



Pós-Graduação em Ciência da Computação

Elaine Guerrero Peña

Algoritmo Evolucionário Multi-Objetivo Dinâmico com Estratégia de Previsão no Espaço Objetivo



Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

Recife
2020

Elaine Guerrero Peña

**Algoritmo Evolucionário Multi-Objetivo Dinâmico com Estratégia de
Previsão no Espaço Objetivo**

Trabalho apresentado ao Programa de Pós-graduação em
Ciência da Computação do Centro de Informática da Univer-
sidade Federal de Pernambuco como requisito parcial para
obtenção do grau de Doutor em Ciência da Computação.

Área de Concentração: Inteligência Computational

Orientador: Professor Dr. Aluizio Fausto Ribeiro Araújo

Recife
2020

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

G934a Guerrero Peña, Elaine
 Algoritmo evolucionário multi-objetivo dinâmico com estratégia de previsão
 no espaço objetivo / Elaine Guerrero Peña. – 2020.
 125 f.: il., fig., tab.

 Orientador: Aluizio Fausto Ribeiro Araújo.
 Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da
 Computação, Recife, 2020.
 Inclui referências.

 1. Inteligência computacional. 2. Otimização. I. Araújo, Aluizio Fausto
 Ribeiro (orientador). II. Título.

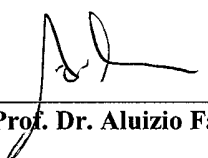
 006.31 CDD (23. ed.) UFPE - CCEN 2020 - 66

Elaine Guerrero Peña

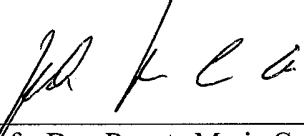
**“Algoritmo Evolucionário Multi-Objetivo Dinâmico com Estratégia de
Previsão no Espaço Objetivo”**

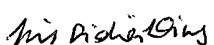
Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutora em Ciência da Computação.

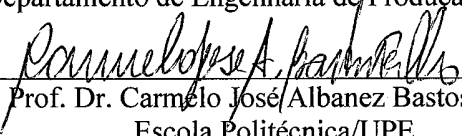
Aprovado em: 05/03/2020.


Orientador: Prof. Dr. Aluizio Fausto Ribeiro Araújo

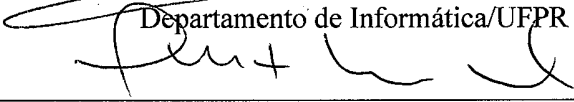
BANCA EXAMINADORA


Prof. Dra. Renata Maria Cardoso Rodrigues de Souza
Centro de Informática


Prof. Dra. Isis Didier Lins
Departamento de Engenharia de Produção/ UFPE


Prof. Dr. Carmelo José Albanez Bastos Filho
Escola Politécnica/UPE


Prof. Dra. Aurora Trinidad Ramirez Pozo
Departamento de Informática/UFPR


Prof. Dr. Luiz Satoru Ochi
Departamento de Ciência da Computação/UFF

A minha família.

AGRADECIMENTOS

Gostaria de agradecer de maneira especial e sincera ao meu orientador Aluizio Araújo por me aceitar sob sua direção no doutorado. A sua orientação, apoio e confiança no meu trabalho têm sido uma contribuição inestimável, não apenas no desenvolvimento desta tese, mas também na minha formação como pesquisadora. Agradeço também por sempre me fornecer meios suficientes para realizar todas as atividades propostas durante o desenvolvimento desta tese.

Aos meus amigos do laboratório que compartilharam conhecimentos e experiências de natureza profissional e pessoal de grande valor, por me ajudar e me corrigir no português, por ser esses amigos com quem sempre posso contar. Aos meus amigos cubanos da vida toda, que embora longe, sempre estiveram aí para me apoiar e dar uma força quando mais precisava. Aos amigos sinceros que fiz aqui no Brasil, muito obrigada a todos!

A todos os meus professores do Centro de Informática, também agradeço sinceramente, pelas aulas, as grandes contribuições em experiências e conhecimentos. Minha gratidão também para ao pessoal da secretaria de pós-graduação, da portaria e demais funcionários do Centro de Informática. Agradeço a todos pela ajuda oferecida nesta viagem.

E, é claro, os agradecimentos mais profundos e sentidos para minha família: meus pais Fidel e Elina, ao meu irmão Fillo e minha irmã Sara. Sem o seu apoio, conselhos, colaboração e inspiração, seria impossível realizar essa empresa difícil. A minha família em Cuba, que tanto apoio brindo-me na distância. Finalmente, e não menos importante, quero agradecer a Fábio e a família dele que com tanto carinho me acolheu. Muito obrigada pelo apoio e a ajuda. Saibam que são minha família brasileira.

Agradeço a todos que me apoiaram e confiaram em mim!

RESUMO

Os Algoritmos Evolucionários (EA) têm sido amplamente explorados e aplicados em problemas de otimização. Nos ambientes dinâmicos, a solução ótima se modifica e o algoritmo de otimização deve encontrá-la ao longo do tempo. Nestas situações, se o tempo disponível entre as alterações dos objetivos for relativamente curto, não haverá tempo hábil para a convergência do algoritmo entre as mudanças. Estudos indicam que um algoritmo bem-sucedido deve produzir algum tipo de diversidade para poder escapar de um conjunto de soluções e convergir rapidamente para um novo conjunto, à medida que as mudanças ocorram no ambiente. Diferentes algoritmos foram propostos modificando a maneira como uma mudança é manipulada. Entre eles, os métodos baseados em previsão têm se mostrado promissores em acelerar a convergência dos EA dinâmicos. Eles fornecem orientação para a evolução da população por meio de um mecanismo de previsão que ajuda o EA a responder rapidamente à nova mudança. Com base nessas estratégias, propomos um algoritmo dinâmico baseado na Evolução Diferencial (DE) com Previsão no Espaço Objetivo (*Dynamic Non-dominated Sorting Differential Evolution improvement with Prediction in the Objective Space*: DOSP-NSDE). DOSP-NSDE é um algoritmo baseado em Pareto e em aprendizagem. A aprendizagem é introduzida no algoritmo a través de uma estratégia de Previsão de Espaço do Objetivo (OSP). O OSP identifica tendências nos movimentos de indivíduos não dominados no espaço objetivo, e então rapidamente encontra regiões promissoras no espaço de busca e gera uma nova população considerando essas regiões. DOSP-NSDE aciona a estratégia OSP quando várias premissas baseadas na métrica Hipervolume Aproximado são alcançadas. Os operadores de variação do DE e a mutação polinomial são usados para produzir a descendência sempre que a condição de OSP não é satisfeita. A proposta usa a estratégia de previsão tanto no processo evolucionário estático (entre mudanças) como no mecanismo de reação à mudança para prever a nova localização da frente ótima. Primeiramente, a eficácia do OSP foi verificada no processo estático. Para isto, foram realizados um conjunto de experimento em 31 funções de múltiplos objetivos bem conhecidas, comparando com algoritmos baseados em EA e outros da inteligência coletiva. Comprovada a eficácia e a eficiência da proposta em ambientes estáticos, foram realizados experimentos em 19 funções de teste multiobjetivas dinâmicas e em um problema do mundo real. As comparações do DOSP-NSDE com diversos algoritmos da literatura considerando três métricas evidenciaram que a proposta é competitiva na maioria dos problemas.

