimo da função  $Z_2$ , com o valor de 63 pessoas contaminadas. A fim de permitir comparar melhor os resultados obtidos, procedeu-se à representação simultânea das duas frentes de Pareto como se pode verificar na Figura 13 c), e) e f).

A análise dos resultados permitiu identificar os conjuntos de parâmetros que permitem uma maior abrangência do espaço de soluções não dominadas, para os diferentes pares de FO, que se encontram resumidos no Quadro 8.

**Quadro 8. Conjuntos de parâmetros de SA que proporcionam maior cobertura do espaço de soluções para a otimização biobjetivo das FO.**

Par de funções a otimizar	Conjunto de Parâmetros				
	T0	$\alpha$	n_passos	Critério paragem	T execução [s]
$Z_1-Z_2$	100	0.01	200	$10^{-2}$	1333
$Z_1-Z_3$				$10^{-3}$	
$Z_1-Z_4$				$10^{-2}$	1871
$Z_2-Z_3$	10			$10^{-2}$	991
$Z_2-Z_4$				$10^{-3}$	1456
$Z_3-Z_4$					1798

Com base nos conjuntos de parâmetros identificados, foram então executadas 11 corridas do processo de otimização para cada par de FO tendo como objetivo perceber em que

medida a agregação das frentes de Pareto obtidas nas diferentes corridas permitiriam melhorar o resultado anteriormente obtido para uma única corrida.

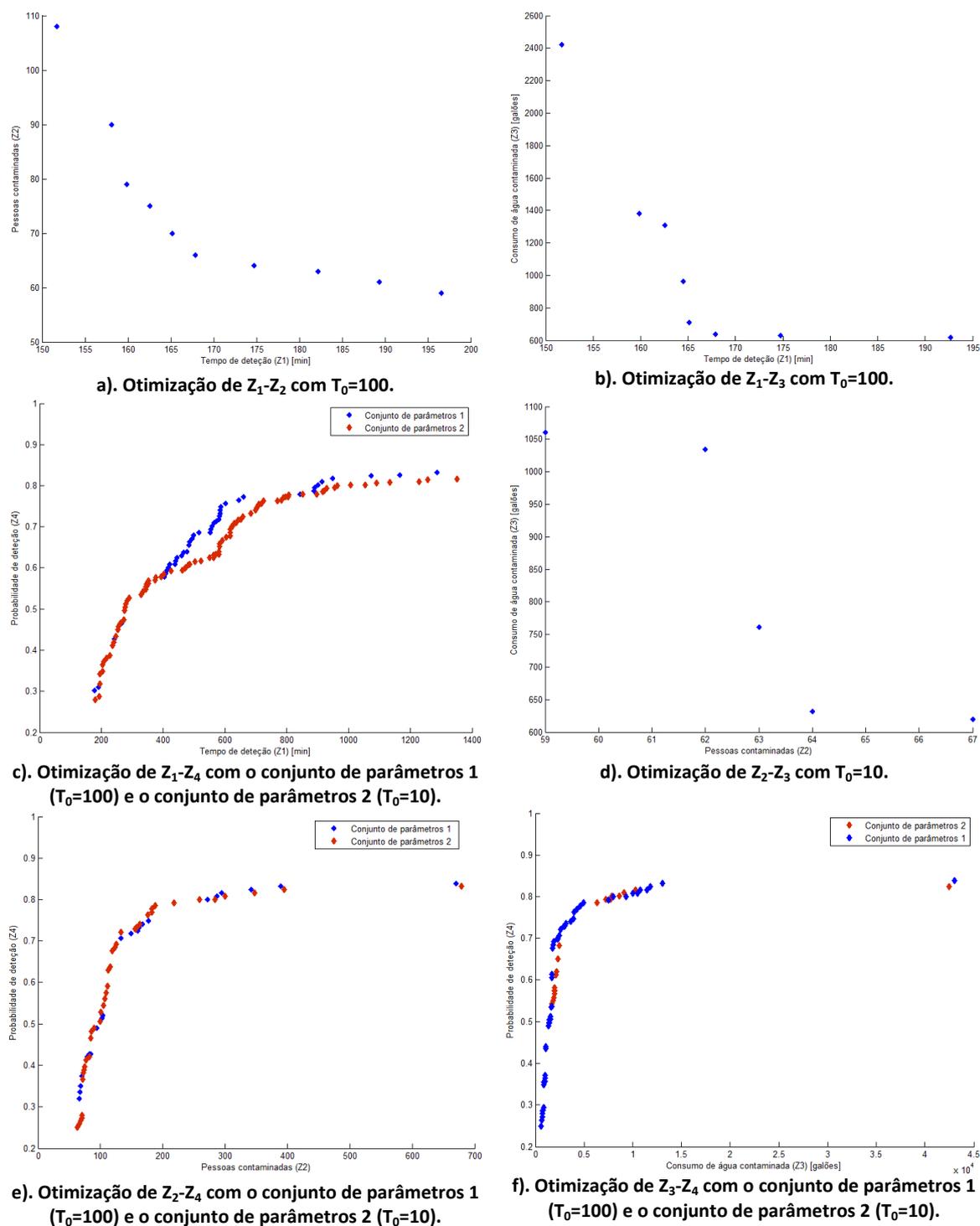


Figura 13. Representação das soluções não dominadas resultantes da otimização biobjetivo.

A Figura 14 mostra que apenas existiram diferenças com algum significado nos pares  $Z_1$ - $Z_4$ ,  $Z_2$ - $Z_4$  e  $Z_3$ - $Z_4$ , onde se notou uma ligeira melhoria na frente de Pareto no sentido de se aproximar do ótimo de  $Z_4$ . No caso de  $Z_1$ - $Z_3$  observou-se uma pequena melhoria com a introdução de uma solução não dominada na frente de Pareto. Para  $Z_1$ - $Z_2$  e  $Z_2$ - $Z_3$  não se verificou qualquer alteração na frente de Pareto, o que leva a crer que serão frentes ótimas.

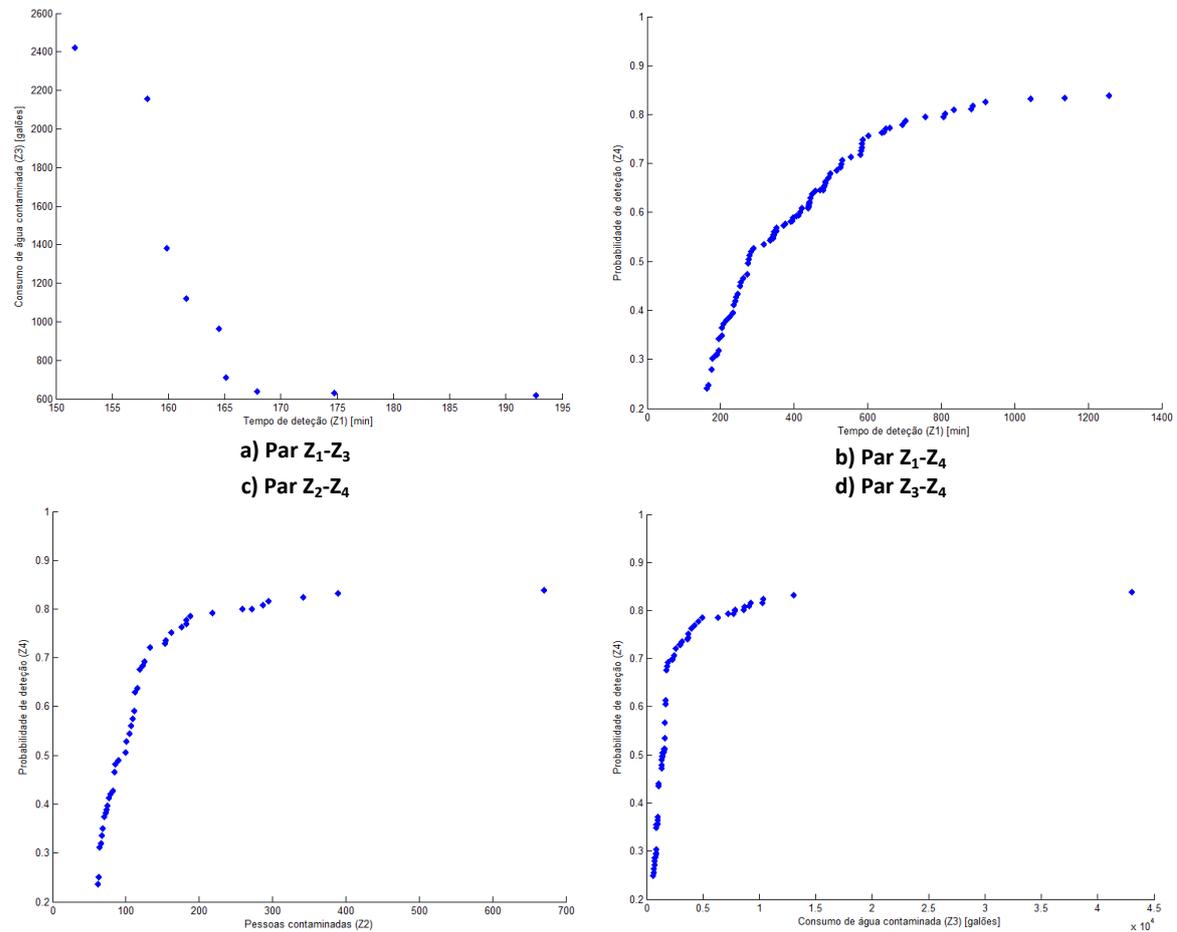


Figura 14. Frente de Pareto na otimização biobjetivo após consolidação com 11 corridas.

Como forma de poder comparar os resultados obtidos nas diferentes abordagens multiobjetivo, optou-se por determinar a distância entre as soluções não dominadas e a solução ideal, que em função do par de objetivos considerado será formada a partir de  $Z_1^*=0$ ,  $Z_2^*=0$ ,  $Z_3^*=0$  e  $Z_4^*=1$ . Para este fim, os valores das soluções não dominadas obtidos foram normalizados, através da divisão pelo máximo valor em cada objetivo ( $Z_1$ ,  $Z_2$ ,  $Z_3$ ). Para que a máxima distância em  $Z_4$  relativamente à situação ideal seja também unitária, normalizou-se o valor complementar de  $Z_4$ . Os resultados obtidos utilizando como métrica a distância Euclideana, encontram-se listados no Quadro 9.

**Quadro 9. Soluções com menor distância Euclideana à solução ideal.**

FO otimizadas	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclideana	Solução				
Z <sub>1</sub> -Z <sub>2</sub>	<b>167.88</b>	<b>66</b>	639.662	24.11%	1.050	22	25	32	35	38
Z <sub>1</sub> -Z <sub>3</sub>	<b>165.16</b>	70	<b>711.596</b>	24.11%	0.906	22	32	35	38	50
Z <sub>1</sub> -Z <sub>4</sub>	<b>531.09</b>	287	4629.819	<b>70.71%</b>	0.572	69	76	84	102	117
Z <sub>2</sub> -Z <sub>3</sub>	174.75	<b>64</b>	<b>631.976</b>	24.11%	1.126	22	27	32	35	38
Z <sub>2</sub> -Z <sub>4</sub>	1074.70	<b>188</b>	7383.323	<b>78.49%</b>	0.397	11	18	32	46	84
Z <sub>3</sub> -Z <sub>4</sub>	886.32	192	<b>3135.934</b>	<b>73.64%</b>	0.358	11	22	69	84	117

Da leitura do Quadro 9 pode-se afirmar que as soluções encontradas na otimização de Z<sub>1</sub>-Z<sub>2</sub>, Z<sub>2</sub>-Z<sub>3</sub> e Z<sub>1</sub>-Z<sub>3</sub> são relativamente próximas e coincidem em 4 sensores, que são o 22, o 32, o 35 e o 38. Já as soluções encontradas na otimização de Z<sub>1</sub>-Z<sub>4</sub>, Z<sub>2</sub>-Z<sub>4</sub> e Z<sub>3</sub>-Z<sub>4</sub>, não sendo tão semelhantes, partilham alguns sensores, nomeadamente o 11, o 69, o 84 e o 117. Na comparação com os resultados das soluções listadas em (Ostfeld *et al.*, 2008), verifica-se a existência de soluções muito próximas das identificadas no Quadro 9 relativas à otimização de Z<sub>1</sub>-Z<sub>4</sub>, Z<sub>2</sub>-Z<sub>4</sub> e Z<sub>3</sub>-Z<sub>4</sub>.

Utilizando como métrica a distância de Chebyshev, obtiveram-se os resultados listados no Quadro 10. Neste conjunto, verifica-se uma coincidência exata na solução obtida para a otimização de Z<sub>1</sub>-Z<sub>2</sub> e Z<sub>1</sub>-Z<sub>3</sub>, que se mantém próxima da solução obtida para Z<sub>2</sub>-Z<sub>3</sub>. Partilham ainda 2 sensores com os restantes pares de funções. A solução Z<sub>2</sub>-Z<sub>4</sub> coincide com a solução obtida utilizando a métrica Euclideana e a solução Z<sub>3</sub>-Z<sub>4</sub> é um pouco mais próxima do ótimo individual de Z<sub>4</sub>, mas Z<sub>3</sub> tem uma degradação significativa.

**Quadro 10. Soluções com menor distância de Chebyshev à solução ideal na otimização biobjetivo.**

FO otimizadas	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Chebyshev	Solução				
Z <sub>1</sub> -Z <sub>2</sub>	<b>159.85</b>	<b>79</b>	1382.349	24.09%	0.813	22	32	36	41	50
Z <sub>1</sub> -Z <sub>3</sub>	<b>159.85</b>	79	<b>1382.349</b>	24.09%	0.830	22	32	36	41	50
Z <sub>1</sub> -Z <sub>4</sub>	<b>516.21</b>	316	5160.346	<b>68.67%</b>	0.413	69	83	102	117	122
Z <sub>2</sub> -Z <sub>3</sub>	182.22	<b>63</b>	<b>760.757</b>	23.47%	0.940	22	27	32	38	109
Z <sub>2</sub> -Z <sub>4</sub>	1074.70	<b>188</b>	7383.323	<b>78.49%</b>	0.281	11	18	32	46	84
Z <sub>3</sub> -Z <sub>4</sub>	1072.80	395	<b>10323.208</b>	<b>82.37%</b>	0.240	11	46	84	117	122

#### 4.2.2.2. Abordagem Triobjetivo

Embora o problema em estudo tenha quatro objetivos, a dificuldade em representar uma frente de Pareto com quatro dimensões, associada à identificação já efetuada da correlação existente entre Z<sub>2</sub> e Z<sub>3</sub>, levou a que se experimentasse a otimização simultânea de apenas Z<sub>1</sub>-Z<sub>2</sub>-Z<sub>4</sub>.

Tal como na abordagem biobjetivo, a diferença mais significativa está na implementação da relação de dominância, que aqui tem que considerar as três FO escolhidas.

Foram de novo testados os parâmetros listados no Quadro 7, considerando como critério de paragem  $T < 10^{-3}$ . O Quadro 11 ilustra os melhores valores obtidos para cada uma das FO no conjunto das soluções não dominadas resultantes de cada otimização efetuada com o conjunto de parâmetros indicado. Assim, a maior aproximação aos ótimos individuais de cada FO, o que corresponde a uma maior cobertura da frente não dominada, foi obtida para uma temperatura de 100 com a alteração de um sensor, e uma temperatura de 1000 com a alteração de dois sensores.

**Quadro 11. Resumo dos melhores resultados obtidos na otimização triobjetivo em função dos parâmetros.**

$T_0$	$\alpha$	n_passos	Estrutura vizinhança	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$
1000	0.01	200	1 sensor	176.523	68.000	1060.339	0.793
100	0.01	200		162.300	67.000	1127.807	0.801
10	0.01	200		189.808	68.000	1148.664	0.771
1000	0.01	200	2 sensores	169.675	63.000	619.114	0.808
100	0.01	200		179.606	65.000	666.480	0.794
10	0.01	200		173.556	65.000	1045.288	0.809

Uma vez que os resultados foram obtidos com uma única corrida, com vista a melhor decidir relativamente aos parâmetros a usar, foram executadas 11 corridas de cada processo. Os resultados mostram que a temperatura de 100 com a estrutura de vizinhança correspondente à alteração de um sensor, é o conjunto de parâmetros que melhor se adequa, confirmando os resultados já obtidos na otimização biobjetivo, tendo tido um tempo de execução de 2 460 segundos.

Tendo em conta as onze corridas da otimização triobjetivo com os parâmetros escolhidos, procedeu-se à reunião das frentes não dominadas, obtendo uma frente de Pareto consolidada que se encontra representada na Figura 15.

Esta representação tridimensional pode ser útil para ajudar um decisor a perceber que compromissos podem ser estabelecidos e quais os seus efeitos, ajudando a reduzir a dimensão do número de soluções relevantes em análise e a identificar uma solução de compromisso satisfatória.

Na sequência da análise já efetuada na otimização biobjetivo, foi calculada a distância Euclideana e a distância de Chebyshev à solução ideal,  $Z_1^*=0$ ,  $Z_2^*=0$  e  $Z_4^*=1$ , como forma de comparar as diferentes abordagens ainda sem estabelecer preferências entre FO. No Quadro 12 são apresentadas as soluções para as quais foram calculadas estas distâncias mínimas.

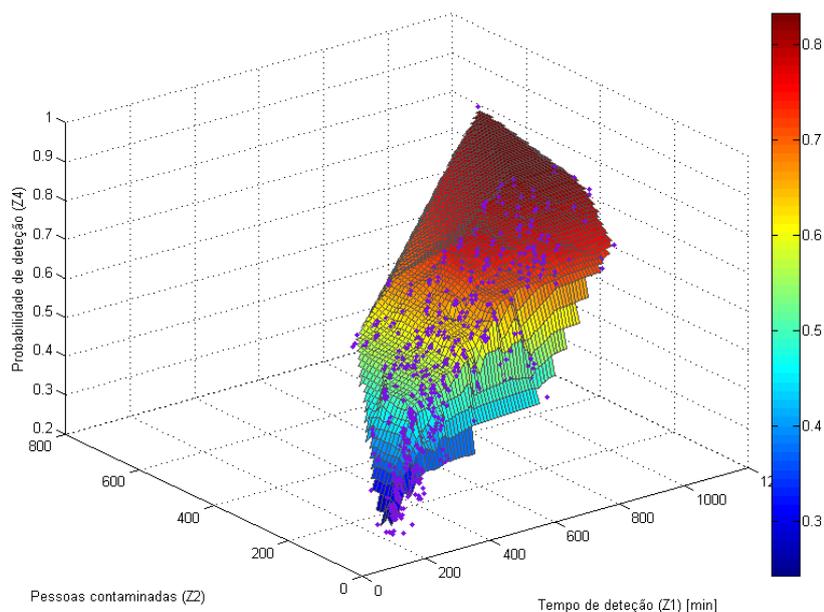


Figura 15. Frente de Pareto para a otimização triobjetivo ( $Z_1, Z_2, Z_4$ ).

Quadro 12. Soluções com menor distância à solução ideal na otimização triobjetivo.

FO otimizadas	$Z_1$	$Z_2$	$Z_3$	$Z_4$	Distância Euclidiana	Distância Chebyshev	Solução				
$Z_1-Z_2-Z_4$	554.48	125	1974.817	69.17%	0.659		18	22	69	84	102
	492.88	220	3971.535	66.34%		0.440	18	69	83	102	121

A Figura 16 ilustra a localização das soluções identificadas na frente de Pareto, podendo verificar-se que se encontram perto do patamar com as melhores soluções em  $Z_4$ . Constata-se também uma proximidade destas soluções a duas das listadas em (Ostfeld *et al.*, 2008).