Palavras-chave: Otimização Multiobjetiva Evolucionária Estática. Otimização Multiobjetiva Evolucionária Dinâmica. Algoritmos Evolucionários baseados em Pareto. Previsão. Modelagem Probabilística. Inferência Variacional.

ABSTRACT

Evolutionary Algorithms (EA) have been extensively explored and applied in optimization problems. In dynamic environments, the optimum solution changes and the optimization algorithm must find it over time. In these situations, if the time available between objective changes is relatively short, there will be no time for algorithm convergence between changes. Therefore, a successful algorithm must produce some diversity to escape from a set of solutions and quickly converge to a new set as changes in the environment occur. Different algorithms have been proposed by modifying the way that a change is handled. Among them, Prediction-based methods have been shown promising to accelerate the convergence of an algorithm. They provide guidance direction for the population evolution through a prediction mechanism that assists the DMOEA to respond quickly to the new change. Based on these strategies, we propose a Dynamic Non-dominated Sorting Differential Evolution improvement with Prediction in the Objective Space (DOSP-NSDE). DOSP-NSDE is an algorithm based on Pareto and learning. Learning is introduced into the algorithm through an Objective Space Prediction (OSP) strategy. OSP identifies trends in the movements of non-dominated individuals in the objective space, and then quickly finds promising regions in the search space and generates a new population considering such regions. DOSP-NSDE triggers the OSP strategy when several premises based on the Approximated Hypervolume metric are achieved. Differential Evolution (DE) variation operators and Polynomial mutation are used to produce the offspring whenever the OSP condition is not satisfied. The proposal uses the OSP in both the static evolutionary process (between changes) and in the change reaction mechanism to predict the new location of the optimal front. The OSP-NSDE effectiveness was verified in 31 well-known functions, comparing with EA-based algorithms and others of collective intelligence. Once the effectiveness and efficiency of the proposed EA in static environments were verified, new experiments were performed in 19 dynamic multi-objective test functions and in a real-world problem. The comparisons of the DOSP-NSDE with several algorithms of the literature considering three metrics showed that the proposal is competitive in most problems.

Keywords: Multi-objective Evolutionary Optimization. Dynamic Multi-objective Evolutionary Optimization. Pareto-based Evolutionary Algorithms. Forecast. Probabilistic Modeling. Variational Inference.

LISTA DE FIGURAS

Figura 1 – Organização da Tese	21
Figura 2 – Nova Taxonomia dos DMOEAs	34
Figura 3 – Representação Gráfica de um Modelo de Variáveis Latentes.	47
Figura 4 – Tipos de Modelos de Variáveis Latentes.	49
Figura 5 – Métodos de Inferência para determinar os parâmetros dos modelos. . .	54
Figura 6 – Representação Gráfica do Modelo de Mistura Gaussiana.	63
Figura 7 – GMM estimado usando VI. O modelo está formado por três gaussianas. .	64
Figura 8 – Evolução Diferencial de Ordenação Não Dominada Melhorada Dinâmica com a Estratégia de Predição do Espaço Objetivo.	65
Figura 9 – Etapas da Estratégia de Predição no Espaço Objetivo. (a) Pré-processamento de dados; (b) Previsão do movimento dos indivíduos não dominados usando modelos de séries temporais; (c) Busca de soluções no espaço de decisão que possuam valores de aptidão no espaço objetivo o mais próximo possível dos da frente prevista; (d) Completar a nova população usando modelos probabilísticos.	68
Figura 10 – Evolução Diferencial com Ordenação por Não Dominância com Previsão no Espaço Objetivo.	72
Figura 11 – Ordenação por dominância de Pareto.	72
Figura 12 – Seleção da próxima geração	76
Figura 13 – Análise da sensibilidade dos parâmetros do DOSP-NSDE ($n_t = 10$, $\tau_t = 10$)	80
Figura 14 – Influência da frequência da mudança.	82
Figura 15 – Influência da severidade da mudança.	83
Figura 16 – Exemplo do funcionamento do critério para o acionamento do OSP. . .	84
Figura 17 – Etapa de previsão no espaço objetivo do OSP.	84
Figura 18 – Etapa de aproximação das soluções líderes do OSP.	85
Figura 19 – Etapa de geração da nova população usando o GMM-LS.	85
Figura 20 – Classificação média baseada no IGD para todos os algoritmos testados. A linha tracejada mostra os valores médios do OSP-NSDE.	92
Figura 21 – Classificação média baseada no HV para todos os algoritmos testados. A linha tracejada mostra os valores médios do OSP-NSDE.	92
Figura 22 – Classificação média baseada na Métrica <i>Ensemble</i> para todos os algoritmos testados. A linha tracejada mostra os valores médios do OSP-NSDE.	92
Figura 23 – Tempo médio de execução dos algoritmos para os diferentes grupos de funções de teste.	93
Figura 24 – Frente de Pareto aproximada encontrada pelos algoritmos para UF7 . .	94