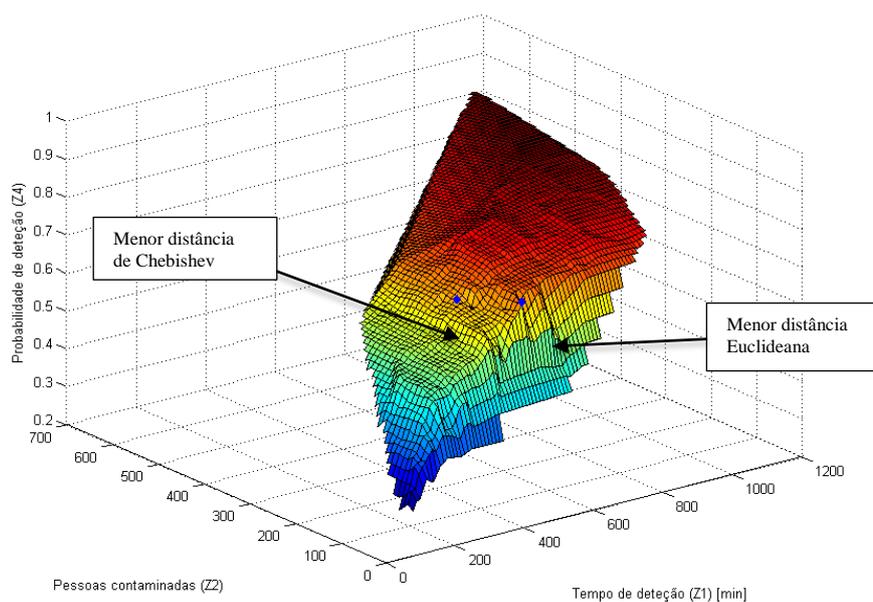


Figura 16. Localização na frente de Pareto das soluções com menor distância à solução ideal.

#### 4.2.2.3. Abordagem Tetra-objetivo

A inclusão da quarta função objetivo cumpre de forma mais exata os objetivos enunciados do problema em estudo. No entanto, dada a forte correlação verificada entre as FO  $Z_2$  e  $Z_3$ , é expectável que os resultados sejam semelhantes

Foram testados de novo os conjuntos de parâmetros listados no Quadro 7, considerando como critério de paragem  $T < 10^{-3}$ . No entanto, tendo-se verificado que o número de soluções não dominadas aumentava continuamente em função do número de iterações, decidiu-se testar também o prolongamento do processo, através da diminuição da temperatura limite associada ao critério de paragem, para verificar se o número de soluções não dominadas atingiria um ponto de estabilização. Esta situação não chegou a ocorrer, como se pode verificar na Figura 17, pelos resultados do Quadro 13, mas verifica-se que esse prolongamento não permitiu obter resultados mais abrangentes, embora implique um tempo de processamento bastante superior.

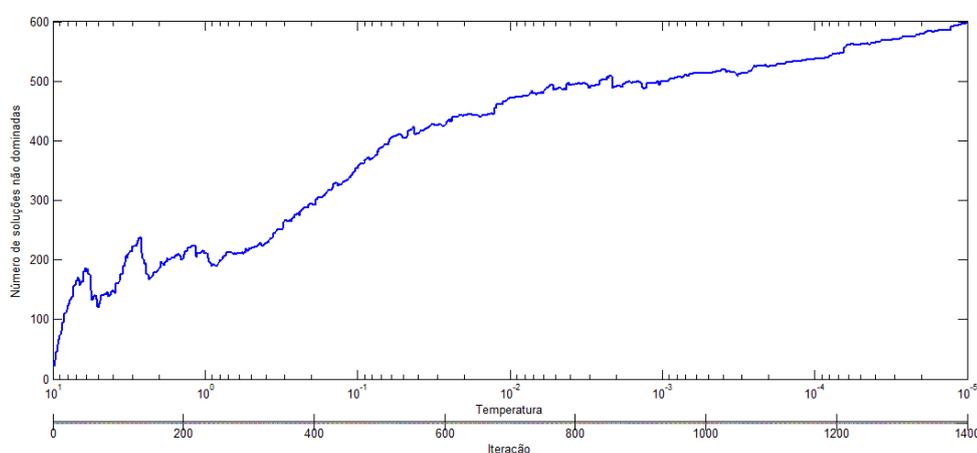


Figura 17. Evolução do número de soluções não dominadas em função da temperatura.

Quadro 13. Resumo da caracterização dos resultados obtidos na otimização tetra-objetivo em função dos parâmetros

$T_0$	$\alpha$	n_passos	Estrutura vizinhança	Critério de paragem	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$
1000	0.01	200	1 sensor	$10^{-3}$	159.854	64	693.189	0.794
100	0.01	200			163.967	61	619.114	0.839
10	0.01	200			165.155	59	619.114	0.802
10	0.01	200		$10^{-4}$	175.301	66	678.610	0.818
10	0.01	200		$10^{-5}$	151.708	61	619.114	0.833
1	0.01	200		$10^{-3}$	171.588	65	655.275	0.818
1000	0.01	200	2 sensores	$10^{-3}$	177.656	66	691.147	0.787
100	0.01	200			178.983	63	678.734	0.818
10	0.01	200			173.590	66	907.580	0.824

Do Quadro 13 resulta que o conjunto de parâmetros em que  $T_0 = 100$  é aquele que permite obter valores mais próximos dos ótimos individuais de cada função, num tempo aceitável



De notar que as duas soluções aqui identificadas são também muito próximas das soluções obtidas a partir das duas métricas utilizadas na otimização biobjetivo  $Z_1$ - $Z_4$ , sendo bastante diferentes das soluções identificadas para qualquer outro par de FO.

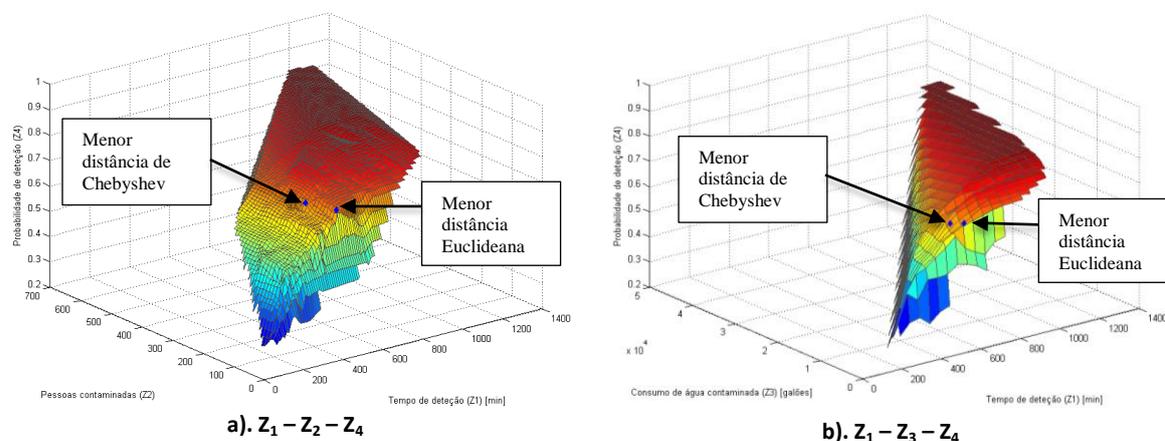


Figura 19. Localização das soluções de menor distância à solução ideal na frente de Pareto com 4 objetivos.

### 4.2.3. Análise de resultados com SA

A decisão sobre qual das soluções não dominadas corresponde à solução a escolher, implica necessariamente compromissos entre as FO que só um decisor final poderá estabelecer. Sem a interação com um decisor real, pode apenas especular-se sobre compromissos aceitáveis ou sugerir estratégias de apoio à sua definição.

Nas secções anteriores, para efeitos de comparação utilizou-se um processo de determinação de soluções com base numa métrica de distância à solução ideal. Porém, essa métrica não reflete qualquer preferência que o decisor possa ter em termos de importância relativa dos objetivos.

Não podendo contar com a interação com um decisor final implicado na gestão de uma RAA real, pode-se no entanto discutir a possível importância dos critérios e de que forma a redução do número de soluções a considerar pode facilitar o processo de decisão. Nesse sentido, de entre as quatro funções objetivo apontados pela BWSN aquele que talvez mereça maior atenção é a probabilidade de deteção ( $Z_4$ ) uma vez que um baixo valor nesta função objetivo significa que um número não desprezável de eventos de contaminação não foi detetado durante o período de simulação, sendo os seus efeitos desconhecidos uma vez que nenhuma das outras funções objetivo os captura. De notar que nem todas as soluções apresentadas à BWSN parecem contemplar esta preocupação, uma vez que existem soluções com uma probabilidade de deteção inferior a 40%. No entanto, constatou-se na otimização biobjetivo a proximidade das soluções que minimizavam a distância à solução ideal para os pares  $Z_i$ - $Z_4$  ( $i=1,2,3$ ) a várias soluções listadas em (Ostfeld *et al.*, 2008), o que dá a entender que este objetivo foi uma preocupação importante para vários autores.

A título de exemplo e tendo em atenção os resultados obtidos, considerou-se que um cenário possível poderia passar por aceitar apenas uma probabilidade de deteção superior a 75%. Considerando este pressuposto, a frente de Pareto resultante da otimização com 4 funções objetivo, seria reduzida à frente representada na Figura 20, representando apenas os objetivos  $Z_1$ ,  $Z_2$  e  $Z_4$ . Neste contexto, é mais fácil interagir com um decisor de forma a identificar na representação tridimensional as soluções que parecem mais atrativas em termos de satisfação dos diferentes objetivos. Assumindo estar na presença de soluções aceitáveis do ponto de vista da probabilidade de deteção, é visível na Figura 20 uma face quase vertical, próxima do vértice correspondente aos melhores valores em  $Z_1$  e  $Z_2$ , na qual foram localizadas três soluções com potencial interesse, por corresponderem possivelmente aos melhores valores de compromisso nestes objetivos, tendo-se optado pela solução intermédia  $Z_1=823$  min,  $Z_2=261$  pessoas,  $Z_3=7927.978$  galões e  $Z_4=78.67\%$ , para o conjunto de sensores indicado no Quadro 15 em termos de índice numérico e em termos de designações dos nós, tal como constam no modelo de rede e como são identificados em Ostfeld *et al.* (2008).

Quadro 15. Solução proposta R1A5, com SA.

Posição	32	46	84	101	117
Etiqueta dos nós	31	45	83	100	118

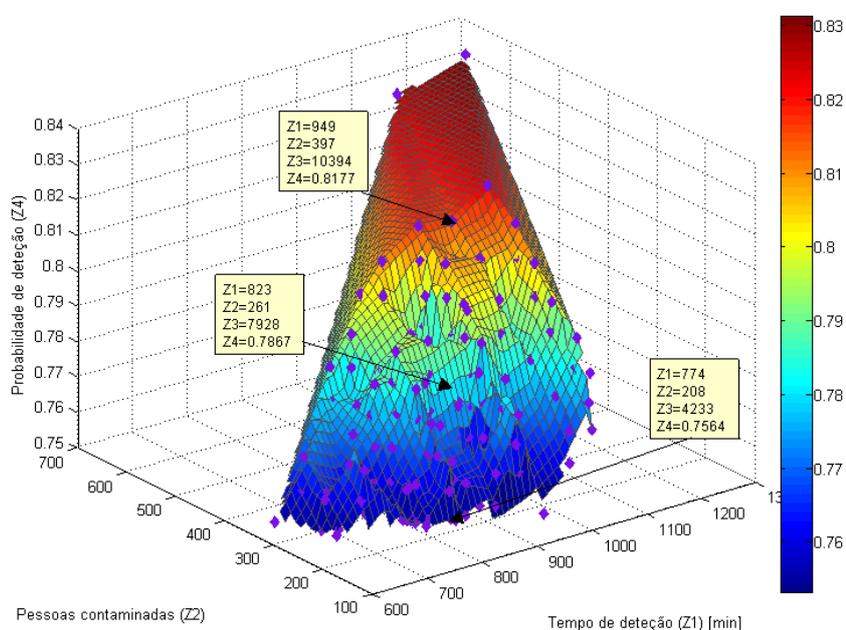
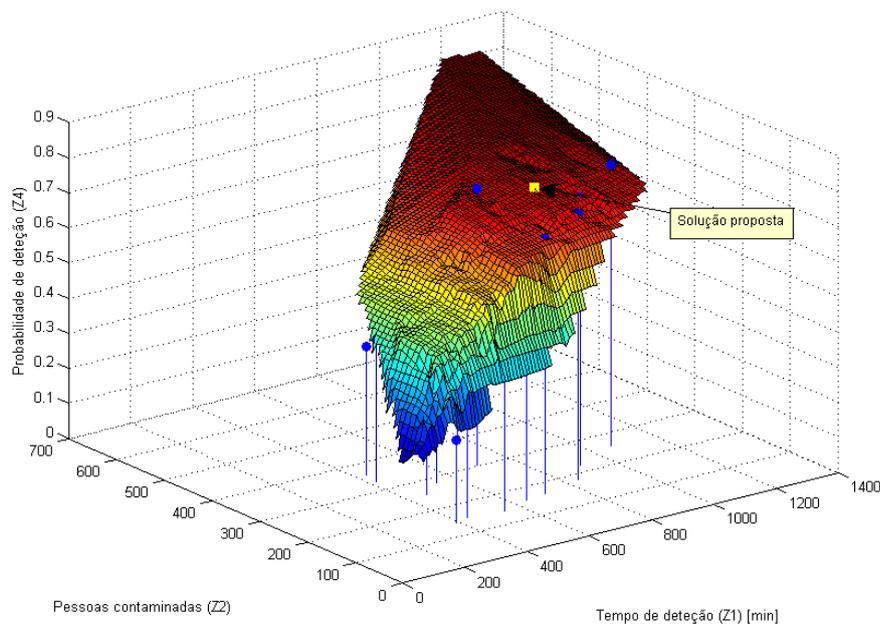


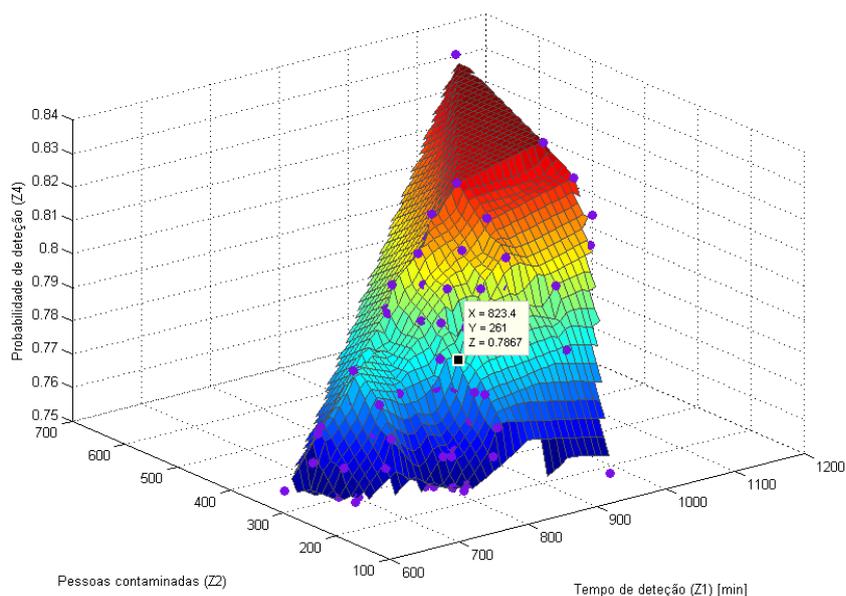
Figura 20. Frente de Pareto com otimização tetra-objetivo, reduzida a soluções com probabilidade de deteção superior a 75%.

Representando na Figura 21 a solução proposta e as soluções apresentadas em Ostfeld *et al.* (2008), resultantes da BWSN, na frente de Pareto, constata-se que a solução agora proposta se insere no centro das soluções da BWSN que se enquadram na frente não dominada. Verifica-se também que seis das catorze soluções apresentadas na BWSN correspondem a soluções dominadas por esta frente. Este facto foi confirmado numericamente uma vez que a quarta dimensão não permitia fazer essa afirmação inequívoca na representação tridimensional.



**Figura 21. Representação da solução proposta na frente de Pareto e comparação com soluções apresentadas na BWSN.**

A repetição do processo atrás descrito com a frente de Pareto resultante da otimização triobjetivo, permite localizar exatamente a mesma solução proposta com base na otimização tetra-objetivo, como se pode verificar na Figura 22.



**Figura 22. Localização da solução proposta na frente de Pareto da otimização triobjetivo.**

Tendo em atenção a constatação efetuada de que a otimização biobjetivo  $Z_1$ - $Z_4$  permitiu resultados muito semelhantes quando se procurou identificar a solução com a distância mínima à solução ideal, foi analisada a frente de Pareto resultante, com os mesmos pressupostos utilizados na procura da solução que se acabou de descrever. Assim, selecionaram-se apenas as soluções não dominadas correspondentes a uma probabilidade de deteção superior a 75%, tendo sido representada a frente de Pareto na Figura 23. Tendo-se observado que a localização das soluções se agrupavam em tornos de três segmentos de reta com declives diferentes e considerando que, estando  $Z_4$  numa zona de aceitabilidade, um eventual decisor tenderia a não deixar já degradar muito o objetivo  $Z_1$ , pelo que uma opção possível seria a indicada, na interseção de duas das retas. A solução encontrada,  $Z_1=704$  min,  $Z_2=303$  pessoas,  $Z_3=8406$  galões e  $Z_4=0,787$ , para os sensores 46, 69, 84, 101, 117, não correspondendo à solução proposta na otimização com quatro objetivos é, porém, muito próxima e corresponde a uma das soluções apresentadas na BWSN (Wu e Walski, 2006).

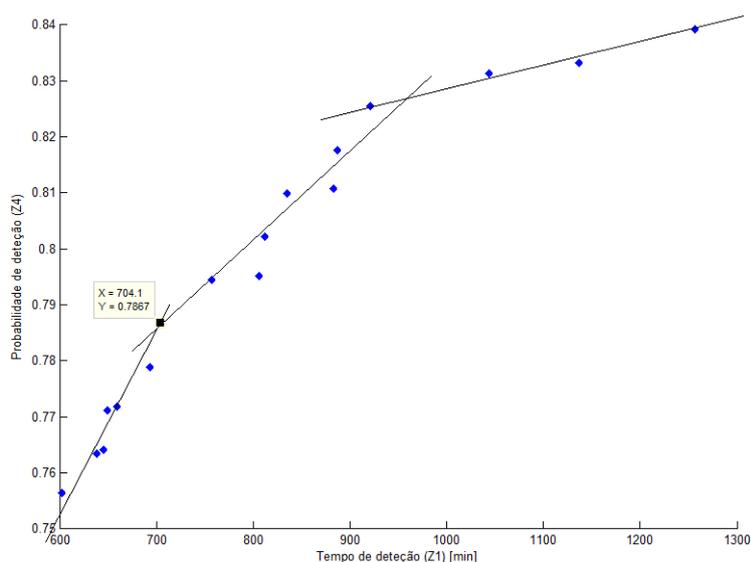


Figura 23. Proposta de solução na frente de Pareto biobjetivo  $Z_1$ - $Z_4$ .

Tendo em atenção o tempo de execução, constata-se que a otimização triobjetivo pode corresponder a uma opção a considerar uma vez que a sua execução leva menos tempo que a otimização tetra-objetivo. Já a otimização biobjetivo, embora seja mais rápida, pode obrigar a procedimentos de análise específicos que levem à anulação dessa vantagem.

### 4.3. Aplicação dos algoritmos genéticos para a determinação da rede de sensores

#### 4.3.1. Otimização mono-objetivo com a implementação base, incluindo elitismo

Tal como no *simulated annealing* e outras metaheurísticas, não existem regras estabelecidas para determinar os diversos parâmetros necessários ao funcionamento do algoritmo genético. Pereira (2012) aponta como valores típicos para o tamanho da população o uso de 100 a 500 elementos, para a probabilidade de recombinação, valores entre 0.7 e 1.0 e para a probabilidade de mutação valores entre 0.001 e 0.1. Deb *et al.* (2002) sugere para este último parâmetro a referência ao número de variáveis de decisão, definindo a probabilidade de mutação  $p_m=1/n$ , sendo  $n$  o número de variáveis, ou a dimensão do cromossoma.

Assim, para o problema da colocação de 5 sensores na Rede 1, e tendo em atenção os parâmetros de maior sucesso na otimização com SA, decidiu-se experimentar o conjunto de parâmetros listado no Quadro 16, para todas as FO, considerando um máximo de 150 gerações e uma percentagem de elitismo de 5%.

Quadro 16. Resumo dos parâmetros testados para o Algoritmo Genético.

População	Probabilidade Mutação	Probabilidade Recombinação
50	0.1	0.7
100	0.15	0.8
200	0.2	

A execução de 11 corridas do algoritmo com cada um dos conjuntos de parâmetros indicados, permitiu confirmar os resultados da otimização mono-objetivo obtidos com o SA, como se pode analisar no Quadro 17. Porém, estes resultados não são facilmente obtidos, sendo as taxas de sucesso relativamente inferiores às que se conseguiram obter com SA, como se pode verificar no Quadro 18.

Quadro 17. Soluções que otimizam cada função objetivo.

Função otimizada	Valor da solução				Sensores				
	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [galões]	Z <sub>4</sub> [%]					
Z <sub>1</sub>	<b>151.71</b>	108	2422.85	20.33%	32	35	38	41	50
Z <sub>2</sub>	196.57	<b>59</b>	1060.83	22.82%	22	27	31	38	55
Z <sub>3</sub>	192.65	67	<b>619.11</b>	24.86%	21	25	32	35	38
Z <sub>4</sub>	1256.88	670	43041.84	<b>83.92%</b>	11	46	84	101	125

Os conjuntos de parâmetros que deram melhores resultados usaram uma probabilidade de mutação mais elevada (0.20). O tamanho da população que produziu os

melhores resultados não foi o mesmo para todas as FO. A otimização da função Z1 e Z4 resultou melhor com um tamanho de população mais reduzido, principalmente em Z1. As estatísticas relativas às soluções obtidas para a otimização de cada FO encontram-se no Apêndice C.

**Quadro 18. Resumo dos parâmetros que permitiram obter a melhor solução.**

Função	Conjunto de Parâmetros					
	População	Prob. Mutação	Prob. Recombinação	Máx. Gerações	Taxa de sucesso	T execução [s]
Z <sub>1</sub>	50	0.20	0.7	150	18%	45
Z <sub>2</sub>	200	0.20	0.7	150	27%	178
Z <sub>3</sub>	200	0.20	0.7	150	45%	177
Z <sub>4</sub> <sup>(1)</sup>	100	0.20	0.7	150	9%	94

<sup>(1)</sup> Só houve um único conjunto de parâmetros a atingir o ótimo.

A convergência muito rápida verifica-se em todas as FO, tal como se ilustra na Figura 24, já que Z<sub>1</sub>, Z<sub>3</sub> e Z<sub>4</sub> convergem ainda antes da geração 50 e Z<sub>2</sub> estabiliza a partir da geração 102, pode justificar os melhores resultados serem obtidos com uma probabilidade de mutação mais elevada. Este facto proporciona potencialmente maior facilidade em escapar a ótimos locais, por garantir maior diversidade genética da população. Os resultados apresentados na Figura 24 foram obtidos considerando uma probabilidade de mutação de 0.20 e uma probabilidade de recombinação de 0.7.

Nota-se, em particular nos casos de Z<sub>1</sub>, Z<sub>2</sub> e um pouco em Z<sub>4</sub>, que o valor médio do desempenho dos elementos da população converge bastante cedo para valores muito próximos do desempenho do melhor elemento, o que reflete uma perda de diversidade da população. O elitismo pode estar a agravar este problema, tendo-se por isso considerado a introdução de um mecanismo de compensação da perda de diversidade, aumentando a probabilidade de mutação quando a variância da população diminui significativamente.

Na tentativa de melhorar o desempenho do processo de otimização foram ainda experimentadas as seguintes variantes, também de uma forma sistemática, com a aplicação de vários conjuntos de parâmetros:

- Aplicação de uma variante de elitismo em que os melhores progenitores competem com os descendentes pela presença na geração seguinte, tal como explicado em 3.3.2.2, que foi testada com os mesmos conjuntos de parâmetros da implementação (Quadro 16);
- Uso de probabilidades de mutação e de recombinação determinadas dinamicamente, conforme explicado em 3.3.2.3. Esta variante foi testada com os tamanhos de população indicados no Quadro 16, com 3% e 5% de elitismo. Para os parâmetros que definem a probabilidade de recombinação e

da probabilidade de mutação, foram adotados os valores indicados em Vasconcelos *et al.* (2001) para uma variação linear de probabilidades, em que a probabilidade de recombinação varia entre 0.5 e 1.0 e a probabilidade de mutação varia entre 0.025 e 0.25.

Em nenhum destes casos houve melhoria dos resultados obtidos, tendo-se verificado resultados muito próximos com a variante de elitismo, mas muito piores com o uso de probabilidades dinâmicas.

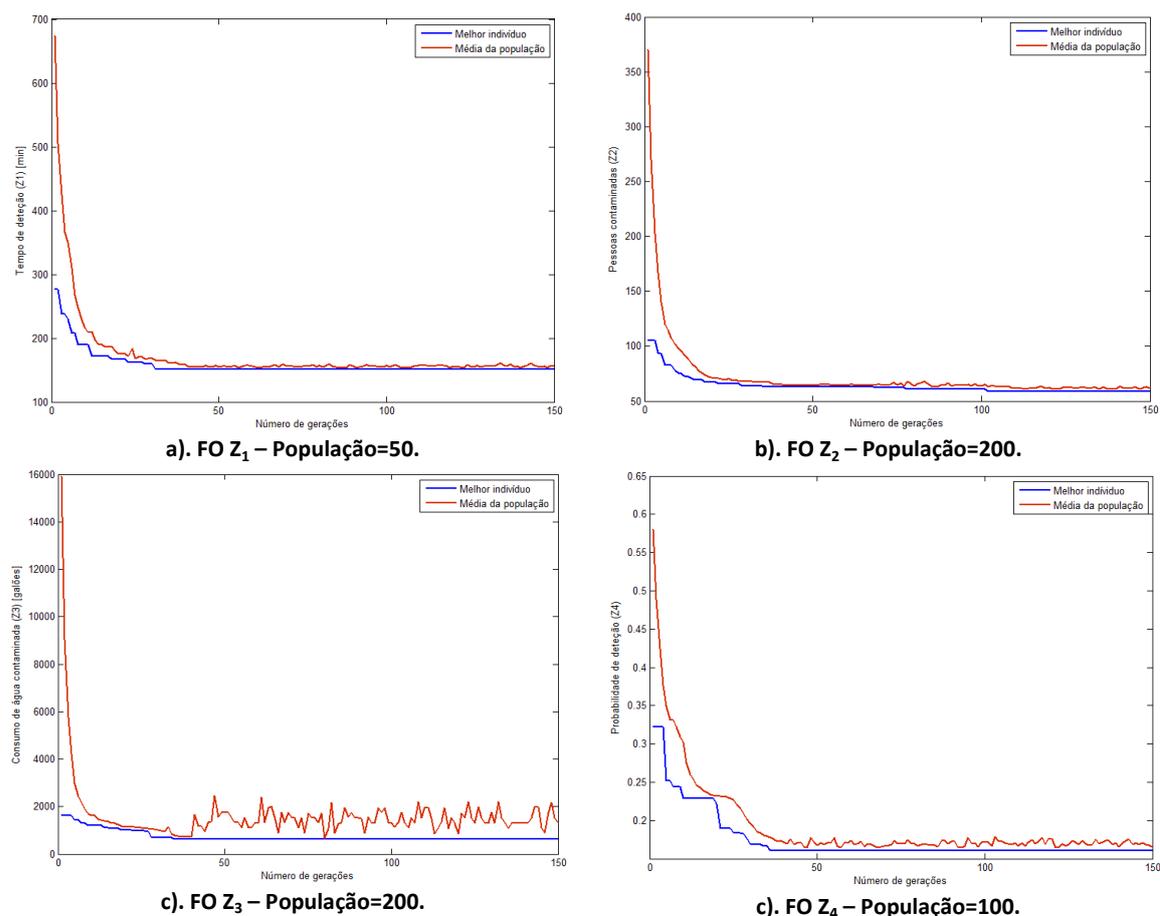


Figura 24. Evolução do melhor elemento e da média da população com o número de gerações.

### 4.3.2. Abordagem multiobjetivo

#### 4.3.2.1. Abordagem biobjetivo

Para a abordagem biobjetivo recorreu-se ao algoritmo NSGA-II de Deb *et al.* (2002), tal como descrito em 3.3.2.4. Este método recorre ao conceito de dominância, agrupando as soluções pertencentes a uma população em frentes de Pareto com diferentes níveis de não dominância, que vão evoluindo com as gerações sucessivas. No final do processo, espera-se que a frente de nível 1 contenha o conjunto de soluções não dominadas.

Com base nos parâmetros que produziram melhores resultados na otimização mono-objetivo com AG, definiram-se os conjuntos de parâmetros listados no Quadro 19, a serem usados na otimização multiobjetivo.

**Quadro 19. Conjunto de parâmetros utilizados nas simulações biobjetivo na otimização com NSGA-II.**

População	Prob. Mutação	Prob. Recombinação	Máx. Gerações
100	0.2	0.8	150
200			

Tendo-se realizado 11 corridas no processo de otimização com cada conjunto de parâmetros, verificou-se que a otimização de  $Z_2$ - $Z_3$  atinge sempre os dois ótimos individuais. Na otimização do par  $Z_1$ - $Z_3$ , verifica-se que em todas as corridas se atinge o ótimo individual de  $Z_3$ , mas apenas se consegue atingir uma vez o ótimo individual de  $Z_1$ , para um tamanho de população de 200 elementos. Relativamente à otimização do par  $Z_1$ - $Z_2$  verifica-se que com qualquer dos conjuntos de parâmetros se obtêm soluções semelhantes, atingindo sempre o ótimo individual de  $Z_2$  e atingindo uma vez o ótimo individual de  $Z_1$ . Na otimização dos pares  $Z_1$ - $Z_4$ ,  $Z_2$ - $Z_4$  e  $Z_3$ - $Z_4$ , apenas se consegue atingir o ótimo de  $Z_4$  e para um tamanho de população de 200 elementos. A Figura 25 mostra as frentes de Pareto para cada par de FO a otimizar. As frentes de Pareto representadas na Figura 25 são praticamente coincidentes com as obtidas em SA, com exceção da otimização do par  $Z_1$ - $Z_3$  em que a frente não inclui a solução com o ótimo individual de  $Z_1$ , o que se verificava na otimização com SA. Os conjuntos de parâmetros que permitiram obter melhores resultados são os constantes no Quadro 20.

**Quadro 20. Conjuntos de parâmetros do NSGA-II que proporcionam maior cobertura do espaço de soluções para otimização biobjetivo das FO.**

Par de funções a otimizar	Conjunto de Parâmetros				
	Tpop	Prob. recombinação	Prob. mutação	Máx. gerações	T execução [s]
$Z_1$ - $Z_2$	100	0.8	0.2	150	272
$Z_1$ - $Z_3$	200				867
$Z_1$ - $Z_4$					272
$Z_2$ - $Z_3$	100				867
$Z_2$ - $Z_4$	200				272
$Z_3$ - $Z_4$					867

Tal como anteriormente, procedeu-se à determinação das distâncias Euclideana e de Chebyshev à solução ideal, ( $Z_1^*=0$ ,  $Z_2^*=0$ ,  $Z_3^*=0$  e  $Z_4^*=1$ ). Como se pode ver no Quadro 21, as soluções com menor distância Euclideana nas frentes relativas a  $Z_1$ - $Z_2$  e  $Z_2$ - $Z_4$  coincidem com as obtidas com SA. As soluções obtidas na otimização dos pares  $Z_1$ - $Z_3$ ,  $Z_1$ - $Z_4$  e  $Z_2$ - $Z_3$ , embora diferentes, são soluções muito próximas das determinadas com SA. Apenas a solução obtida para

o par  $Z_3$ - $Z_4$  é um pouco diferente. As soluções relativas aos pares  $Z_1$ - $Z_2$ ,  $Z_1$ - $Z_3$  são iguais e diferindo apenas num sensor da solução relativa ao par  $Z_2$ - $Z_3$ , muito próximo do obtido em SA, para os mesmos pares de FO.

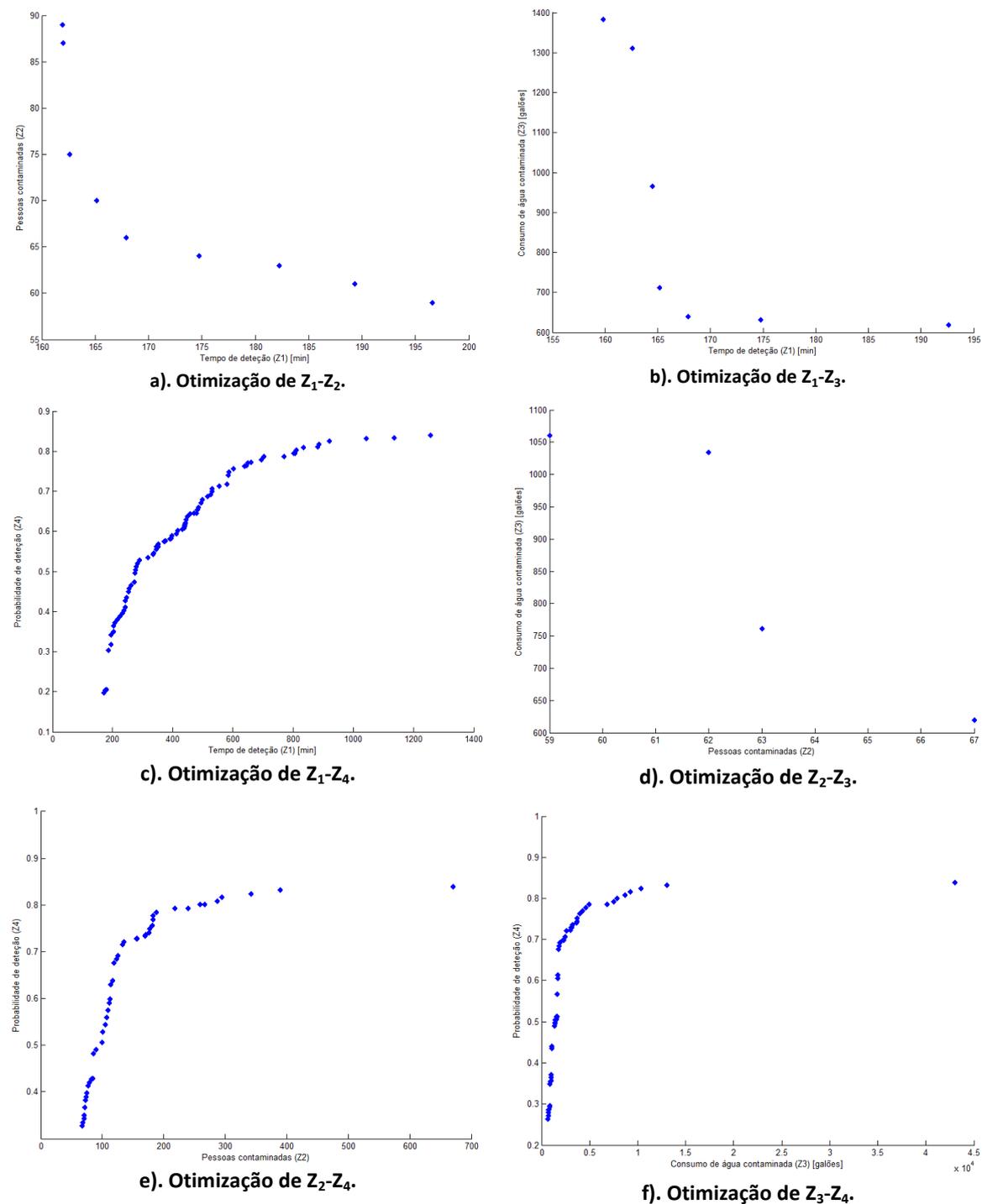


Figura 25. Representação das soluções não dominadas resultantes da otimização biobjetivo com NSGA-II.

**Quadro 21. Soluções com menor distância Euclideana à solução ideal na otimização biobjetivo com NSGA-II.**

FO otimizadas	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclideana	Solução				
Z <sub>1</sub> -Z <sub>2</sub>	<b>167.88</b>	<b>66</b>	639.662	24.11%	1.131	22	25	32	35	38
Z <sub>1</sub> -Z <sub>3</sub>	<b>167.88</b>	66	<b>639.662</b>	24.11%	0.987	22	25	32	35	38
Z <sub>1</sub> -Z <sub>4</sub>	<b>530.70</b>	286	4568.397	<b>69.94%</b>	0.558	69	75	84	99	117
Z <sub>2</sub> -Z <sub>3</sub>	192.65	<b>67</b>	<b>619.114</b>	24.86%	1.158	21	25	32	35	38
Z <sub>2</sub> -Z <sub>4</sub>	1074.70	<b>188</b>	7383.323	<b>78.49%</b>	0.425	11	18	32	46	84
Z <sub>3</sub> -Z <sub>4</sub>	1030.89	260	<b>4966.651</b>	<b>78.57%</b>	0.313	11	32	84	117	122

O Quadro 22 mostra os resultados obtidos com a distância de Chebyshev, sendo visível que as otimizações dos pares Z<sub>2</sub>-Z<sub>3</sub>, Z<sub>2</sub>-Z<sub>4</sub> e Z<sub>3</sub>-Z<sub>4</sub> são exatamente os mesmos dos obtidos com SA, sendo os restantes um pouco diferentes. No entanto, Z<sub>1</sub>-Z<sub>2</sub>, Z<sub>1</sub>-Z<sub>3</sub> e Z<sub>2</sub>-Z<sub>3</sub> apresentam soluções próximas, partilhando as duas primeiras 4 sensores. Verifica-se ainda que a solução obtida para o par Z<sub>2</sub>-Z<sub>4</sub> é a mesma com as duas métricas.