Figura 25 – Frente de Pareto aproximada encontrada pelos algoritmos para ZDT3 .	94
Figura 26 – Frente de Pareto aproximada encontrada pelos algoritmos para DTLZ7	95
Figura 27 – Frente de Pareto aproximada encontrada pelos algoritmos para WFG6	95
Figura 28 – Tempo médio de execução dos algoritmos baseados em PSO e o OSP- NSDE para os diferentes grupos de funções de teste.	97
Figura 29 – Frente de Pareto aproximada obtida pelo DOSP-NSDE em todas as mudanças, nas funções: (a) FDA1; (b) FDA3; (c) DMOP2; (d) UDF5. .	100
Figura 30 – Rastreamento do controlador MFAC para os sinais de referência e seus sinais de controle, usando DOSP-NSDE para determinar os parâmetros ótimos do MFAC.	105

LISTA DE TABELAS

Tabela 1 – Tipos de DMOPs segundo (FARINA; DEB; AMATO, 2004).	25
Tabela 2 – Problemas de Otimização Multiobjetivo Dinâmicos e os estudos em que foram usados.	26
Tabela 3 – Indicadores de desempenho.	39
Tabela 4 – Condições Experimentais.	79
Tabela 5 – Média e desvio padrão da métrica IGD para as estratégias comparadas com diferentes MOEAs.	86
Tabela 6 – Média e desvio padrão do IGD para os três algoritmos baseados em NSDE. Os melhores resultados médios em cada caso estão em negrito. Os símbolos " + ", " $\approx$ " e " – " indicam que o algoritmo na coluna correspondente foi significativamente melhor que, equivalente ao, ou pior do que o OSP-NSDE.	87
Tabela 7 – Média e desvio padrão do IGD para os algoritmos comparados. Os melhores resultados médios em cada instância estão em negrito. Os símbolos " + ", " $\approx$ " e " – " indicam onde o algoritmo na coluna correspondente funciona significativamente melhor que, equivalente a, ou pior do que o OSP-NSDE.	89
Tabela 8 – Média e desvio padrão do HV para os algoritmos comparados. Os melhores resultados médios em cada instância estão em negrito. Os símbolos " + ", " $\approx$ " e " – " indicam onde o algoritmo na coluna correspondente funciona significativamente melhor que, equivalente a, ou pior do que o OSP-NSDE.	90
Tabela 9 – Frequência de classificação de cada algoritmo para o IGD, HV e Métrica <i>Ensemble</i>	91
Tabela 10 – Média e desvio padrão do IGD para os algoritmos baseados em PSO comparados. Os melhores resultados médios em cada instância estão em negrito. Os símbolos " + ", " $\approx$ " e " – " indicam onde o algoritmo na coluna correspondente funciona significativamente melhor que, equivalente a, ou pior do que o OSP-NSDE.	96
Tabela 11 – Média e desvio padrão do IGD para os DMOEAs comparados.	97
Tabela 12 – Média e desvio padrão do SP para os DMOEAs comparados.	98
Tabela 13 – Média e desvio padrão do HV para os DMOEAs comparados.	98
Tabela 14 – Média e desvio padrão do IGD para os DMOEAs baseados em predição atuais.	99

Tabela 15 – Média e variância da Soma dos Quadrados dos Erros (SSE), e média e variância do sinal de controle para o controlador MFAC ajustado com os parâmetros encontrados pelos otimizadores comparados.	104
---	-----

SUMÁRIO

1	INTRODUÇÃO	14
1.1	MOTIVAÇÃO	16
1.2	OBJETIVOS	17
1.3	METODOLOGIA	18
1.4	CONTRIBUIÇÕES	19
1.5	ESTRUTURA DO DOCUMENTO	20
2	ALGORITMOS EVOLUCIONÁRIOS MULTIOBJETIVOS EM AMBIENTES DINÂMICOS	22
2.1	PROBLEMAS DE OTIMIZAÇÃO MULTIOBJETIVO DINÂMICOS .	23
2.2	TAXONOMIA DOS ALGORITMOS EVOLUCIONÁRIOS MULTI- OBJETIVO DINÂMICOS	28
2.2.1	Estrutura do Otimizador Usado	29
2.2.2	Operadores de Variação	30
2.2.3	Mecanismo de Detecção da Mudança	32
2.2.4	Mecanismo de Reação à Mudança (CRM)	33
2.3	INDICADORES DE DESEMPENHO	38
2.3.1	Limitações dos Indicadores de Desempenho	42
2.4	CONCLUSÕES DO CAPÍTULO	44
3	MODELAGEM PROBABILÍSTICA	46
3.1	MODELO DE VARIÁVEIS LATENTES	46
3.1.1	Tipos de Modelos	48
3.2	INFERÊNCIA DA PROBABILIDADE A POSTERIORI	53
3.2.1	Métodos Determinísticos	55
3.2.2	Métodos de Simulação Estocástica	56
3.3	DISTRIBUIÇÃO PREDITIVA	58
3.4	CONCLUSÕES DO CAPÍTULO	58
4	ALGORITMO EVOLUCIONÁRIO MULTIOBJETIVOS DI- NÂMICO COM PREDIÇÃO NO ESPAÇO OBJETIVO . . .	60
4.1	FUNDAMENTOS PROBABILÍSTICOS PARA O MODELO PROPOSTO	61
4.1.1	Modelo AutoRegressivo (AR)	61
4.1.2	Modelo de Mistura Gaussiana (GMM)	62
4.1.3	Inferência Variacional (VI)	63