**Quadro 22. Soluções com menor distância de Chebyshev à solução ideal na otimização biobjetivo com NSGA-II.**

FO otimizadas	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Chebyshev	Solução				
Z <sub>1</sub> -Z <sub>2</sub>	<b>165.16</b>	<b>70</b>	711.596	24.11%	0.840	22	32	35	38	50
Z <sub>1</sub> -Z <sub>3</sub>	<b>164.50</b>	76	<b>964.906</b>	24.74%	0.854	22	32	35	36	50
Z <sub>1</sub> -Z <sub>4</sub>	<b>497.96</b>	360	6646.745	<b>67.89%</b>	0.400	71	83	102	117	121
Z <sub>2</sub> -Z <sub>3</sub>	182.22	<b>63</b>	<b>760.757</b>	23.47%	0.940	22	27	32	38	109
Z <sub>2</sub> -Z <sub>4</sub>	1074.70	<b>188</b>	7383.323	<b>78.49%</b>	0.320	11	18	32	46	84
Z <sub>3</sub> -Z <sub>4</sub>	1072.80	395	<b>10323.208</b>	<b>82.37%</b>	0.240	11	46	84	117	122

#### 4.3.2.2. Abordagem triobjetivo

Seguindo o procedimento executado com SA, foi realizada a otimização simultânea das funções objetivo Z<sub>1</sub>, Z<sub>2</sub> e Z<sub>4</sub> com base na identificação de uma forte correlação existente entre Z<sub>2</sub> e Z<sub>3</sub>. Tendo em atenção que na otimização com SA se verificou a existência de 530 soluções não dominadas, considerou-se que seria importante acrescentar aos conjuntos de parâmetros listados no Quadro 19 um conjunto de parâmetros com um tamanho de população de 500 elementos. O Quadro 23 mostra os resultados obtidos, onde se vê que esta última hipótese é a que melhor permite cobrir a frente de Pareto, mas, com uma diferença reduzida face à opção com tamanho da população de 200. Esta mesma diferença reduzida é confirmada pela determinação da mesma solução quando se procura a mínima distância Euclideana à solução ideal. O mesmo já não acontece relativamente à solução que minimiza a distância de Chebyshev, embora se obtenha o mesmo valor de distância. Estes resultados estão apresentados no Quadro 24. Porém, dada a diferença não desprezável no tempo de cálculo, considerou-se preferível a opção com menor

tamanho da população, cujo tempo de execução é 1124s por corrida. De notar que, devido aos resultados terem sido obtidos após a reunião dos conjuntos de soluções não dominadas, resultantes das 11 corridas do processo de otimização, o conjunto final de soluções não dominadas é constituído por 419 soluções.

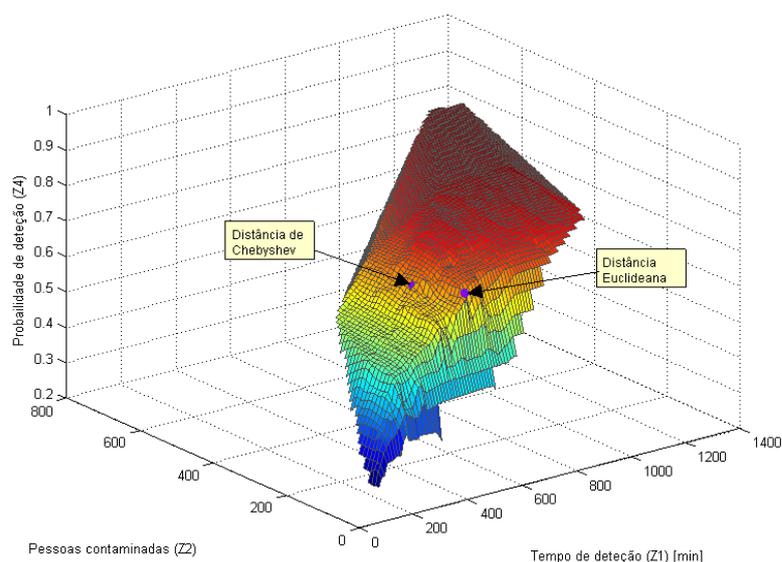
**Quadro 23. Resumo dos melhores resultados obtidos em cada FO, na otimização triobjetivo em função dos parâmetros, na otimização com NSGA-II.**

População	Prob. Mutação	Prob. Recombinação	Máx. Gerações	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [gal]	Z <sub>4</sub>
100	0.2	0.8	200	171.68	63	976.532	0.818
200				167.55	63	691.147	0.839
500				158.12	61	619.114	0.824

**Quadro 24. Soluções com menor distância à solução ideal na otimização triobjetivo, na otimização com NSGA-II**

FO otimizadas	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclideana	Distância Chebyshev	Solução				
Z <sub>1</sub> -Z <sub>2</sub> -Z <sub>4</sub>	554.48	125	1974.817	69.17%	0.613		18	22	69	84	102
	510.53	234	4358.600	67.12%		0.409	18	69	83	102	122

A Figura 26 representa as duas soluções na frente de Pareto, sendo de assinalar que a solução que minimiza a distância Euclideana corresponde à solução determinada com a mesma métrica com SA e a solução que minimiza a distância de Chebyshev apenas difere da solução análoga obtida com SA num sensor (122 por troca com 121, um sensor adjacente).



**Figura 26. Localização na frente de Pareto das soluções com menor distância à solução ideal, na otimização com NSGA-II.**

#### 4.3.2.3. Abordagem Tetra-objetivo

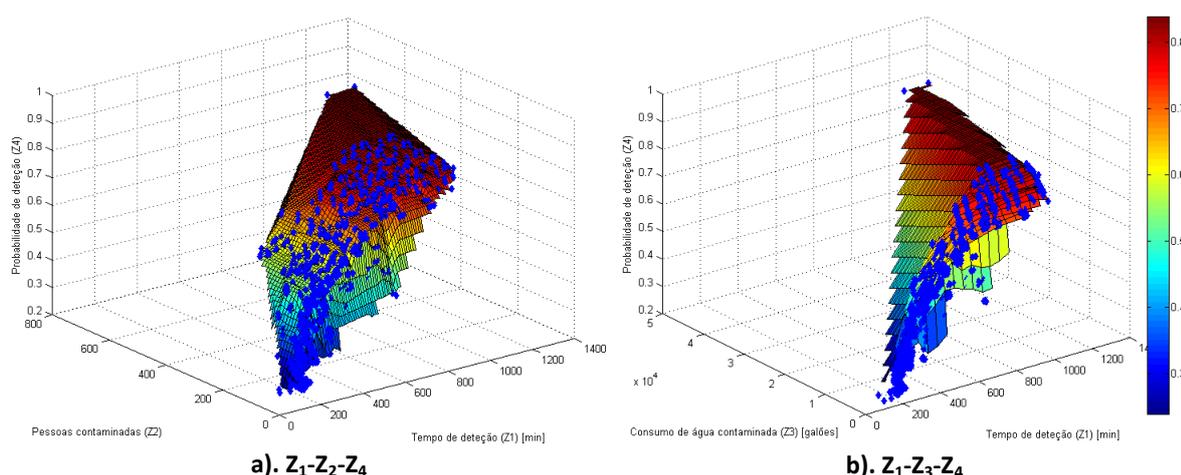
Tal como na secção anterior, procurou-se reproduzir a otimização simultânea das quatro FO repetindo os mesmos conjuntos de parâmetros, nomeadamente com tamanho da população de 100, 200 e 500. Os resultados deste processo, obtidos após 11 corridas estão

indicados no Quadro 25. Nestes resultados salienta-se que com uma população de 500 elementos é possível atingir os ótimos individuais de cada uma das FO em 5729 segundos, embora com qualquer um dos conjuntos de parâmetros se tenham obtido resultados muito abrangentes.

**Quadro 25. Resumo dos melhores resultados das diferentes FO, obtidos na otimização tetra-objetivo em função dos parâmetros, com NSGA-II.**

População	Prob. Mutação	Prob. Recombinação	Máx. Gerações	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [gal]	Z <sub>4</sub>
100	0.2	0.8	200	162.58	59	619.114	83.32%
200				156.48	61	619.114	81.77%
500				151.71	59	619.114	83.92%

A Figura 27 mostra duas representações tridimensionais da frente de Pareto à semelhança do que foi efetuado com SA para a otimização tetra-objetivo, sendo visível a semelhança com a Figura 18.



**Figura 27. Visualização tridimensional dos resultados obtidos na otimização tetra-objetivo com NSGA-II.**

Da mesma forma, o cálculo das métricas de distância à solução ideal, representadas no Quadro 26, permitiu obter soluções muito próximas, sendo a solução que minimiza a distância Euclidiana a mesma que foi obtida com SA, coincidindo também nos resultados da otimização triobjetivo com ambas as heurísticas. A solução que minimiza a distância de Chebyshev difere da solução obtida na otimização triobjetivo em apenas um sensor (69 pelo 71).

**Quadro 26. Soluções com menor distância à solução ideal na otimização tetra-objetivo, com NSGA-II.**

FO otimizadas	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclidiana	Distância Chebyshev	Solução				
							18	22	69	84	102
Z <sub>1</sub> -Z <sub>2</sub> -Z <sub>3</sub> -Z <sub>4</sub>	554.48	125	1974.817	69.17%	0.617		18	22	69	84	102
	505.24	264	5553.856	67.12%		0.413	18	83	71	102	122

### 4.3.3. Análise de resultados com AG

Replicando a análise efetuada em 4.2.3, em que um cenário possível seria só aceitar soluções com uma probabilidade de deteção superior a 75%, determinou-se a frente de Pareto resultante da otimização com 4 FO, condicionada por esta restrição. O resultado deste processo foi praticamente igual ao obtido com SA, encontrando-se as mesmas 3 soluções e a mesma superfície ilustrada na Figura 20. Por esse motivo entendeu-se ser redundante qualquer prosseguimento da análise de resultados, uma vez que esta já tinha sido efetuada em 4.2.3.

## 4.4. Determinação de soluções para alguns dos restantes casos propostos na BWSN

Após a estabilização do processo de otimização mais adequado para a resolução do caso A da rede 1, tal como proposto na competição BWSN, nesta secção pretendeu-se analisar o que seria a aplicação desse processo aos restantes casos propostos, incluindo a implantação de 20 sensores, para a rede 1 (caso R1A20) e o caso A para a rede 2 com 20 sensores (caso R2A20).

Os casos B, C e D para a rede 1 não foram analisados por se entender que acrescentam muito pouco ao processo, uma vez que apenas os dados numéricos são diferentes. Considerou-se igualmente pouco relevante o caso R2A5 por parecer inverosímil proteger uma rede de 12 523 nós com apenas 5 sensores.

### 4.4.1. Rede 1 com 20 sensores

A otimização da rede 1 com 5 sensores serviu de base para definir os conjuntos de parâmetros a utilizar na otimização mono-objetivo da rede para a implantação de 20 sensores, primeiro com SA, utilizando uma estrutura de vizinhança com a troca de apenas um sensor e uma temperatura mínima de  $10^{-3}$ , conforme se pode verificar pelo Quadro 27, tendo-se obtido os ótimos individuais de cada FO, listados no Quadro 28.

Quadro 27. Resumo dos parâmetros testados na otimização mono-objetivo com SA, para o caso R1A20.

Função otimizada	$T_0$	$\alpha$	n_passos
$Z_1$	1000	0.01	200
$Z_2$	500		
$Z_3$	1000	0.05	
$Z_4$	1	0.01	

**Quadro 28. Soluções para cada função objetivo no caso R1A20, na otimização mono-objetivo com SA.**

Função otimizada	Valor da solução				Sensores									
	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [galões]	Z <sub>4</sub> [%]										
Z <sub>1</sub>	20.43	47	713.22	20.09%	24	25	27	30	31	35	36	38	40	41
					42	43	44	45	49	50	51	52	54	124
Z <sub>2</sub>	137.15	22	157.22	35.84%	22	25	27	29	31	32	33	38	40	41
					42	43	51	56	57	59	66	93	109	114
Z <sub>3</sub>	105.71	27	64.59	33.24%	21	23	28	30	31	32	35	36	40	43
					44	51	52	53	54	56	59	65	66	109
Z <sub>4</sub>	615.25	436	17171.27	89.99%	11	13	20	35	36	46	62	66	80	84
					86	92	99	101	107	109	113	122	123	125

De igual modo, procedeu-se à otimização com AG usando os conjuntos de parâmetros que permitiram obter melhores resultados para a rede com 5 sensores. No entanto, apenas se conseguiu obter o valor de Z<sub>4</sub> indicado no Quadro 28, sendo os valores obtidos nas outras FO piores que os verificados com SA, como se pode constatar pelo Quadro C 9 do Apêndice C.

Com base na informação relativa aos ótimos individuais, procedeu-se à otimização triobjetivo, já que da experiência anterior se tinha concluído que era suficiente, devido à correlação existente entre Z<sub>2</sub> e Z<sub>3</sub>, diminuindo assim o tempo de execução. Tal como anteriormente, foram efetuadas 11 corridas do processo, primeiro com SA, a partir das quais se obteve a frente de Pareto consolidada apresentada na Figura 28.

À semelhança do que foi feito anteriormente, foram também calculadas as métricas de distância à solução ideal  $Z^* = (0,0,1)$ , tendo-se obtido os valores apresentados no Quadro 29.

**Quadro 29. Soluções com menor distância à solução ideal na otimização triobjetivo (caso R1A20), com SA.**

FO otimizadas	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclideana	Distância Chebyshev	Solução									
Z <sub>1</sub> -Z <sub>2</sub> -Z <sub>4</sub>	210.43	60	605.68	78.53%	0.58		7	15	18	22	32	36	47	50	65	69
							71	75	77	81	83	85	91	102	103	121
	210.50	80	710.95	79.96%		0.35	5	22	32	35	36	46	65	69	73	76
							77	83	85	91	99	100	111	116	119	121

As duas soluções indicadas no Quadro 29 apresentam resultados muito próximos, sendo a solução determinada pela menor distância Euclideana melhor em Z<sub>2</sub> e Z<sub>3</sub>, enquanto a distância de Chebyshev resulta numa solução com melhores valores em Z<sub>4</sub>. Pela Figura 29 verifica-se que as duas soluções se encontram muito perto do patamar que corresponderá às melhores soluções em Z<sub>4</sub>, pelo que se admite que um decisor final potencialmente optaria por uma solução nessa região, que minimizasse as restantes funções objetivo de acordo com uma qualquer estrutura de preferências.

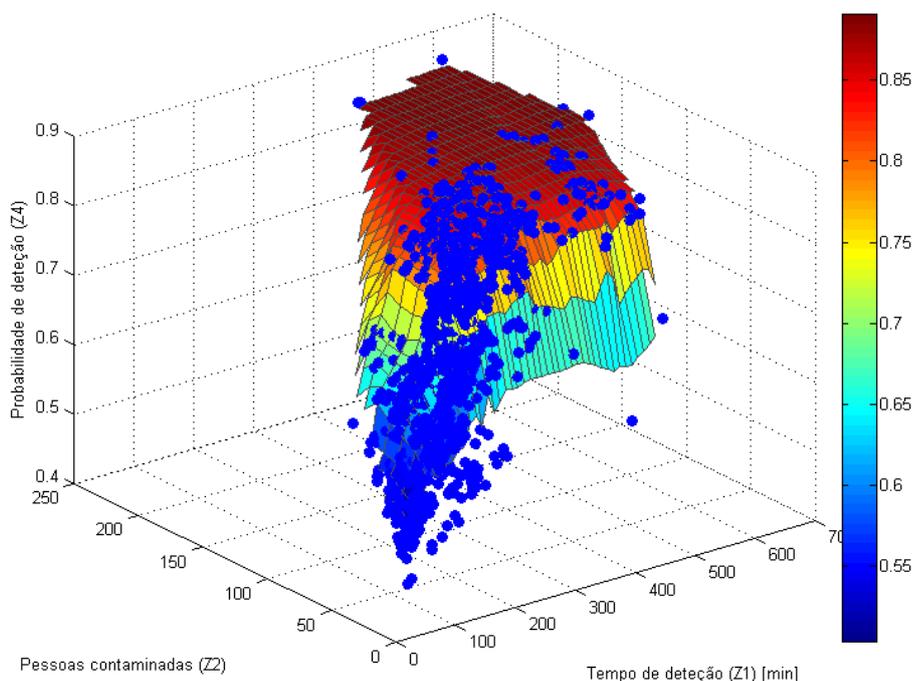


Figura 28. Frente de Pareto para a otimização triobjetivo ( $Z_1$ ,  $Z_2$ ,  $Z_4$ ), para o caso R1A20, com SA.

Assim, considerando o mesmo tipo de pressupostos anteriormente utilizados na análise da localização para 5 sensores, admitiu-se, a título de exemplo, que com a gama de resultados obtidos se procurariam soluções com uma probabilidade de detecção ( $Z_4$ ) superior a 85%. O resultado da imposição desse limite de aceitabilidade encontra-se ilustrado na Figura 30. Com a frente reduzida resultante será mais fácil ajudar um decisor final a identificar a solução que melhor satisfaz os seus requisitos, tendo-se ilustrado na figura duas possíveis opções, que não diferem muito em termos de probabilidade de detecção mas que ainda apresentam alguma diferenciação nas restantes funções objetivo.

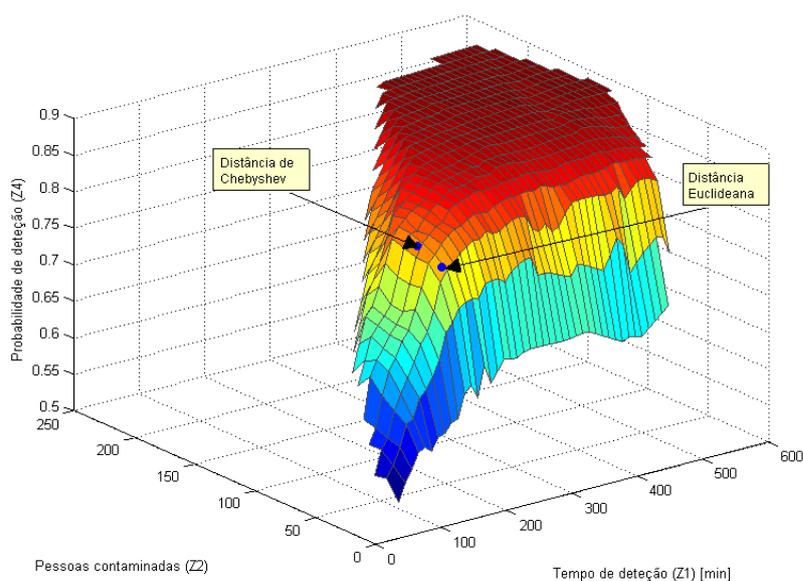


Figura 29. Localização na frente de Pareto das soluções com menor distância à solução ideal, para o caso R1A20, com SA.

Em comparação com os resultados apresentados por Ostfeld *et al.* (2008), verifica-se que a frente de Pareto com 928 soluções se ajusta muito bem às soluções aí propostas (Figura 31). De notar que das 13 soluções listadas em (Ostfeld *et al.*, 2008) relativamente a este caso, apenas 8 eram soluções não dominadas entre elas. Destas, 4 são dominadas pelas soluções pertencentes à frente de Pareto que, no entanto, também tem 35 soluções que são dominadas pelas apresentadas na BWSN. A solução escolhida de entre as duas discutidas no parágrafo anterior mantém-se como uma solução não dominada, e apresenta como resultados  $Z_1=341.5$  min,  $Z_2=91$  pessoas;  $Z_3=988$  galões;  $Z_4=87.01\%$ , correspondendo ao conjunto de sensores indicado no Quadro 30, ficando muito próxima de algumas apresentadas na BWSN, nomeadamente a apresentada por Eliades & Polycapou (2006).

Quadro 30. Solução proposta R1A20.

Posição	11	13	18	20	22	32	36	46	66	68
	73	75	84	91	99	101	104	111	117	122
Etiqueta dos nós	10	12	17	19	21	31	35	45	65	67
	72	74	83	90	98	100	103	112	118	123

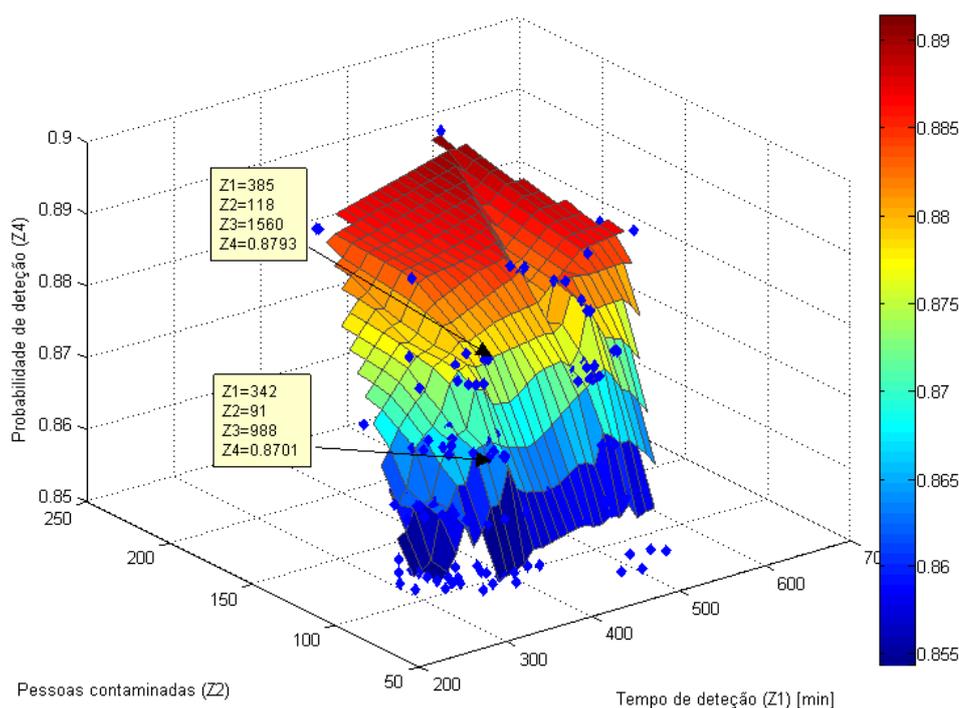
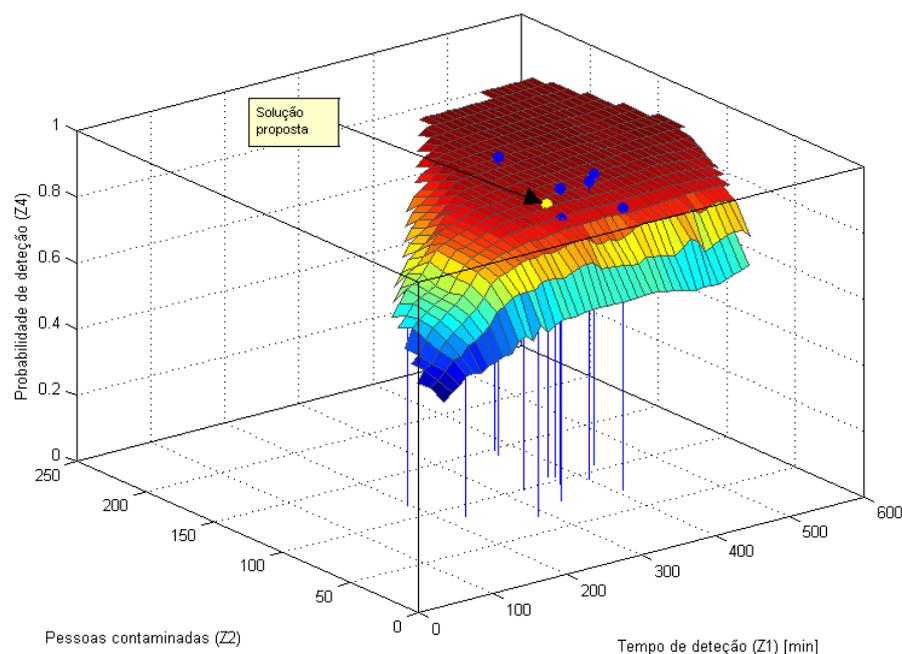


Figura 30. Frente de Pareto reduzida a soluções com probabilidade de deteção superior a 85%, para o caso R1A20, com SA.



**Figura 31. Representação da solução proposta na frente de Pareto e comparação com soluções apresentadas na BWSN para o caso R1A20.**

#### 4.4.2. Rede 2 com 20 sensores

Como já referido anteriormente, a segunda rede de teste (rede 2) possui 12 523 nós, colocando problemas adicionais à aplicação do método usado para a rede 1. Este problema começa nos requisitos de memória para armazenar os resultados das simulações hidráulicas e de qualidade da água, mas é particularmente notório no tempo necessário para esse procedimento, que pode levar mais de uma semana, mesmo simulando apenas dois eventos aleatórios em cada nó tal como foi realizado na BWSN. Para este trabalho, optou-se por usar as matrizes disponibilizadas pela Universidade de Exeter, para evitar, por um lado, o tempo de execução inerente e, por outro, para garantir a comparabilidade de resultados uma vez que, não tendo uma simulação de todos os eventos possíveis, não seria possível obter exatamente as mesmas matrizes de avaliação. Para minimizar os requisitos de memória optou-se por armazenar as matrizes  $Z_1$  como *array* de inteiros de 16 bits, e a matriz  $Z_4$ , por ser constituída apenas por 0 e 1, como *array* de inteiros de 8 bits. As matrizes  $Z_2$  e  $Z_3$ , sendo de números reais, foram armazenadas em *arrays* de *float* (32 bits) por estarem nos ficheiros originais com apenas 7 algarismos significativos.

O problema agora em estudo, consiste na colocação de 20 sensores numa rede com 12523 nós, o que corresponde a  $3642.6 \times 10^{60}$  combinações, ou seja com uma dimensão impossível de estudar de forma exaustiva, sendo no entanto passível de análise por uma metaheurística com vista à obtenção de uma solução de “boa qualidade”.

Seguindo a metodologia que tem vindo a ser seguida, em primeiro lugar determinaram-se os ótimos individuais de cada FO aplicando as metaheurísticas SA e AG, com parâmetros baseados nos que foram usados anteriormente com a Rede 1, mas que foram depois sujeitos a aperfeiçoamento iterativo. O Quadro 31 mostra os melhores valores obtidos com SA e o Quadro 34 os obtidos com AG.

Os resultados obtidos mostram mais uma vez que é mais fácil conseguir parametrizar SA para obter melhores resultados, embora implicando um maior tempo de cálculo. Os valores obtidos nas FO para  $Z_1$ ,  $Z_2$  e  $Z_3$  foram em qualquer dos casos muito inferiores aos indicados nas soluções listadas em (Ostfeld *et al.*, 2008), o que permitia esperar bons resultados na otimização multiobjetivo. Porém, o melhor valor de  $Z_4$  é inferior ao valor indicado numa das soluções apresentadas à BWSN, e apenas foi obtido com SA uma vez em onze corridas, com um conjunto de parâmetros que implicou um tempo de computação bastante longo, como se pode verificar no Quadro 32. Os resultados com AG, embora piores, foram obtidos num tempo muito menor como se pode verificar no Quadro 33. No entanto, experiências posteriores com alterações significativas nos parâmetros, incluindo o número máximo de gerações, não permitiram melhorar os resultados de forma visível.

**Quadro 31. Soluções para cada função objetivo para o caso R2A20, na otimização mono-objetivo com SA.**

Função otimizada	Valor da solução				Sensores						
	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [galões]	$Z_4$ [%]							
$Z_1$	<b>111.26</b>	183	4706.57	4.14%	716	1733	1753	2030	3073	4713	5552
					5724	6804	7201	7279	7966	8710	8784
					8792	9107	9599	10826	12110	12235	
$Z_2$	315.10	<b>125</b>	2327.26	6.43%	1795	1864	2887	2990	3329	3358	4714
					5295	5686	6393	6458	6491	6722	6784
					6804	7984	8123	8883	10068	11454	
$Z_3$	201.93	87	<b>1259.90</b>	5.28%	1733	1801	2328	2986	4410	5110	5240
					6016	6076	6393	7178	7337	7656	8032
					8147	8383	8715	10359	11238	11766	
$Z_4$	1665.61	2497.00	260553.3	<b>41.62%</b>	501	734	1487	1522	2668	2731	2866
					3165	3670	3724	4209	4368	6960	7437
					7929	9047	9111	9996	10645	10781	

**Quadro 32. Conjunto de parâmetros utilizados na otimização mono-objetivo, para o caso R2A20 com SA.**

Função	$T_0$	$\alpha$	n_passos	$T_{\text{mínimo}}$	Tempo execução [s]	Iterações
$Z_1$	500	0.01	200	0.0001	4430	1543
$Z_2$					4126	1543
$Z_3$					4164	1543
$Z_4$	1	0.01		0.000001	4117	1382

Quadro 33. Resumo dos parâmetros usados para o caso R2A20 com AG.

População	Probabilidade Recombinação	Probabilidade Mutação	Máximo Gerações	Elitismo	Tempo execução [s]
100	0.8	0.2	150	5%	231

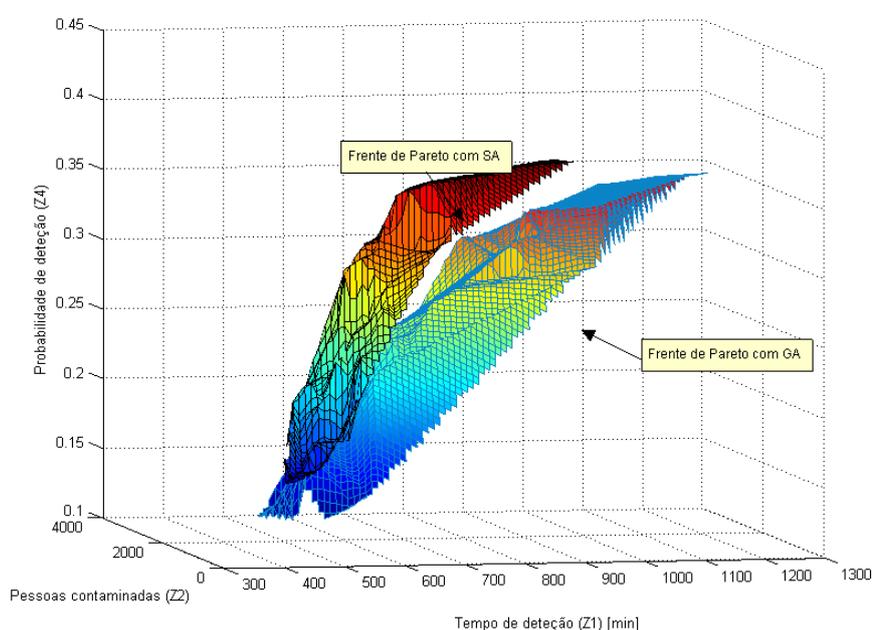
Quadro 34. Soluções para cada função objetivo para o caso R2A20, na otimização mono-objetivo com AG.

Função otimizada	Valor da solução				Sensores						
	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [galões]	Z <sub>4</sub> [%]							
Z <sub>1</sub>	<b>327.46</b>	418	13025.28	11.61%	110	1035	1445	1760	1827	2539	2546
					2683	4043	4470	4747	5208	6113	6523
					7975	8061	9141	11178	11240	12043	
Z <sub>2</sub>	436.75	<b>359</b>	11121.51	9.09%	2111	2531	2594	3314	3365	4441	5759
					5960	6402	6467	7247	7463	7942	7962
					7984	8883	9136	10866	11593	11670	
Z <sub>3</sub>	329.85	334	<b>8219.88</b>	10.53%	1645	2059	2180	4033	4378	5115	5653
					6864	7112	7199	7262	8668	8710	8860
					9419	9905	10380	11221	11314	12250	
Z <sub>4</sub>	1162.41	1969	173295.31	<b>37.73%</b>	12	936	1100	1853	3601	3677	3713
					3863	4207	4378	5996	6587	7320	7703
					8200	9259	9365	10649	10875	12378	

À semelhança dos casos anteriores, optou-se por realizar a otimização triobjetivo devido à correlação existente entre Z<sub>2</sub> e Z<sub>3</sub>. Os resultados obtidos com SA e AG permitem mais uma vez constatar uma maior facilidade em parametrizar SA para a obtenção de melhores resultados, embora sempre à custa de um tempo de computação considerável. Na Figura 32 representam-se as duas frentes de Pareto obtidas, verificando-se que a frente obtida com SA domina a frente obtida com AG. As distâncias Euclideana e de Chebyshev à solução ideal são dadas no Quadro 35. Verifica-se que a solução com menor distância Euclideana é uma solução com bons valores em Z<sub>1</sub>, Z<sub>2</sub> e Z<sub>3</sub>, mas correspondendo a uma probabilidade de deteção muito baixa. Já a solução que minimiza a distância de Chebyshev melhorou significativamente o valor do objetivo Z<sub>4</sub> aproximando-se do melhor valor encontrado na otimização individual. Os resultados obtidos com AG são genericamente piores, como se pode verificar no Quadro C 11 do Apêndice C.

**Quadro 35. Soluções para o caso R2A20, com a menor distância à solução ideal na otimização triobjetivo, com SA**

Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclideana	Distância Chebyshev	Solução				
417.97	491	14420.47	15.44%	1.16		652	1494	2224	2260	2384
						2580	3891	5308	5309	5791
						6726	7682	8755	8838	9000
						9747	9933	10947	11178	11252
707.52	917	43952.34	36.75%		0.75	637	1579	3236	3329	3674
						3750	3838	4033	4814	6518
						6584	6874	7206	7312	8177
						9230	10408	10875	11168	11218



**Figura 32. Sobreposição das frentes de Pareto obtidas para o caso R2A20.**

Embora não seja possível determinar definitivamente uma das soluções não dominadas como sendo a melhor solução, o que requereria a intervenção da expressão de preferências do decisor final, procedeu-se a uma análise semelhante à anteriormente efetuada, identificando a solução não dominada com os valores  $Z_1=697$  min;  $Z_2=863$  pessoas;  $Z_3=41256$  galões e  $Z_4=36,6\%$ , correspondendo ao conjunto de sensores indicado no Quadro 36 e que está representada graficamente na Figura 33, como podendo satisfazer um decisor que procurasse um equilíbrio entre a melhor garantia de detecção dos eventos com o menor impacto em termos de tempo de detecção e pessoas contaminadas.

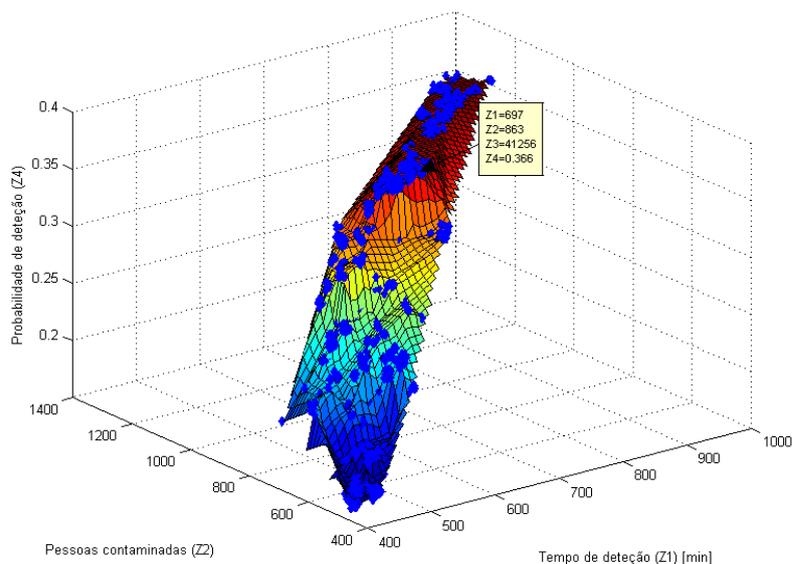


Figura 33. Análise da Frente de Pareto resultante da otimização triobjetivo, para o caso R2A20.

Quadro 36. Solução proposta R2A20.

Posição	637	1580	3127	3236	3329	3673	3750	3837	4033	4814
	6523	6584	6986	7161	7312	8177	10408	10741	10875	11168
Etiqueta dos nós	636	1579	3126	3235	3328	3672	3749	3836	4032	4813
	6522	6583	6985	7160	7311	8176	10407	10740	10874	11167

Os resultados obtidos, por comparação com as soluções listadas em (Ostfeld *et al.*, 2008), revelaram menor capacidade do método anteriormente aplicado nesta dissertação em gerar soluções de compromisso que garantissem “bons” valores em  $Z_4$ . Assim, dada a importância atribuída a esta última FO, tornou-se necessário procurar formas de melhorar o desempenho neste domínio. Uma hipótese contemplada e que consta como pista de trabalho em (Ostfeld *et al.*, 2008), consiste em efetuar uma agregação da rede, obtendo uma rede reduzida equivalente, com menos nós. Para o efeito, efetuou-se uma análise dos resultados das simulações, nomeadamente da matriz  $Z_4$  que indica a capacidade de deteção por cada nó, de eventos ocorridos em qualquer ponto da rede. A Figura 34 mostra que existe um número significativo de nós (aproximadamente 7000) que deteta menos de 500 eventos de contaminação, ou seja, menos de 2% dos eventos simulados.

Testou-se assim uma abordagem que consiste em restringir a análise aos nós com mais de 500 deteções, para a qual foi necessário reimplementar a função de caracterização de rede, com vista a reestabelecer as ligações entre nós após a eliminação das colunas das matrizes de  $Z_1$  a  $Z_4$  correspondentes à restrição imposta. As rotinas anteriormente implementadas podem ser assim aplicadas às matrizes reduzidas, com a lista de nós vizinhos adaptada pela rotina descrita no Algoritmo 10, sendo depois necessário fazer a correspondência entre a nova numeração dos nós e a numeração e etiquetas originais.