4.2	EVOLUÇÃO DIFERENCIAL DE ORDENAÇÃO NÃO DOMINADA MELHORADA DINÂMICA COM A ESTRATÉGIA DE PREDIÇÃO DO ESPAÇO OBJETIVO, DOSP-NSDE	65
4.2.1	Mecanismo de Detecção da Mudança	66
4.2.2	Estratégia de Previsão no Espaço Objetivo	67
4.2.3	Mecanismo de Reação de Mudança (CRM)	71
4.2.4	Evolução Diferencial com Ordenação por Não Dominância com Previsão no Espaço Objetivo, OSP-NSDE	71
4.2.5	Complexidade Computacional do DOSP-NSDE	76
5	RESULTADOS EXPERIMENTAIS	78
5.1	ANÁLISE DA SENSIBILIDADE DOS PARÂMETROS DO DOSP-NSDE	80
5.2	INFLUÊNCIA DA FREQUÊNCIA E DA SEVERIDADE DA MUDANÇA NO DOSP-NSDE	81
5.3	ANÁLISE DA ESTRATÉGIA OSP	83
5.3.1	Eficácia do Procedimento OSP como Mecanismo de Reação à Mudança	84
5.3.2	Eficácia do Procedimento OSP entre Mudanças (OSP-NSDE) .	86
5.4	COMPARAÇÃO DE DESEMPENHO DO DOSP-NSDE COM DMO-EAS BASEADOS EM PREVISÃO	97
5.5	RESOLUÇÃO DE UM PROBLEMA DO MUNDO REAL: OTIMIZAÇÃO DINÂMICA DOS PARÂMETROS DO CONTROLADOR ADAPTATIVO LIVRE DE MODELO PARA REATOR DE TANQUE AGITADO CONTINUAMENTE (CSTR)	99
6	CONCLUSÃO	106
6.1	LIMITAÇÕES E TRABALHOS FUTUROS.	107
6.2	TRABALHOS PUBLICADOS	108
	REFERÊNCIAS	111

1 INTRODUÇÃO

A maior parte dos problemas reais encontrados na área de otimização, - como por exemplo problemas de agendamento (*scheduling*), roteamento de veículos e problemas de sistemas de controle - envolve diversos objetivos que devem ser atingidos simultaneamente. Eles podem ser conflitantes, isto é, não existe uma solução única que otimize todos ao mesmo tempo. Problemas dessa natureza são chamados Problemas de Otimização Multiobjetivo (MOP) e, para eles, devem-se buscar um conjunto de soluções com compromissos entre seus objetivos. Os MOPs possuem um conjunto de funções objetivo a serem otimizadas (maximizar ou minimizar) e restrições que devem ser satisfeitas para que uma solução seja factível para o problema.

Nos meados do século XX, a programação multiobjetivo era baseada nas técnicas de otimização matemática clássicas: no método Simplex (DANTZIG et al., 1955) ou no método *Non-Inferior Set Estimation* (NISE) (COHON; CHURCH; SHEER, 1979). Tais métodos têm garantias matemáticas de obtenção do conjunto de soluções ótimas condicionadas a determinadas condições específicas de um dado problema, tipicamente, linearidade das funções objetivo e das restrições, ou convexidade do espaço de busca. Outras propostas com diferentes metodologias foram sendo empregadas (DEB, 2001): Método da Soma Ponderada, Método ε -Restrito e Método de Programação de Metas. Outra abordagem que não tardou a ser aplicada aos MOPs, e que vem sendo tema de intensa pesquisa atualmente nesse campo, são os métodos baseados em meta-heurísticas. Entre os algoritmos pioneiros de meta-heurísticos estão os Algoritmos Evolucionários (EAs) e, atualmente, há uma intensa pesquisa no desenvolvimento de outras meta-heurísticas com destaque para *Particle Swarm Optimization* (PSO) (KENNEDY; EBERHART, 1995). As meta-heurísticas multiobjetivo apresentam sobre as técnicas de programação matemática a vantagem de serem menos suscetíveis à forma do espaço de busca. No entanto, não há garantias matemáticas que um conjunto de soluções encontrado por esses métodos, seja o conjunto real de soluções ótimas para um dado MOP.

Os EAs têm sido largamente explorados em problemas de otimização. Uma das características mais importantes destes algoritmos é que possibilitam encontrar soluções ótimas ou adequadas para um problema complexo sem usar informação adicional sobre a paisagem de aptidão. A possibilidade de trabalhar com várias soluções simultaneamente, e poder escapar de ótimos locais fazem dos EAs uma técnica promissora a ser empregada nos MOPs.

Os EAs que são usados para resolver problemas de otimização de múltiplos objetivos simultaneamente são chamados Algoritmos Evolucionários Multiobjetivo (MOEA). Eles vieram dos EAs que foram inicialmente projetados para resolver problemas de objetivo único (MITCHELL, 1998; EIBEN; SMITH, 2003). O propósito principal dos MOEAs consiste

em encontrar um conjunto de soluções factíveis, ótimas e bem distribuídas (DEB, 2008). Estes algoritmos otimizam funções objetivo que são estáticas, porém diversos outros problemas do mundo real necessitam de otimização multiobjetivo *on-line*, onde as funções objetivo, restrições ou os parâmetros do problema podem variar ao longo do tempo (dinâmicos). Nestas situações, a utilização de um MOEA para realizar a otimização pode não ser eficiente se o tempo disponível entre duas alterações consecutivas do problema for relativamente curto, não havendo tempo hábil para a convergência do algoritmo entre as alterações ocorridas no ambiente. Por tanto, os MOEAs dinâmicos têm que convergir mais rapidamente e ser menos custosos que usar os MOEAs várias vezes.

Recentemente, a pesquisa sobre MOEAs propiciou sua adaptação e criação de novos algoritmos destinados à solução de problemas de otimização multiobjetivo dinâmicos (DMOPs). Um MOEA dinâmico (DMOEA) deve lidar com quaisquer dificuldades relacionadas ao espaço de busca ao longo do processo evolucionário. Portanto, um DMOEA bem-sucedido deve se caracterizar por ter rápida convergência e preservar diversidade necessária para se afastar de um conjunto de soluções e convergir para um novo, em seguida a mudanças no ambiente (FARINA; DEB; AMATO, 2004).