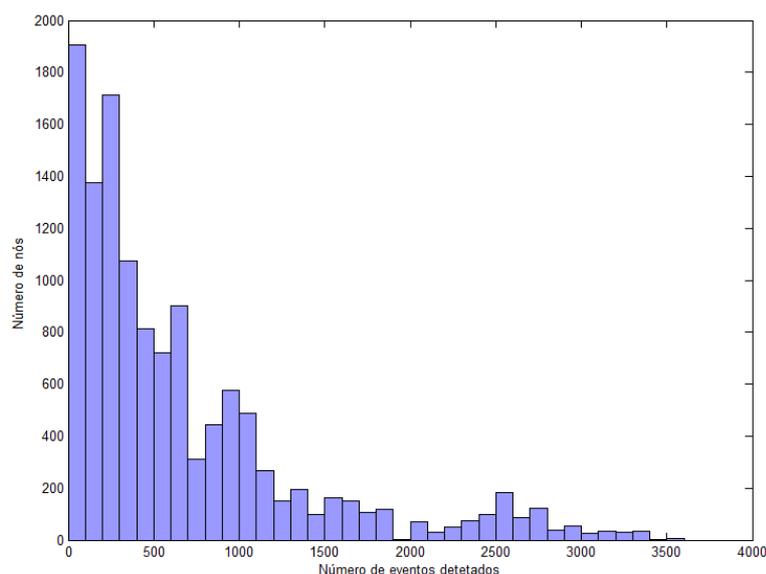
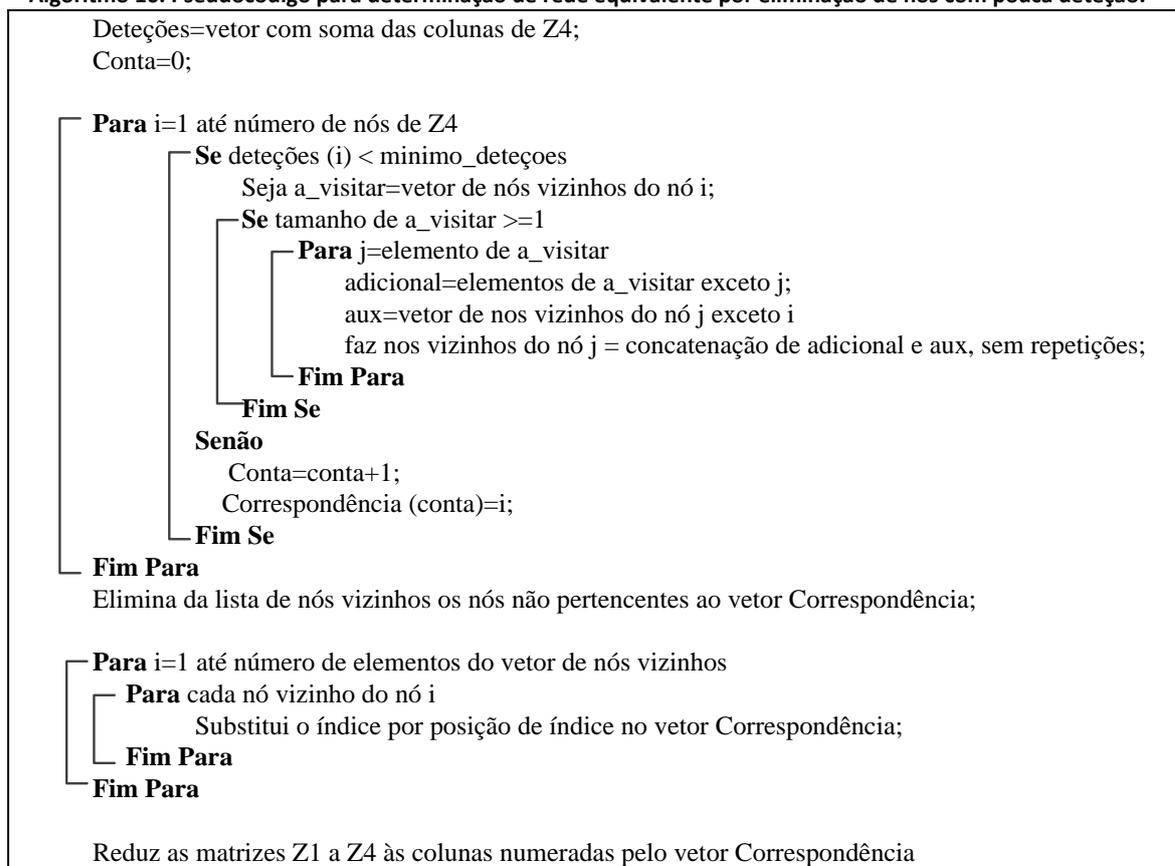


Figura 34. Histograma do número de detecções em cada nó para o caso R2A20.

Os resultados deste processo estão descritos no Quadro 37 em termos de otimização mono-objetivo, representando a Figura 36 a frente de Pareto de uma otimização triobjetivo em  $Z_1$ - $Z_2$ - $Z_4$ , onde estão representadas as soluções que minimizam as métricas de distância euclidiana e de Chebyshev à solução ideal (Quadro 38) e uma exemplificação do que poderia ser o estabelecimento de compromissos através da análise do gráfico, em que se procurou reproduzir o raciocínio anteriormente usado. Nestas condições é possível constatar a obtenção de um maior número de soluções com valores de  $Z_4$  próximos dos melhores valores apresentados à BWSN. A comparação entre a frente de Pareto obtida e as soluções listadas em (Ostfeld *et al.*, 2008) permite identificar que um pequeno número destas últimas dominam a frente de Pareto obtida, mas as restantes enquadram-se ou são dominadas por esta, como pode ser observado na Figura 36.

A solução identificada pelo estabelecimento de compromissos aceitáveis para um potencial decisor, a partir de uma abordagem interativa tendo por base a representação tridimensional da frente de Pareto, e que corresponde a uma potencial boa solução, resultou novamente de se procurar a zona em que ocorre uma maior variação positiva em  $Z_4$  com menor degradação nos restantes objetivos, através da identificação de uma face quase vertical na frente. A solução escolhida localiza-se no limite superior dessa face, com os valores de  $Z_1=689$  min;  $Z_2=872$  pessoas;  $Z_3=40333$  galões e  $Z_4=37,11\%$ , correspondendo ao conjunto de sensores indicado no Quadro 39. Esta solução é muito próxima da anteriormente determinada considerando a totalidade dos nós, sendo ligeiramente melhor em todos os objetivos com exceção de  $Z_2$ .

**Algoritmo 10. Pseudocódigo para determinação de rede equivalente por eliminação de nós com pouca detecção.****Quadro 37. Soluções para cada função objetivo para o caso R2A20, na otimização mono-objetivo, desprezando nós com menos de 500 deteções, com SA.**

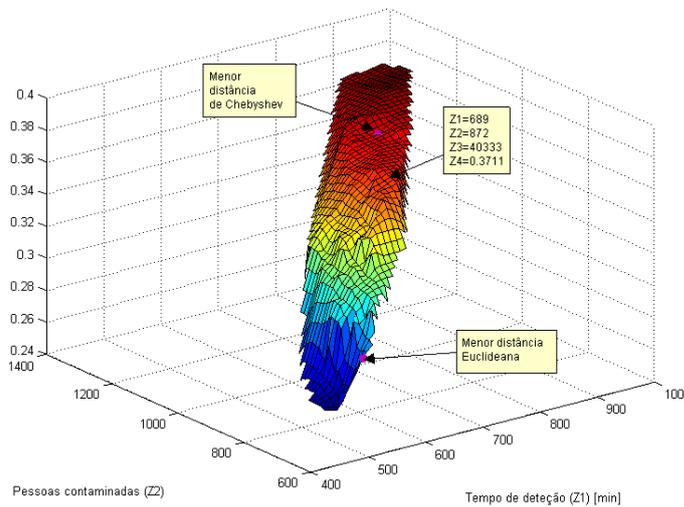
Função otimizada	Valor da solução				Sensores						
	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [galões]	Z <sub>4</sub> [%]	81	109	592	1827	2074	4033	4253
Z <sub>1</sub>	247.62	378	10769	9.26%	4644	4745	4794	5586	5641	6286	7223
					7616	9000	10363	11127	11246	11919	
Z <sub>2</sub>	288.41	425	11496.69	13.69%	977	1242	2074	4033	4137	4245	4992
					5076	5115	5586	6650	6752	7223	7256
					7661	9000	10595	11127	11909	12326	
Z <sub>3</sub>	289.24	405	10670.23	14.56%	977	1198	1242	2074	4033	4137	4233
					4245	5115	5586	6650	7223	7256	7275
					9000	10412	10628	11192	11228	11909	
Z <sub>4</sub>	1582.23	2476	251309.14	44.55%	303	376	501	624	734	814	823
					1487	2866	3670	3724	3834	4209	4307
					7437	7666	8355	8493	8954	9047	

**Quadro 38. Soluções para o caso R2A20, com a menor distância à solução ideal na otimização triobjetivo, com SA, desprezando nós com menos de 500 detecções.**

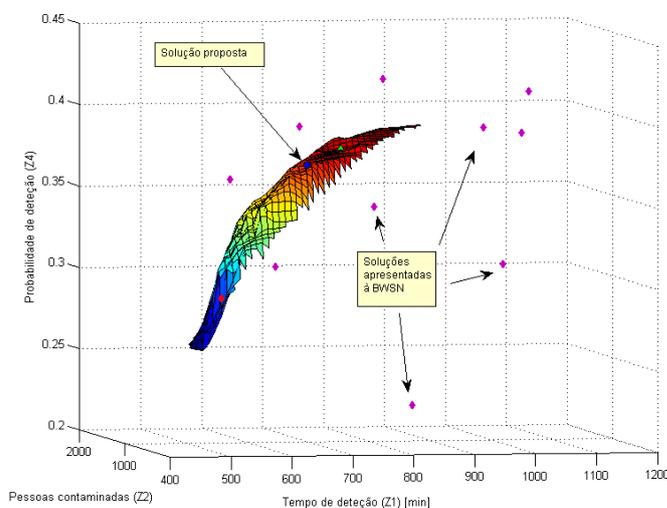
Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclidiana	Distância Chebyshev	Solução						
534.92	674	26211.210	29.16%	1.245		298	743	1002	1918	3358	3383	4047
						4222	4553	4814	5586	7998	9000	9205
						9684	10127	10615	10704	10875	11192	
755.06	1014	51882.838	37.95%		0.834	637	839	878	1487	3138	3358	4039
						4235	4826	4991	5586	7437	7994	8390
						9049	9625	9978	10307	10875	11178	

**Quadro 39. Solução proposta para o caso R2A20, na otimização triobjetivo com SA.**

Posição	637	1479	2291	3359	3444	3574	3678	3769	4160	4183
	5775	6310	7257	7437	7577	7627	9000	9143	9365	10704
Etiqueta dos nós	636	1478	2290	3358	3443	3573	3677	3768	4159	4182
	5774	6309	7256	7436	7576	7626	8999	9142	9364	10703



**Figura 35. Representação de soluções propostas na frente de Pareto para a otimização triobjetivo (Z<sub>1</sub>, Z<sub>2</sub>, Z<sub>4</sub>), para o caso R2A20, com SA, desprezando nós com menos de 500 detecções.**



**Figura 36. Comparação da frente de Pareto da otimização triobjetivo (Z<sub>1</sub>, Z<sub>2</sub>, Z<sub>4</sub>) com SA, para o caso R2A20, desprezando nós com menos de 500 detecções, com as apresentadas em (Ostfeld *et al.*, 2008).**

## 5. CONCLUSÕES

### 5.1. Trabalho desenvolvido

Esta dissertação tinha como objetivo a aplicação de métodos de otimização baseados em metaheurísticas a um problema de engenharia que tem merecido a atenção de várias entidades (órgãos governamentais, reguladores e entidades gestoras), consistindo na determinação da melhor localização para a implantação de sensores com vista à detecção de eventos de contaminação em sistemas de abastecimento de água, acidentais ou propositados. Este assunto tem suscitado o interesse de vários investigadores, tendo existido uma competição específica, *The Battle of the Water Sensor Networks* (Ostfeld *et al.* 2008), que se tornou uma referência para esta área de trabalho, definindo regras de comparação de resultados, nomeadamente as redes de teste e as quatro funções objetivo que foram usadas nesta dissertação: o tempo de deteção ( $Z_1$ ), a população afetada ( $Z_2$ ), o consumo de água contaminada ( $Z_3$ ) e a probabilidade de deteção de eventos ( $Z_4$ ).

Dado que as metaheurísticas não garantem a obtenção da solução “ótima”, um aspeto a que se atribuiu particular importância foi procurar garantir que o processo de pesquisa utilizado proporcionasse a maior abrangência do espaço de soluções. Nesse sentido, as metaheurísticas foram implementadas em variantes mono e multiobjetivo com o objetivo de, por um lado, determinar os melhores valores individuais de cada uma das funções objetivo consideradas e, por outro lado, explorar as relações entre as funções objetivo, analisando *trade-offs* entre elas e determinando soluções de compromisso. Procurou-se igualmente analisar a maior ou menor relevância do uso de todas as funções objetivo consideradas. As abordagens testadas foram baseadas em *simulated annealing* e *algoritmos genéticos*, procurando explorar a diferença entre um método de trajetória, com base na pesquisa em torno de uma solução, e um método de população, com uma pesquisa paralela de várias soluções. Na otimização multiobjetivo utilizou-se uma implementação do *simulated annealing* baseada em (Suppaitnarm e Parks, 1999) e o NSGA-II de (Deb *et al.*, 2002).

Uma das características inerentes ao uso de metaheurísticas é a necessidade de definir um conjunto de parâmetros que são fundamentais para o bom desempenho das metodologias. Nesse sentido, uma parte substancial da dissertação consistiu em procurar determinar de uma forma sistemática os parâmetros que levavam ao melhor desempenho acima descrito, isto é, permitindo atingir de uma forma consistente os ótimos individuais das funções nas abordagens mono-objetivo, e aproximando-se o mais possível desses valores no contexto das abordagens multiobjetivo, de forma a assegurar a maior abrangência possível do espaço de

soluções. Os resultados obtidos mostraram uma maior facilidade em conseguir melhores soluções com *simulated annealing*. Embora as abordagens baseadas em algoritmos genéticos, em alguns casos proporcionassem soluções equivalentes e com um menor esforço computacional, a menor taxa de sucesso verificada obriga à necessidade de efetuar um número mais elevado de corridas do processo de otimização para garantir os melhores resultados, o que acaba por anular parcialmente a vantagem.

O trabalho nesta dissertação incidiu essencialmente na primeira rede de teste com 129 nós e para a localização de cinco sensores, caso identificado como R1A5. Os resultados obtidos mostraram que dada a correlação existente entre os objetivos  $Z_2$  e  $Z_3$ , a otimização com 4 e com 3 objetivos, elegendo apenas  $Z_1$ ,  $Z_2$  e  $Z_4$ , conduzia a soluções idênticas com vantagem para a otimização triobjetivo devido a um menor tempo computacional, embora a otimização apenas com os objetivos  $Z_1$  e  $Z_4$  tenha identificado soluções muito próximas das atrás referidas. De salientar que a frente de Pareto obtida enquadrava todas as soluções identificadas em (Ostfeld *et al.*, 2008), dominando seis das apresentadas neste artigo.

O caso da colocação de 20 sensores na rede 1 não foi alvo de uma pesquisa sistemática de parâmetros, tendo-se adotado os que produziram melhores resultados na situação anterior, para executar uma otimização apenas triobjetivo em  $Z_1$ ,  $Z_2$  e  $Z_4$ . Assim, muito embora se tenha obtido uma frente de Pareto com 928 soluções não dominadas, verificou-se que 35 soluções da frente de Pareto eram dominadas por algumas das soluções de (Ostfeld *et al.*, 2008), embora também se tenha verificado o contrário, e apenas 4 das 13 soluções ali apresentadas pudessem constar de uma reunião das frentes de Pareto, a qual teria assim 897 soluções não dominadas. Porém, as soluções que previamente tinham sido identificadas como potencialmente interessantes permaneceram não dominadas, mantendo-se assim como hipóteses elegíveis para uma solução de compromisso final.

Segundo Antunes *et al.* (2012), num problema multiobjetivo torna-se necessário fazer intervir no processo de decisão não apenas meios técnicos de calcular soluções não dominadas, mas também informação sobre as preferências do decisor que permitam discriminar soluções. O problema multiobjetivo pode então colocar-se como a escolha, de entre os elementos do conjunto de soluções não dominadas, de uma que constitua uma solução de compromisso aceitável pelo decisor tendo em atenção as suas preferências. No entanto, não se encontraram nas regras da BWSN, nem foram indicadas em (Ostfeld *et al.*, 2008), quaisquer elementos sobre esta questão que permitissem explicar totalmente as soluções ali listadas. Em todos os artigos consultados referentes às participações na competição, que usaram abordagens multiobjetivo, a única estratégia apresentada em alguns dos casos, foi a atribuição de pesos iguais a todas as funções objetivo.

Para poder comparar soluções entre as diferentes abordagens optou-se nesta dissertação por definir duas formas de decidir entre as soluções não dominadas. A primeira por

determinação da solução com menor distância à solução ideal, definida com base nas funções objetivo considerados em cada caso, de acordo com duas métricas habitualmente usadas para este fim, a distância Euclideana e a distância de Chebyshev. A segunda, por definição interativa de compromissos com um potencial decisor, partindo do conjunto de soluções resultantes do estabelecimento de um limite mínimo para a probabilidade de deteção, critério que foi assumido como fundamental, uma vez que um baixo valor nesta função objetivo, aliado à não contabilização nos outros objetivos dos eventos não detetados implicaria a não consideração de consequências potencialmente significativas dos eventos de contaminação. As soluções determinadas com base neste critério estão representadas na Figura 37, verificando-se que a solução obtida no caso R1A5 é um subconjunto da solução R1A20 e resultam ambas muito próximas das propostas por Eliades e Polycarpou (2006).

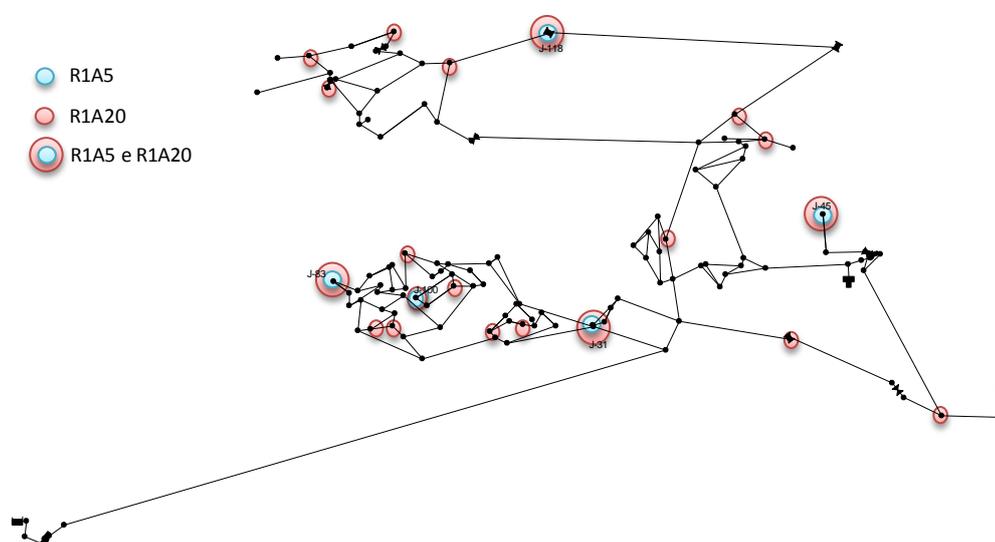


Figura 37. Representação das soluções propostas para a rede 1 com 5 e 20 sensores.

Embora se admita que as diferenças nos processos, nomeadamente na forma de normalizar as escalas dos diferentes objetivos, não permitam uma equivalência exata, o que explica as diferenças verificadas, foi possível identificar proximidades entre as soluções determinadas com base na distância mínima à solução ideal e várias soluções listadas em (Ostfeld *et al.*, 2008), dependendo da métrica escolhida ou do conjunto de objetivos otimizados, o que permite inferir algumas preferências estabelecidas para a escolha dessas soluções.

Relativamente ao estudo de caso correspondente à 2ª rede de teste verificou-se que a sua grande dimensão acarreta maiores dificuldades na resolução do problema, que começam pelos requisitos em termos de memória e tempo de processamento inicial, mas que também se refletem na maior dificuldade em atingir uma frente de Pareto que contenha de facto as melhores soluções. Verificou-se, em particular, alguma dificuldade em encontrar soluções com probabilidades de deteção ao nível das melhores soluções neste objetivo entre as listadas em (Ostfeld *et al.*, 2008).

Por este motivo, foi testada uma abordagem alternativa na qual se desprezaram os nós da rede que não permitissem um mínimo de 500 detecções, cerca de 2% dos eventos simulados, após se verificar que essa restrição reduziria a rede a cerca de metade. O resultado da abordagem revelou-se promissor, tendo determinado um maior número de soluções com um valor mais elevado de probabilidade de detecção.

## 5.2. Pistas de trabalho futuro

O trabalho realizado nesta dissertação permitiu identificar um conjunto de itens que ficaram por explorar por escassez de tempo e que poderão proporcionar avanços interessantes no tema desenvolvido.

O sucesso das metaheurísticas aplicadas à resolução do problema é muito influenciado por um conjunto de parâmetros cuja determinação não é fácil. Embora se tenha procurado responder a este problema através de uma pesquisa sistemática, reconhece-se que outras abordagens poderão vir a simplificar este processo, quer seja pela adoção de parâmetros dinâmicos, tal como experimentado com a variação da probabilidade de mutação e de recombinação em função da diversidade populacional no algoritmo genético, quer seja pela sua determinação através de uma combinação de metaheurísticas. No entanto, no caso do problema em análise, a morosidade dos cálculos associada à avaliação das soluções, nomeadamente no caso de redes de grande dimensão como a rede 2, cria dificuldades a este processo.