Uma dificuldade de se utilizar EAs em ambientes dinâmicos decorre da perda de diversidade genética da população, prejudicando a sua habilidade de adaptação (BUI et al., 2008). Vários modelos têm sido propostos na literatura para lidar com a dinâmica do ambiente, em especial com a perda de diversidade, variando a maneira como aumentam a diversidade através do Mecanismo de Reação à Mudança (CRM). Esses modelos podem ser divididos em abordagens baseadas em memória, os quais armazenam as melhores soluções encontradas em cada geração (YOUNES, 2006; DEB; KARTHIK et al., 2007; AZZOUZ; BECHIKH; Ben Said, 2014; CHEN; ZHANG; ZENG, 2015); abordagens baseadas em população as que utilizam subpopulações para explorar diversas áreas simultaneamente em uma superfície que se modifica periodicamente (CÁMARA; ORTEGA; TORO, 2007); abordagens baseadas em imigrantes que inserem novas soluções na população (DEB; KARTHIK et al., 2007; AZEVEDO; ARAÚJO, 2011b; WAN; WANG, 2013); ou abordagens baseadas em predição (KOO; GOH; TAN, 2010; LI et al., 2014; MURUGANANTHAM; TAN; VADAKKEPAT, 2016) as quais orientam o deslocamento da população através de um mecanismo de previsão. Há indicações relevantes de desempenho positivo dos métodos baseados em previsão (AZZOUZ; BECHIKH; SAID, 2017) por analisarem dados anteriores (conhecimento anterior ou soluções anteriores) para diminuir o tempo necessário à adaptação, unindo evolução ao longo das gerações e potencial incorporação de conhecimento para cada geração.

A pesquisa sobre aprendizagem de máquina estatística aborda o desafio da análise de dados mediante o desenvolvimento da área de modelagem probabilística, um campo que oferece uma abordagem formal para desenvolver novas metodologias para análise (PEARL, 1988; JORDAN, 1999; BISHOP, 2007; KOLLER; FRIEDMAN, 2009; MURPHY, 2012). Esta área articula a relação entre as variáveis escondidas e observadas com uma distribuição de

probabilidade fatorizada. Na modelagem probabilística são utilizados modelos de variáveis latentes ^[1] (LVM), que pressupõem que um complexo conjunto de dados observados exibe padrões mais simples mas não observados (BLEI, 2014).

A modelagem probabilística usa algoritmos de inferência para estimar os parâmetros do modelo através do cálculo da distribuição a posteriori: a distribuição condicional da estrutura oculta dada obtida a partir de observações. Os correspondentes algoritmos de inferência a posteriori permitem analisar os dados sob essas premissas, isto é, inferir a estrutura oculta que melhor explique as observações (HOFFMAN et al., 2013).

Para muitos modelos de interesse, uma distribuição a posteriori pode não ser tratável computacionalmente, implicando no uso de métodos aproximados (HOFFMAN et al., 2013). As duas estratégias de aproximação mais empregadas em estatística e aprendizagem de máquina são *Markov Chain Monte Carlo* (MCMC) e Inferência Variacional (VI). Na amostragem MCMC, constrói-se uma cadeia de Markov sobre as variáveis ocultas cuja distribuição estacionária é a distribuição a posteriori de interesse (METROPOLIS et al., 1953; HASTINGS, 1970; GEMAN; GEMAN, 1984; GELFAND; SMITH, 1990; ROBERT; CASELLA, 2004). Em Inferência Variacional, define-se uma família flexível de distribuições sobre as variáveis ocultas, indexada por parâmetros livres (JORDAN et al., 1999; WAINWRIGHT; JORDAN, 2008), encontrando a configuração dos parâmetros (o membro da família), que é mais próximo da distribuição a posteriori. Assim, resolve-se o problema de inferência através da resolução de um problema de otimização (HOFFMAN et al., 2013). A VI é uma das técnicas mais conhecidas e rápidas para estimar os parâmetros de um modelo.

Normalmente, são realizadas dois tipos de tarefas com um modelo: exploração e predição. Na exploração usam-se inferências sobre as variáveis ocultas, geralmente através da estimação da distribuição a posteriori, para resumir e visualizar os dados, ou conhecer suas características, isto é, dividi-lo em grupos com estruturas ditadas pelas inferências. Na predição, são projetados ou estimados os dados futuros usando a distribuição preditiva a posteriori. Ela permite que se façam previsões em modelos de regressão e em estudos que envolvem um componente temporal (KINAS; ANDRADE, 2014).

A combinação da modelagem probabilística com os EAs pode resultar interessante para manter um equilíbrio entre exploração e exploração durante o processo evolucionário, assim como para lidar de maneira rápida com as alterações do ambiente e para acelerar processos nos quais tendências comportamentais são detectadas.

1.1 MOTIVAÇÃO

Há alguns pontos relevantes que motivaram a escolha do tópico de pesquisa desta tese. Entre eles, os principais estão enunciados a seguir.

^[1] São variáveis que não podem ser medidas nem observadas diretamente pelo investigador. Por isso, explica-se a sua influência através de hipóteses, e do estudo de outras variáveis.

Característica favorável dos EAs. Observa-se que, apesar da literatura para solucionar DMOPs, o uso de EAs para a solução destes tipos de problemas é um tópico de pesquisa ainda em aberto. Os EAs são candidatos para resolver problemas multiobjetivo dinâmicos devido à elevada complexidade do espaço de busca destes problemas: a exploração progressiva de ambientes desconhecidos está entre as melhores características dos EAs desde a sua proposição na década de 1960.

Aumento da velocidade de convergência dos DMOEAs. Uns dos requisitos mais desejados de algoritmos evolucionários é aumentar a velocidade de convergência dos algoritmos. Em particular, os DMOPs demandam encontrar novas soluções ótimas antes que aconteça uma nova mudança do ambiente.

Adaptação às mudanças nos DMOEAs. Quando um AE converge ou se aproxima de uma solução, ele pode ter diversidade reduzida provocando dificuldades para buscar soluções após uma mudança no ambiente. Uma solução simples é reiniciar o algoritmo após cada mudança no ambiente, alternativa potencialmente custosa e ineficiente, pois informações valiosas sobre o histórico de busca são descartadas a cada reinício. O uso de métodos que se adaptem rapidamente à nova paisagem podem mostrar índices de sucesso maior em relação à reinicialização da população.