O estudo efetuado seguiu de perto as definições do desafio da BWSN procurando a colocação de um número pré-definido de sensores. Porém, o número adequado de sensores será certamente função da dimensão da rede, sendo provável que exista um ponto de equilíbrio adequado a cada caso, pelo que esta deveria ser um objetivo a incluir na análise, como identificado em (Eliades *et al.*, 2014).

As definições da BWSN estabeleceram que as funções objetivo  $Z_1$ ,  $Z_2$  e  $Z_3$  com as quais se avalia cada solução proposta, são determinados apenas para os eventos detetados por essa solução. A argumentação presente em (Ostfeld *et al.*, 2008) é de que a consideração dos eventos não detetados levaria a colocações não intuitivas dos sensores em locais que minimizassem a não detecção em detrimento da minimização do tempo de detecção. Para compensar este efeito foi definido o objetivo  $Z_4$  que mede a probabilidade de detecção, a maximizar. Porém, essa definição não parece garantir que entre duas soluções com igual probabilidade de detecção se consiga discriminar aquela em que os eventos não detetados afetam um número limitado de pessoas relativamente a outra em que, pelo contrário, não seja detetado um evento de grande dimensão. Este problema foi identificado em (Dorini *et al.*, 2010; Eliades e Polycarpou, 2006; Krause *et al.*, 2006). Hipóteses para ultrapassar este problema incluem a alteração da fórmula de cálculo com a consideração de todos os eventos detetados ou não, e a inclusão de um objetivo adicional que

permitisse diferenciar essas soluções. No entanto, tal pode implicar recalcular as matrizes em que se baseia a avaliação das soluções, um processo computacionalmente muito pesado, principalmente para a rede de grande dimensão, podendo corresponder a várias semanas de cálculo. Importa por isso procurar soluções que permitam dotar o processo de maior eficiência sem grande perda de qualidade dos resultados.



## REFERÊNCIAS BIBLIOGRÁFICAS

- Ailamaki, A., Faloutsos, C., Fischbeck, P.S., Small, M.J., VanBriesen, J., 2003. An Environmental Sensor Network to Determine Drinking Water Quality and Security. *SIGMOD Rec* 32, 47–52.
- Alperovits, E., Shamir, U., 1977. Design of optimal water distribution systems. *Water Resour. Res.* 13, 885–900.
- Alves, M.J.T.G., 2001. *Apoio à decisão em problemas de programação inteira e inteira-mista multiobjectivo: contribuições metodológicas*. Tese de doutoramento. Universidade de Coimbra
- Antunes, C.H., Alves, M.J., Clímaco, J., 2012. Tomada de decisão em ambiente multiobjectivo, in: *Manual de Computação Evolutiva e Metaheurística*. Imprensa da Universidade de Coimbra, pp. 323–356.
- Antunes, C.H., Lima, P., Oliveira, E., Pires, D.F., 2011. A multi-objective simulated annealing approach to reactive power compensation. *Eng. Optim.* 43, 1063–1077.
- Bandyopadhyay, S., Saha, S., Maulik, U., Deb, K., 2008. A Simulated Annealing-Based Multiobjective Optimization Algorithm: AMOSA. *IEEE Trans. Evol. Comput.* 12, 269–283.
- Berry, J.W., Fleischer, L., Hart, W.E., Phillips, C.A., Watson, J.-P., 2005. Sensor Placement in Municipal Water Networks. *J. Water Resour. Plan. Manag.* 131, 237–243.
- Blum, C., Roli, A., 2003. Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison. *ACM Comput Surv* 35, 268–308.
- Conselho Europeu, 1998. Diretiva 98/83/CE do Conselho de 3 de Novembro relativa à qualidade da água destinada ao consumo humano, *Jornal Oficial das Comunidades Europeias*.
- Czyżak, P., Jaszkiwicz, A., 1997. Pareto Simulated Annealing, in: Fandel, P.D.G., Gal, P.D.D.T. (Eds.), *Multiple Criteria Decision Making, Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, pp. 297–307.
- Dai, C.H., Zhu, Y.F., Chen, W.R., 2006. Adaptive Probabilities of Crossover and Mutation in Genetic Algorithms Based on Cloud Model, in: *IEEE Information Theory Workshop, 2006. ITW '06 Chengdu*, pp. 710–713.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6, 182–197.
- Dorini, G., Jonkergouw, P., Kapelan, Z., Savic, D., 2010. SLOTS: Effective Algorithm for Sensor Placement in Water Distribution Systems. *J. Water Resour. Plan. Manag.* 136, 620–628.
- Eliades, D., 2009. EPANET MATLAB Toolkit.  
<http://www.mathworks.com/matlabcentral/fileexchange/25100-epanet-matlab-toolkit>
- Eliades, D.G., Kyriakou, M., Polycarpou, M.M., 2014. Sensor Placement in Water Distribution Systems Using the S-PLACE Toolkit. *Proceedings of the 12th International Conference on Computing and Control for the Water Industry, CCWI2013* 70, 602–611.
- Eliades, D., Polycarpou, M., 2006. Iterative Deepening of Pareto Solutions in Water Sensor Networks, in: *Water Distribution Systems Analysis Symposium*. American Society of Civil Engineers, pp. 1–19.

- Fonseca, C.M., Fleming, P.J., 1993. Genetic Algorithms for Multiobjective Optimization: Formulation Discussion and Generalization, in: *Proceedings of the 5th International Conference on Genetic Algorithms*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 416–423.
- Gaspar Cunha, A., Takahashi, R., Antunes, C.H., 2012. *Manual de computação evolutiva e metaheurística*. Imprensa da Universidade de Coimbra e Editora da Universidade Federal de Minas Gerais, Coimbra.
- Gibbs, M.S., Maier, H.R., Dandy, G.C., 2010. Comparison of Genetic Algorithm Parameter Setting Methods for Chlorine Injection Optimization. *J. Water Resour. Plan. Manag.* 136, 288–291.
- Glover, F., Kochenberger, G.A., 2003. *Handbook of metaheuristics*. Kluwer Academic Publishers, Boston.
- Halhal, D., G.A. Walters, 1997. Water network rehabilitation with structured messy genetic algorithm. *J. Water Resour. Plan. Manag.* 123, 137–147.
- Hart, W.E., Murray, R., 2010. Review of sensor placement strategies for contamination warning systems in drinking water distribution systems. *J. Water Resour. Plan. Manag.* 136, 611–619.
- Hilaco, S.I. da C., 2012. *Implementação do plano de segurança da água para consumo humano em Portugal* (Tese de Mestrado). Universidade Nova de Lisboa.
- Horn, J., Nafpliotis, N., Goldberg, D.E., 1994. A niched Pareto genetic algorithm for multiobjective optimization, in: *Proceedings of the First IEEE Conference on Evolutionary Computation*, 1994, pp. 82–87 vol.1.
- Huang, J.J., McBean, E.A., 2008. Using Bayesian Statistics to Estimate Chlorine Wall Decay Coefficients for Water Supply System. *J. Water Resour. Plan. Manag.* 134, 129–137.
- Huang, J., McBean, E., James, W., 2006. Multi-objective Optimization for Monitoring Sensor Placement in Water Distribution Systems, in: *Water Distribution Systems Analysis Symposium*. American Society of Civil Engineers, pp. 1–14.
- Kang, D., Lansey, K., 2012. Revisiting Optimal Water-Distribution System Design: Issues and a Heuristic Hierarchical Approach. *J. Water Resour. Plan. Manag.* Vol. 138 Issue 3, p208–217.
- Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by Simulated Annealing. *Science* 220, 671–680.
- Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., VanBriesen, J., Small, M., Fischbeck, P., 2006. Optimizing Sensor Placements in *Water Distribution Systems Using Submodular Function Maximization*. American Society of Civil Engineers, pp. 1–17.
- Lee, B.H., Deininger, R.A., 1992. Optimal Locations of Monitoring Stations in Water Distribution System. *J. Environ. Eng.* 118, 4–16.
- Loureiro, D., Coelho, S.T., 2002. *EPANET 2.0 em português (Manual de Utilização)*. Laboratório Nacional de Engenharia Civil, Lisboa.
- Marques, J.A.S., Sousa, J.J. de O., 2008. *Hidráulica Urbana. Sistemas de Abastecimento de Água e de Drenagem de Águas Residuais*, 1<sup>a</sup> ed. Imprensa da Universidade de Coimbra, Coimbra.
- Mohammed, N.G.N., Abdulrahman, A., 2009. Water Supply Network System Control based on Model Predictive Control, in: *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Newswood Ltd.: International Association of Engineers, Hong Kong.

- Munavalli, G.R., Kumar, M.S.M., 2003. Optimal Scheduling of Multiple Chlorine Sources in Water Distribution Systems. *J. Water Resour. Plan. Manag.* 129, 493–504.
- Murray, R., Hart, W.E., Phillips, C.A., Berry, J., Boman, E.G., Carr, R.D., Riesen, L.A., Watson, J.-P., Haxton, T., Herrmann, J.G., Janke, R., Gray, G., Taxon, T., Uber, J.G., Morley, K.M., 2009. US Environmental Protection Agency Uses Operations Research to Reduce Contamination Risks in Drinking Water. *Interfaces* 39, 57–68.
- Murray, R., Haxton, T., Sean A. McKenna, David B. Hart, Katherine Klise, Mark Koch, Eric D. Vugrin, Shawn Martin, Mark Wilson, Victoria Cruz, Laura Cutler, 2010. *Water Quality Event Detection Systems for Drinking Water Contamination Warning Systems: Development, Testing, and Application of CANARY*.
- Mutikanga, H.E., Sharma, S.K., Vairavamoorthy, K., 2013. Methods and Tools for Managing Losses in Water Distribution Systems. *J. Water Resour. Plan. Manag.* 139, 166–174.
- Ostfeld, A., Uber, J.G., Salomons, E., 2005. Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms - Detailed Problem Description and Rules. *Presented at the Water Distribution Systems Analysis 2006*.
- Ostfeld, A., Uber, J.G., Salomons, E., Berry, J.W., Hart, W.E., Phillips, C.A., Watson, J.-P., Dorini, G., Jonkergouw, P., Kapelan, Z., di Pierro, F., Khu, S.-T., Savic, D., Eliades, D., Polycarpou, M., Ghimire, S.R., Barkdoll, B.D., Gueli, R., Huang, J.J., McBean, E.A., 2008. The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms. *J. Water Resour. Plan. Manag.* 134, 556–568.
- Peng, W., Zhang, Q., Li, H., 2009. Comparison between MOEA/D and NSGA-II on the Multi-Objective Travelling Salesman Problem, in: Goh, C.-K., Ong, Y.-S., Tan, K.C. (Eds.), *Multi-Objective Memetic Algorithms, Studies in Computational Intelligence*. Springer Berlin Heidelberg, pp. 309–324.
- Pereira, F. J. B., 2012. Algoritmos Genéticos, in: *Manual de Computação Evolutiva E Metaheurística*. Imprensa da Universidade de Coimbra, pp. 25–48.
- Philadelphia Water Department, CH2M HILL, 2013. *Selection of Online Water Quality Monitoring Technologies and Station Design* (White Paper Submitted to EPA). Philadelphia Water Department.
- Pirlot, M., 1996. General local search methods. *Eur. J. Oper. Res.* 92, 493–511.
- Rosen, J.S., Bartrand, T., 2013. Security and Preparedness -- Using Online Water Quality Data to Detect Events in a Distribution System. *J. - Am. Water Works Assoc.* 105, 22–26.
- Rossman, L., 2000. EPANET 2 Users Manual.
- Savic, D.A., Walters, G.A., 1995. *Genetic algorithm techniques for calibrating network models* (No. 95/12). Centre for Systems and Control Engineering, University of Exeter, Devon, United Kingdom.
- Savic, D.A., Walters, G.A., 1997. Genetic algorithms for least-cost design of water distribution networks. *J. Water Resour. Plan. Manag.* 123, 67–78.
- Silberholz, J., Golden, B., 2010. Comparison of Metaheuristics, in: Gendreau, M., Potvin, J.-Y. (Eds.), *Handbook of Metaheuristics*, International Series in Operations Research & Management Science. Springer US, pp. 625–640.
- Srinivas, M., Patnaik, L.M., 1994. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Trans. Syst. Man Cybern.* 24, 656–667.
- Srinivas, N., Deb, K., 1994. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evol Comput* 2, 221–248.

- Suman, B., 2002. Multiobjective simulated annealing - a metaheuristic technique for multiobjective optimization of a constrained problem. *Found. Comput. Decis. Sci.* Vol. 27, No. 3, 171–191.
- Suman, B., 2003. Simulated annealing-based multiobjective algorithms and their application for system reliability. *Eng. Optim.* 35, 391–416.
- Suman, B., Kumar, P., 2005. A survey of simulated annealing as a tool for single and multiobjective optimization. *J. Oper. Res. Soc.* 57, 1143–1160.
- Suppavitnarm, A., Parks, G.T., 1999. Simulated annealing: An alternative approach to true multiobjective optimization, in: *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 1999)*. Morgan Kaufmann Publishers, Orlando, Florida, pp. 406–407.
- Takahashi, R.H.C., Cunha, A.G., 2012. Introdução, in: *Manual de Computação Evolutiva e Metaheurística*. Imprensa da Universidade de Coimbra, pp. 1–21.
- Takahashi, R.H.C., Gaspar-Cunha, A., Fonseca, C.M., 2012. Algoritmos evolutivos multi-objectivo, in: *Manual de Computação Evolutiva e Metaheurística*. Imprensa da Universidade de Coimbra, pp. 357–380.
- Ulungu, E.L., Teghem, J., Ost, C., 1998. Efficiency of interactive multi-objective simulated annealing through a case study. *J. Oper. Res. Soc.* 49, 1044–1050.
- US Army, 2009. *Drinking water contamination warning systems - Design and implementation of an online water quality monitoring system for military water system*. Technical Information Paper. US Army.
- US EPA, 2005. *Water sentinel system architecture*. US EPA Office of Research and Development, Washington, DC.
- Vasconcelos, J., Ramirez, J., Takahashi, R.H.C., Saldanha, R.R., 2001. Improvements in genetic algorithms. *IEEE Trans. Magn.* 37, 3414–3417.
- Vieira, J.M.P., Morais, C., 2005. *Planos de Segurança da água para consumo humano em sistemas públicos de abastecimento*. Instituto Regulador de Águas e Resíduos e Universidade do Minho.
- Watson, J., Murray, R., Hart, W., 2009. Formulation and Optimization of Robust Sensor Placement Problems for Drinking Water Contamination Warning Systems. *J. Infrastruct. Syst.* 15, 330–339.
- WHO, 2008. *Guidelines for drinking-water quality Third edition incorporating the first and second addenda* (No. Volume 1 - Recommendations). World Health Organization, Geneva.
- Wu, Z.Y., Sage, P., Turtle, D., 2010. Pressure-Dependent Leak Detection Model and Its Application to a District Water System. *J. Water Resour. Plan. Manag.* 136, 116–128.
- Wu, Z.Y., Walski, T., 2006. Multi-Objective Optimization of Sensor Placement in Water Distribution Systems, in: *Water Distribution Systems Analysis Symposium*. American Society of Civil Engineers, pp. 1–11.
- Xu, J., Fischbeck, P.S., Small, M.J., VanBriesen, J.M., Casman, E., 2008. Identifying Sets of Key Nodes for Placing Sensors in Dynamic Water Distribution Networks. *J. Water Resour. Plan. Manag.* 134, 378–385.
- Zitzler, E., Laumanns, M., Thiele, L., 2001. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. In Giannakoglou *et al.*, *Evolutionary methods for design, optimization and control*, CIMNE, Barcelona.
- Zou, R., Lung, W.-S., 2004. Robust Water Quality Model Calibration Using an Alternating Fitness Genetic Algorithm. *J. Water Resour. Plan. Manag.* 130, 471–479.