Abordagens de EAs baseadas em aprendizagem. Estes métodos vêm ganhado aceitação entre os pesquisadores, evidenciado pelo número crescente de publicações na área nos últimos anos. Novas possibilidades de pesquisa para EAs podem incluir outro tipo de conhecimento, extraindo o comportamento probabilístico da população e a prole gerada, com representações dos indivíduos no espaço probabilístico. Esta experiência pode ser relevante para compreender a dinâmica do algoritmo nas etapas de exploração e de exploração. Um solucionador bem-sucedido deve ter capacidade de manter a diversidade ao longo do processo de busca através da retenção de informações úteis do passado. Este requisito adiciona outra dimensão à questão tradicional de balancear a exploração e a exploração.

1.2 OBJETIVOS

Esta tese tem como objetivo principal elaborar e implementar um modelo de algoritmo evolucionário empregando conceitos da modelagem probabilística para resolver problemas de otimização multiobjetivos dinâmicos, de maneira que possa rastrear e prever movimentações da fronteira do Pareto, bem como acelerar a convergência do algoritmo antes de ocorrer uma nova modificação do ambiente, melhorando também a qualidade das soluções obtidas em relação aos algoritmos da literatura.

Os objetivos específicos da pesquisa são:

- Descrever o estado da arte dos Algoritmos Evolucionários Multiobjetivos em ambientes dinâmicos.

- Propor nova taxonomia baseada na forma em que os MOEAs dinâmicos inserem diversidade.
- Apresentar o estado da arte da Modelagem Probabilística e analisar a sua aplicabilidade em conjunto com os EAs.
- Propor e implementar um DMOEA probabilístico para reagir rapidamente às mudanças e melhorar a convergência do processo evolucionário.
- Validar e comparar o método proposto considerando diferentes funções de teste de múltiplos objetivos dinâmicos.

1.3 METODOLOGIA

Para atingir o objetivo principal, esta pesquisa está focada nos algoritmos baseados em predição. Geralmente, estes DMOEAs executam o procedimento de regressão no espaço de decisão, geralmente um espaço com dimensões significativamente mais altas que o espaço objetivo, espaço onde estão as funções objetivo. Apesar dessa prática, as informações extraídas do espaço objetivo tendem a ser mais confiáveis e consistentes para estabelecer as tendências dos movimentos da frente no espaço objetivo que tendem a ocorrer em direção à Frente Ótima de Pareto POF através da inserção de novas soluções não dominadas. Portanto, um deslocamento da população baseado em movimentos anteriores de indivíduos no espaço de decisão pode levar o processo a regiões não promissoras do espaço objetivo, prejudicando a convergência do algoritmo.

Com base nessas estratégias e tendo em mente o pressuposto mencionado, esta tese propõe um algoritmo dinâmico baseado na Evolução Diferencial com Ordenação pela Não Dominância NSDE com Previsão no Espaço Objetivo (*Dynamic Non-dominated Sorting Differential Evolution improvement with Prediction in the Objective Space* (DOSP-NSDE)). Quando várias premissas baseadas na métrica de Hipervolume Aproximado são alcançadas, o DOSP-NSDE aciona a estratégia de Previsão de Espaço do Objetivo (OSP). O OSP estima a movimentação de indivíduos não dominados no espaço objetivo que possibilita encontrar rapidamente regiões promissoras no espaço de busca e, portanto, gerar uma nova população considerando essas regiões. Os operadores de variação da Evolução Diferencial (DE) e a mutação polinomial são usados para produzir a descendência sempre que a condição de OSP não for satisfeita. Os resultados obtidos nas simulações de diversas pesquisas mostrou as vantagens de usar a combinação destes operadores para melhorar o desempenho dos algoritmos evolucionários. A proposta desta tese emprega a estratégia de previsão tanto na fase estática (sem mudanças no ambiente) do processo evolucionário como no mecanismo de reação à mudança para prever a nova localização da frente ótima.

Primeiramente, a eficácia do OSP foi verificada no processo estático. Para isto, foram realizados experimentos para 31 funções multiobjetivas bem conhecidas, comparando o

OSP-NSDE (a versão estática do DOSP-NSDE) com outros MOEAs e com modelos de inteligência coletiva. Comprovada a eficácia e a eficiência da proposta em ambientes estáticos, foram realizados experimentos em 19 funções de teste multiobjetivas dinâmicas e em um problema do mundo real. As comparações do DOSP-NSDE com diversos algoritmos da literatura empregaram três métricas diferentes e evidenciaram que a proposta é competitiva na maioria dos problemas, atingindo-se o objetivo principal da pesquisa.

1.4 CONTRIBUIÇÕES

As principais contribuições da tese são listadas a seguir:

- Uma nova taxonomia para os DMOEAs é proposta considerando como os algoritmos dinâmicos aumentam a diversidade. Assim, estes algoritmos evolucionários são agrupados em DMOEAs Endógenos, DMOEAs Exógenos e DMOEAs Híbridos.
- Ao contrário da maioria dos DMOEAs existentes, o DOSP-NSDE emprega métodos probabilísticos para prever tanto o processo evolucionário quanto as mudanças no ambiente.
- O DOSP-NSDE considera as informações de distribuição da população e as informações locais de cada indivíduo. As informações de distribuição da população são consideradas em algumas gerações do processo evolucionário usando a modelagem probabilística. Por outro lado, as informações locais de cada indivíduo são consideradas usando os operadores de variação.
- O DOSP-NSDE aciona a modelagem probabilística dependendo de um critério baseado na métrica Hipervolume aproximada que estabelece tendências do movimento para as soluções não dominadas no espaço objetivo. A métrica HV foi escolhida porque a maioria dos problemas reais possui uma Fronteira ou Frente Ótima de Pareto (POF) desconhecida e porque esta métrica apresenta correlação com a diversidade (AZEVEDO; ARAÚJO, 2011a).
- DOSP-NSDE extrai informações do espaço de decisão e do espaço objetivo. A informação estatística levantada a partir do espaço objetivo, usada no processo de previsão, pode aumentar a velocidade de convergência, em relação às gerações, por captar tendências do comportamento do virtual movimento dos indivíduos não dominados. A informação gerada a partir do espaço de decisão, empregada na geração de indivíduos que complementarão uma nova população, cria uma população com soluções mais esparsadas na vizinhança de uma frente prevista.
- O DOSP-NSDE usa métodos estatísticos formais para realizar uma busca efetiva orientada no espaço de soluções. O método proposto prevê o emprego de um modelo de série temporal (ARX) considerando a linha do tempo das gerações. Além disso,

a Inferência Variacional (VI) é usada para determinar os parâmetros de um Modelo de Mistura Gaussiana (GMM).

- Este estudo demonstra que mesmo procedimentos complexos podem ser executados em tempo hábil nos DMOEAs e não acrescentar altos custos para os modelos.
- Extensos experimentos foram realizados para analisar o desempenho da proposta na aceleração da convergência do processo evolucionário. Foram escolhidos alguns problemas de teste MOPs bem conhecidos: os problemas UF, os problemas de Zitzler-Deb-Thiele (ZDT), e os conjuntos de testes DTLZ e WFG. Os resultados das comparações com onze MOEAs selecionados mostram uma vantagem significativa no emprego da proposta desenvolvida.
- Extensos experimentos foram realizados para comparar o DOSP-NSDE com outros algoritmos de última geração, em três conjuntos de testes: os problemas FDA, conjunto de testes UDF e os problemas DF. Os resultados mostram que o DOSP-NSDE é competitivo de acordo com três indicadores de desempenho utilizados nas comparações.
- Além dos problemas de teste, o DOSP-NSDE é usado para resolver um problema de otimização multiobjetivo dinâmico no mundo real: Otimização Dinâmica dos Parâmetros do Controlador Adaptativo Livre de Modelo para um Reator de Tanque Agitado Continuamente (CSTR)

1.5 ESTRUTURA DO DOCUMENTO

Além deste capítulo, este documento possui em mais seis capítulos (ver Figura 1) que estão organizados da seguinte forma:

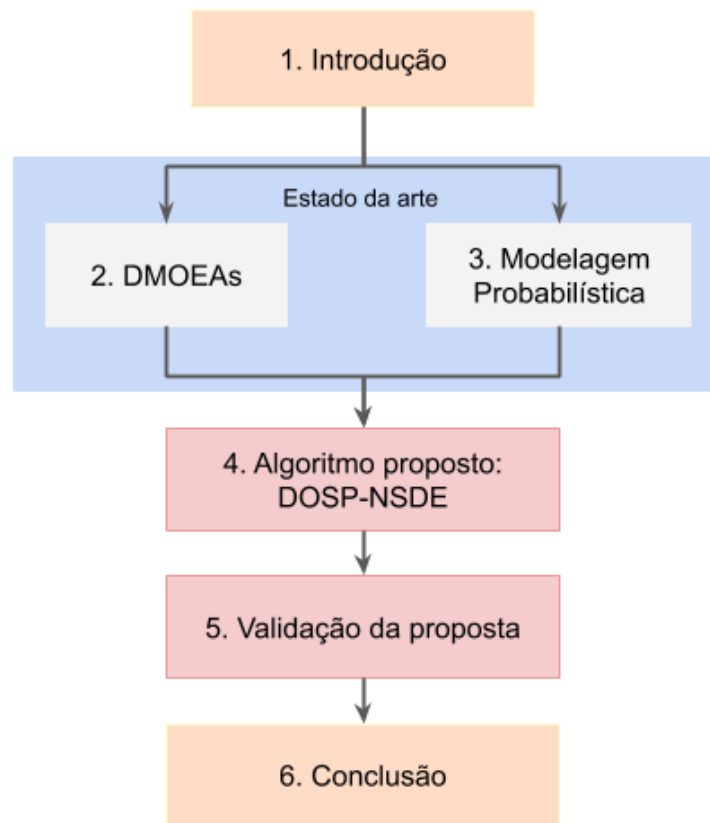
Capítulo 2: Apresenta o estado da arte dos algoritmos evolucionários de múltiplos objetivos em ambientes dinâmicos. Alguns conceitos importantes, tal como a definição de DMOP, assim como uma breve revisão das funções de teste mais usadas pelos pesquisadores é também mostrada. Por último, no capítulo são tratadas as medidas e métricas mais utilizadas para avaliar os DMOEAs, assim como os problemas que podem surgir devido à escolha inadequada das métricas.

Capítulo 3: Apresenta o estado da arte dos modelos de variáveis latentes, assim como dos métodos de inferência usados para estimar os parâmetros dos mesmos.

Capítulo 4: Introduz os fundamentos probabilísticos usados no DMOEA proposto nesta tese, (DOSP-NSDE). Além disso, descreve-se o DMOEA proposto com a estratégia de predição no espaço objetivo. Os passos desta estratégia OSP são também apresentados.

Capítulo 5: Descreve os experimentos realizados para a validação e comparação do modelo proposto com diferentes funções de teste multiobjetivo dinâmicos. Também, são

Figura 1 – Organização da Tese



Fonte: Autor

apresentados os resultados obtidos na resolução de um problema real de controle usando o DOSP-NSDE.

Capítulo 6: Apresenta as considerações finais, limitações e sugestões para trabalhos futuros.

2 ALGORITMOS EVOLUCIONÁRIOS MULTIOBJETIVOS EM AMBIENTES DINÂMICOS

Existem problemas que, com o passar do tempo, tanto o seu ambiente quanto as suas características e restrições podem sofrer mudanças. Estas mudanças podem acarretar a modificação da solução ótima do problema e, portanto, as mudanças devem ser consideradas no processo de otimização (HELBIG; ENGELBRECHT, 2014). Muitos problemas do mundo real são dinâmicos, como por exemplo, problemas de agendamento (*scheduling*), roteamento de veículos e problemas de sistemas de controle.