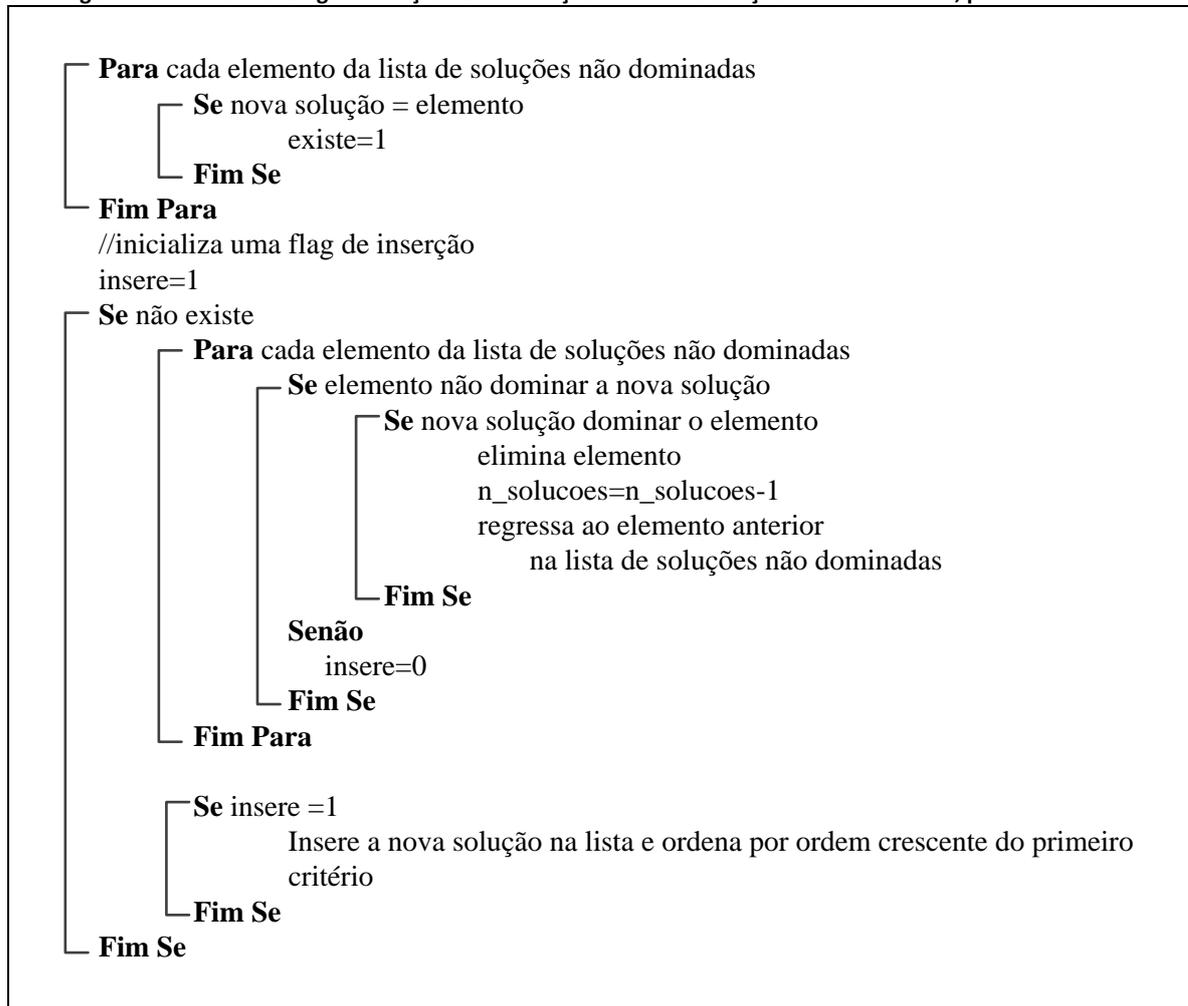
## **APÊNDICE A - PSEUDOCÓDIGOS**



## Lista de Pseudocódigos não inseridos no texto

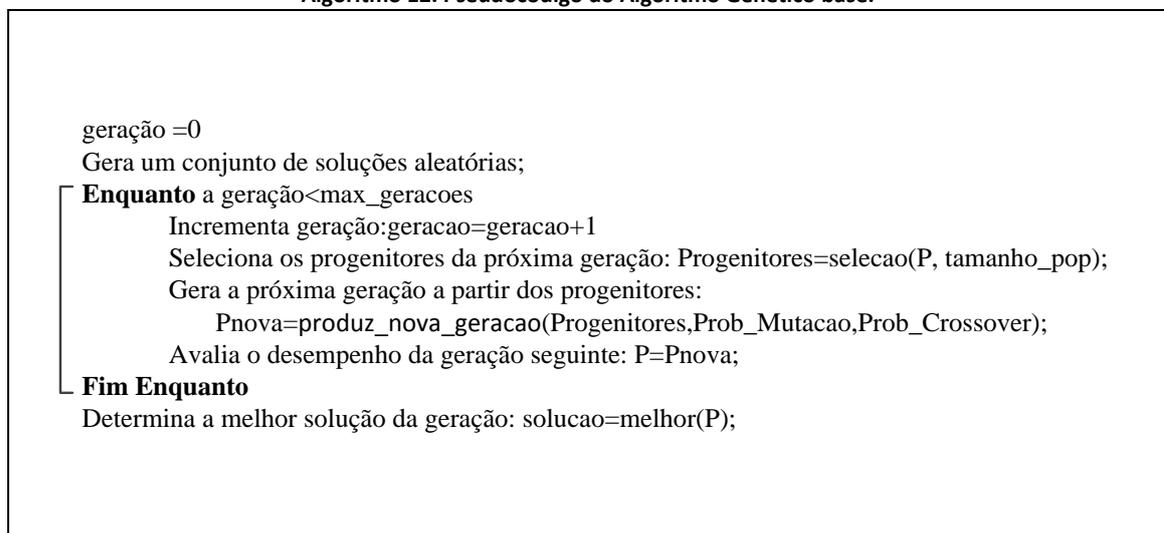
### *Simulated annealing*

**Algoritmo 11.** Pseudocódigo da função de atualização da lista de soluções não dominadas, para a SMOSA.



### **Algoritmo Genético**

**Algoritmo 12.** Pseudocódigo do Algoritmo Genético base.



**Algoritmo 13. Pseudocódigo para a seleção por torneio binário.**

```

Para i=1 até número de pais
  Inicializa os índices dos indivíduos a escolher, com um vetor de zeros de dimensão igual ao tamanho do torneio
  Inicializa o contador de elementos do torneio: i=1
  Enquanto não forem escolhidos os 2 elementos do torneio: i<=2
    Determina aleatoriamente o índice de um dos elementos da pop.: temp=rand[1,tamanho_pop]
    Assume inicialmente que o novo elemento não foi ainda escolhido: ja_existe=0
    //Verifica se não foi já escolhido
    Para j=1:i-1
      Se o índice(j) for igual a temp, já_existe=1
    Fim Para
    Se não foi ainda escolhido ( já_existe=0),
      Insere esse índice na lista de índices: indice(i)=temp
      Incrementa o contador de elementos do torneio: i=i+1
    Fim Se
  Fim enquanto

  //Compara os elementos seleccionados e escolhe o melhor
  Se fitness do elemento com o índice 1 for melhor que o elemento com o índice 2
    O escolhido é o elemento com o índice 1
  Senão
    O escolhido é o elemento com o índice 2
  Fim Se
Fim para
  
```

**Algoritmo 14. Pseudocódigo para produzir nova geração.**

```

Dado o tamanho da população: tamanho_pop
Para cada par de progenitores: i=1 até tamanho_pop, de dois em dois
  Obtém o descendente Pnova(i) e o descendente Pnova(i+1), com a função recombinação,
  para uma probabilidade de crossover especificada
  Se rand[0,1]<Prob_Mutacao
    Pnova(i)=mutacao(Pnova(i));
  Fim Se
  Se rand[0,1]<Prob_Mutacao
    Pnova(i+1)=mutacao(Pnova(i+1));
  Fim Se
Fim Para
  
```

## **APÊNDICE B - CORRESPONDÊNCIA ENTRE A POSIÇÃO DO NÓ COM A ETIQUETA DO NÓ NO EPANET**



Quadro D 1. Rede

Posição do nó	Etiqueta do nó	Posição do nó	Etiqueta do nó	Posição do nó	Etiqueta do nó
1	JUNCTION-0	43	JUNCTION-42	85	JUNCTION-84
2	JUNCTION-1	44	JUNCTION-43	86	JUNCTION-85
3	JUNCTION-2	45	JUNCTION-44	87	JUNCTION-86
4	JUNCTION-3	46	JUNCTION-45	88	JUNCTION-87
5	JUNCTION-4	47	JUNCTION-46	89	JUNCTION-88
6	JUNCTION-5	48	JUNCTION-47	90	JUNCTION-89
7	JUNCTION-6	49	JUNCTION-48	91	JUNCTION-90
8	JUNCTION-7	50	JUNCTION-49	92	JUNCTION-91
9	JUNCTION-8	51	JUNCTION-50	93	JUNCTION-92
10	JUNCTION-9	52	JUNCTION-51	94	JUNCTION-93
11	JUNCTION-10	53	JUNCTION-52	95	JUNCTION-94
12	JUNCTION-11	54	JUNCTION-53	96	JUNCTION-95
13	JUNCTION-12	55	JUNCTION-54	97	JUNCTION-96
14	JUNCTION-13	56	JUNCTION-55	98	JUNCTION-97
15	JUNCTION-14	57	JUNCTION-56	99	JUNCTION-98
16	JUNCTION-15	58	JUNCTION-57	100	JUNCTION-99
17	JUNCTION-16	59	JUNCTION-58	101	JUNCTION-100
18	JUNCTION-17	60	JUNCTION-59	102	JUNCTION-101
19	JUNCTION-18	61	JUNCTION-60	103	JUNCTION-102
20	JUNCTION-19	62	JUNCTION-61	104	JUNCTION-103
21	JUNCTION-20	63	JUNCTION-62	105	JUNCTION-104
22	JUNCTION-21	64	JUNCTION-63	106	JUNCTION-105
23	JUNCTION-22	65	JUNCTION-64	107	JUNCTION-106
24	JUNCTION-23	66	JUNCTION-65	108	JUNCTION-109
25	JUNCTION-24	67	JUNCTION-66	109	JUNCTION-110
26	JUNCTION-25	68	JUNCTION-67	110	JUNCTION-111
27	JUNCTION-26	69	JUNCTION-68	111	JUNCTION-112
28	JUNCTION-27	70	JUNCTION-69	112	JUNCTION-113
29	JUNCTION-28	71	JUNCTION-70	113	JUNCTION-114
30	JUNCTION-29	72	JUNCTION-71	114	JUNCTION-115
31	JUNCTION-30	73	JUNCTION-72	115	JUNCTION-116
32	JUNCTION-31	74	JUNCTION-73	116	JUNCTION-117
33	JUNCTION-32	75	JUNCTION-74	117	JUNCTION-118
34	JUNCTION-33	76	JUNCTION-75	118	JUNCTION-119
35	JUNCTION-34	77	JUNCTION-76	119	JUNCTION-120
36	JUNCTION-35	78	JUNCTION-77	120	JUNCTION-121
37	JUNCTION-36	79	JUNCTION-78	121	JUNCTION-122
38	JUNCTION-37	80	JUNCTION-79	122	JUNCTION-123
39	JUNCTION-38	81	JUNCTION-80	123	JUNCTION-124
40	JUNCTION-39	82	JUNCTION-81	124	JUNCTION-125
41	JUNCTION-40	83	JUNCTION-82	125	JUNCTION-126
42	JUNCTION-41	84	JUNCTION-83	126	JUNCTION-128

Posição do nó	Etiqueta do nó
127	RESERVOIR-129
128	TANK-130
129	TANK-131

**Quadro D 2. Rede 2**

Posição do nó	Etiqueta do nó
1	JUNCTION-0
2	JUNCTION-1
3	JUNCTION-2
4	JUNCTION-3
5	JUNCTION-4
6	JUNCTION-5
7	JUNCTION-6
8	JUNCTION-7
⋮	⋮
12518	JUNCTION-12517
12519	JUNCTION-12518
12520	JUNCTION-12519
12521	JUNCTION-12520
12522	JUNCTION-12521
12523	JUNCTION-12522
12524	RESERVOIR-12523
12525	RESERVOIR-12524
12526	TANK-12525
12527	TANK-12526

## **APÊNDICE C - RESULTADOS DOS PROCEDIMENTOS DE OTIMIZAÇÃO**



Quadro C 1. Estatísticas referentes à otimização mono-objetivo de  $Z_1$  com SA, R1A5.

Melhor conjunto de parâmetros	Solução	$T_0$	$\alpha$	n_passos	Estrutura de Vizinhança	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$	iteracao	T execução [s]	Solução				
	Melhor	1000	0.01	200	Ambas	151.71	108	2422.855	20.33%	1152	1633	32	35	38	41	50
	Pior					167.68	85	1588.906	20.95%			22	25	31	36	41
	Média					153.90										
	Mediana					151.71	108	2422.855	20.33%			32	38	35	41	50
	Desvio padrão					5.18										
Análise global	Pior	500	0.05	100	1 sensor	203.27	110	1670.064	32.09%	263	194	10	22	32	36	121
	Média					159.55										
	Desvio padrão					12.26										

Quadro C 2. Estatísticas referentes à otimização mono-objetivo de  $Z_2$  com SA, R1A5.

Melhor conjunto de parâmetros	Solução	$T_0$	$\alpha$	n_passos	Estrutura de Vizinhança	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$	iteracao	T execução [s]	Solução				
	Melhor	500	0.01	50	1 sensor	196.571	59	1060.83	22.82%	1082	392	22	27	31	38	55
	Pior					320.722	84	1748.93	28.86%			21	27	32	96	97
	Média						62									
	Mediana					196.571	59	1060.83	22.82%			22	27	31	38	55
	Desvio padrão						7.97									
Análise global	Pior	500	0.05	50	1 sensor	346.918	93	1816.49	33.51%	263	98	21	32	65	92	108
	Média						62.40									
	Desvio padrão						5.76									

Quadro C 3. Estatísticas referentes à otimização mono-objetivo de  $Z_3$  com SA, R1A5.

Melhor conjunto de parâmetros	Solução	$T_0$	$\alpha$	n_passos	Estrutura de Vizinhança	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$	iteracao	T execução [s]	Solução				
	Melhor	1000	0.05	200	1 sensor	192.65	67	619.114	24.86%	277	393	21	25	35	32	38
	Pior					192.65	67	619.114	24.86%			21	25	35	32	38
	Média							619.114								
	Mediana					192.65	67	619.114	24.86%			21	25	35	32	38
	Desvio padrão							0.000								
Análise global	Pior	100	0.05	200	2 sensores	565.22	273	14044	25.38%	231	360.92	44	69	113	124	127
	Média							1253.12								
	Desvio padrão							1953.74								

Quadro C 4. Estatísticas referentes à otimização mono-objetivo de  $Z_4$  com SA, R1A5.

Melhor conjunto de parâmetros	Solução	$T_0$	$\alpha$	n_passos	Estrutura de Vizinhança	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$	iteracao	T execução [s]	Solução				
	Melhor	1	0.01	100	1 sensor	1256.88	670	43041.836	83.92%	691	522	11	46	84	101	125
	Pior					1055.08	688	41118.828	81.77%			46	84	101	122	125
	Média								83.20%							
	Mediana					1256.88	670	43041.836	83.92%			11	46	84	101	125
	Desvio padrão								1.01%							
Análise global	Pior	5	0.05	100	1 sensor	993.48	609	27211.06	73.77%	125	92	20	84	96	101	125
	Média								41.64%							
	Desvio padrão								29.73%							

Quadro C 5. Estatísticas referentes à otimização mono-objetivo de  $Z_1$  com AG com elitismo, R1A5.

Melhor conjunto de parâmetros	Solução	População	Prob. Mutação	Prob. Crossover	Máx. Gerações	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$	Elitismo	T execução [s]	Solução				
	Melhor	50	0.2	0.7	150	151.71	108	2422.855	20.33%	5%	45	32	35	38	41	50
	Pior					245.43	207	2227.145	40.62%	5%		22	32	36	117	121
	Média					187.75										
	Mediana					182.84	74	821.438	28.59%	5%		10	22	32	35	36
	Desvio padrão					28.56										
Análise global	Pior	50	0.2	0.7	150	317.86	108	2422.855	20.33%	5%	45	32	35	38	41	50
Média						189.97										
Desvio padrão						24.88										

Quadro C 6. Estatísticas referentes à otimização mono-objetivo de  $Z_2$  com AG com elitismo, R1A5.

Melhor conjunto de parâmetros	Solução	População	Prob. Mutação	Prob. Recombinação	Máx. Gerações	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$	Elitismo	T execução [s]	Solução				
	Melhor	200	0.2	0.7	150	196.57	59	1060.83	22.82%	0.05	178	22	27	31	38	55
	Pior					295.17	79	1589.28	30.41%	0.05		21	25	32	96	98
	Média						67.82									
	Mediana					175.49	70	1265.46	23.46%	0.05		22	27	32	38	41
	Desvio padrão						7.00									
Análise global	Pior	50	0.15	0.7	150	823.8034	144	3697.315	55.99%	5%	46	10	22	29	32	82
Média							73.75									
Desvio padrão							12.61									

Quadro C 7. Estatísticas referentes à otimização mono-objetivo de  $Z_3$  com AG com elitismo, R1A5.

Melhor conjunto de parâmetros	Solução	População	Prob. Mutação	Prob. Recombinação	Máx. Gerações	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$	Elitismo	T execução [s]	Solução				
	Melhor	200	0.2	0.7	150	192.65	67	619.11	24.86%	5%	177	21	25	32	35	38
	Pior					279.44	97	1355.98	41.26%	5%		19	23	38	69	91
	Média							967.64								
	Mediana					204.33	68	1137.66	28.08%	5%		21	25	32	38	92
	Desvio padrão							341.08								
Análise global	Pior	50	0.15	0.7	150	390.1922	125	2303	48.24%	5%	45	18	32	86	91	102
	Média							1167.79								
	Desvio padrão							350.74								

Quadro C 8. Estatísticas referentes à otimização mono-objetivo de  $Z_4$  com AG com elitismo, R1A5.

Melhor conjunto de parâmetros	Solução	População	Prob. Mutação	Prob. Recombinação	Máx. Gerações	$Z_1$ [min]	$Z_2$ [pessoas]	$Z_3$ [gal]	$Z_4$	Elitismo	T execução [s]	Solução				
	Melhor	100	0.2	0.7	150	1256.88	670.00	43041.84	83.92%	0.05	94	11	46	84	101	125
	Pior					1277.41	778.00	60646.52	73.27%	0.05		12	20	74	84	122
	Média								76.91%							
	Mediana					1393.03	748.00	53239.11	75.92%	0.05		11	20	84	113	125
	Desvio padrão								3.60%							
Análise global	Pior	50	0.15	0.7	150	1187.769	595	48659.24	68.35%	5%	49	12	20	59	75	84
	Média								77.45%							
	Desvio padrão								3.17%							

**Quadro C 9. Soluções para cada função objetivo no caso R1A20, na otimização mono-objetivo com AG**

Função otimizada	Valor da solução				Sensores						
	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [galões]	Z <sub>4</sub> [%]							
Z <sub>1</sub>	<b>64.23</b>	39	250.12	35.57%	6	10	16	22	27	32	33
					34	35	36	41	42	44	50
					52	57	59	60	112	115	
Z <sub>2</sub>	116.41	<b>24</b>	171.95	42.04%	6	7	16	19	22	25	29
					31	32	38	42	43	52	57
					59	65	66	92	109	115	
Z <sub>3</sub>	118.89	31	<b>100.47</b>	41.77%	21	27	30	31	35	36	40
					43	44	50	52	55	59	61
					66	69	96	98	99	102	
Z <sub>4</sub>	554.05	207	3270.23	<b>89.99%</b>	5	11	12	15	20	21	28
					35	36	40	46	59	84	85
					101	107	113	122	123	125	

**Quadro C 10. Soluções com a menor distância à solução ideal obtidas, no caso R1A20, na otimização triobjetivo com NSGA-II**

FO otimizadas	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclideana	Distância Chebyshev	Solução						
Z <sub>1</sub> -Z <sub>2</sub> -Z <sub>4</sub>	206.99	58	524.01	70.81%	0.60		6	15	20	22	27	32	36
							63	66	71	75	77	80	85
							86	91	99	103	109	121	
	229.97	76	663.89	73.75%		0.38	13	15	18	20	22	28	32
							35	36	47	69	73	86	95
							102	113	115	117	119	121	

**Quadro C 11. Soluções para a R2A20, com a menor distância à solução ideal na otimização triobjetivo, com NSGA-II**

Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclideana	Distância Chebyshev	Solução						
457.25	632	21901.38	15.58%	1.05		1729	2641	3144	3166	3218	3315	4238
						5492	5725	5781	5932	6584	6918	6996
						7364	7686	9197	10491	10608	11379	
923.05	1552	96511.73	34.19%		0.75	3	116	898	1925	2036	2096	2128
						2653	3329	3663	6167	6588	7924	8749
						9183	9548	9964	10875	11583	11881	

**Quadro C 12. Soluções para a R2A20, com a menor distância à solução ideal na otimização triobjetivo, com NSGA-II, desprezando nós com menos de 500 detecções.**

Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Distância Euclideana	Distância Chebyshev	Solução						
<b>526.83</b>	<b>910</b>	37617.65	<b>24.89%</b>	1.235		69	154	1023	1330	3840	4137	6571
						6985	7547	7789	8260	9143	9772	10085
						10202	10332	10351	10738	11265	12457	
<b>868.50</b>	<b>1254</b>	77666.49	<b>35.10%</b>		0.838	124	2302	2869	3236	3554	3660	4243
						4686	4920	6736	7378	7544	7627	7834
						9469	10520	10716	10755	11977	12165	

**Quadro C 13. Soluções para cada função objetivo para o caso R2A20, na otimização mono-objetivo, desprezando nós com menos de 500 detecções, com AG.**

Função otimizada	Valor da solução				Sensores						
	Z <sub>1</sub> [min]	Z <sub>2</sub> [pessoas]	Z <sub>3</sub> [galões]	Z <sub>4</sub> [%]							
<b>Z<sub>1</sub></b>	<b>407.01</b>	803	41463	18.38%	519	1189	1900	1968	3319	3429	5115
					5254	6650	7756	8248	8377	9000	10359
					10453	10615	11035	11193	11246	11305	
<b>Z<sub>2</sub></b>	454.00	<b>580</b>	23007.17	20.45%	416	851	1002	1037	2074	4033	4245
					4570	5039	5046	5112	7223	7381	7496
					7992	8273	8553	9000	9143	11192	
<b>Z<sub>3</sub></b>	437.87	582	<b>19779.01</b>	21.03%	1469	2074	2293	3333	3434	3631	4033
					4154	4553	4854	6630	6650	7256	7992
					8377	9143	10064	10127	10197	10408	
<b>Z<sub>4</sub></b>	1329.02	2266	199845.44	<b>40.13%</b>	511	572	1479	1905	2503	3231	3487
					3765	3790	3834	3860	4108	4613	4956
					7442	7666	8328	9051	9626	12277	