Se a dinamicidade dos problemas for tratada através de uma nova otimização desde o seu início, os métodos de otimização existentes para o problema estático podem ser aplicados. A utilização de MOEAs para problemas dinâmicos surgiu para estender o uso dos EAs de múltiplos objetivos aos problemas dinâmicos. No entanto, a reinicialização pode elevar o tempo do algoritmo, demorar a convergência e até pode não ser encontrada a nova solução. Neste caso, não haverá tempo hábil para otimizar o problema e, assim, o algoritmo poderá não ser adequado.

Os Algoritmos Evolucionários Multiobjetivos Dinâmicos (DMOEAs) devem convergir em um tempo razoável, evitando que as suas soluções se tornem obsoletas. Nestes algoritmos há necessidade de aplicação de algum mecanismo de preservação de diversidade, já que ela tende a ser reduzida devido à necessidade de rápida convergência. A severidade e a frequência das mudanças podem também influenciar no funcionamento do algoritmo: se a severidade da mudança for muito alta, o algoritmo deverá aumentar a sua diversidade e exploração, mas ainda assim poderá não encontrar o ótimo ou ficar preso em algum ótimo local, se o aumento da diversidade não for suficiente. Se a frequência de ocorrência das mudanças for alta, a convergência deverá ser suficientemente rápida para que a sua velocidade de evolução consiga atingir as soluções ótimas antes da ocorrência da próxima mudança. Além disto, o custo computacional é um fator determinante para um DMOEA ser bem-sucedido.

Portanto, um DMOEA deve lidar com problemas de otimização multiobjetivo dinâmicos, cujas soluções ótimas são modificadas ao longo do processo de otimização decorrente de:

- Mudança do ambiente;
- Mudança das restrições;
- Mudança dos objetivos;
- Mudança das configurações do problema (representação, por exemplo).

Na próxima secção são apresentados os conceitos básicos relativo a DMOPs. Em seguida, uma nova taxonomia dos MOEAs para ambientes dinâmicos é proposta na Seção 2.2, baseada na forma em que os algoritmos inserem diversidade. Na Seção 2.3 são abordados os indicadores de desempenho mais usados para avaliar os DMOEAs.

2.1 PROBLEMAS DE OTIMIZAÇÃO MULTIOBJETIVO DINÂMICOS

Um Problema de Otimização Multiobjetivo Dinâmico (DMOP) se caracteriza por possuir mais que uma função objetivo (MOP) onde, com o passar do tempo, o seu ambiente, suas características e suas restrições podem sofrer mudanças. O DMOP é matematicamente definido como:

$$\begin{aligned} \min_{\mathbf{x}} \mathbf{f}(\mathbf{x}, t) &= [f_1(\mathbf{x}, t) \ f_2(\mathbf{x}, t) \ \dots \ f_m(\mathbf{x}, t)]^T \\ \text{sujeito a:} \\ \mathbf{x} &\in [x_{min}^i, x_{max}^i]^n \\ g_j(\mathbf{x}, t) &\leq 0, j = 1, \dots, q \\ h_j(\mathbf{x}, t) &= 0, j = q + 1, \dots, p \end{aligned} \tag{2.1}$$

onde $\mathbf{x} \in \mathbb{R}^n$ é o vetor de decisão também denominado vetor de solução, e n é o número total de variáveis de decisão. Os valores $[x_{min}, x_{max}]^n$ representam os limites inferior e superior, respectivamente, para as variáveis de decisão. Esses limites, também chamados restrições paramétricas, definem um espaço n -dimensional $S \in \mathbb{R}^n$. O espaço de decisão factível $S_{dec} \subseteq S$ é definido pelo conjunto de p restrições adicionais lineares ou não lineares. q é o número de restrições de desigualdade e $p - q$ é o número de restrições de igualdade. $f_i(\mathbf{x}, t) \in S_{obj}$ é a i -ésima função objetivo, e S_{obj} é o espaço objetivo. m é o número total de objetivos, $\mathbf{f}(\mathbf{x}, t) \in S_{obj} \subseteq \mathbb{R}^m$ é o vetor de funções objetivo. Para cada solução $\mathbf{x}$ no espaço de decisão, existe um $\mathbf{f}(\mathbf{x})$ em S_{obj} .

As funções objetivo empregadas nos MOPs tanto estático como dinâmico podem ser conflitantes entre si. Uma função objetivo f_1 é conflitante com uma outra função f_2 quando não é possível melhorar o valor de uma delas sem piorar o da outra. Neste tipo de problemas, para comparar duas soluções factíveis do problema, emprega-se o conceito de dominância de Pareto.

Definição 1 *Dominância de um vector de decisão: Uma solução $\mathbf{x}$ domina uma outra solução $\mathbf{y}$ (representado como $\mathbf{x} \preceq \mathbf{y}$), em um processo onde as funções objetivo devem ser minimizadas, se e só se:*

- A solução $\mathbf{x}$ não é pior que $\mathbf{y}$ em todos os objetivos, isto é, $f_j(\mathbf{x}) \leq f_j(\mathbf{y}), \forall j = 1, 2, \dots, m$.
- A solução $\mathbf{x}$ é estritamente melhor que $\mathbf{y}$ pelo menos em um objetivo, isto é, $\exists i = 1, \dots, m : f_i(\mathbf{x}) < f_i(\mathbf{y})$.

Se ambas condições da Definição 1 forem satisfeitas, pode-se dizer que: 1) $\mathbf{y}$ é dominada por $\mathbf{x}$; 2) $\mathbf{x}$ é não-dominada por $\mathbf{y}$; 3) $\mathbf{x}$ é não inferior que $\mathbf{y}$.

A relação de dominância satisfaz as seguintes propriedades:

1. Não é reflexiva. Conforme a Definição 1, uma solução não pode ser dominada por si mesma;
2. Não é simétrica, isto é, $\mathbf{x} \preceq \mathbf{y}$ implica que $\mathbf{y} \preceq \mathbf{x}$ é falso;
3. Transitiva, se $\mathbf{x} \preceq \mathbf{y}$ e $\mathbf{y} \preceq \mathbf{z}$ então $\mathbf{x} \preceq \mathbf{z}$.

Um conjunto de soluções pode ser dividido em conjunto de soluções dominadas e não-dominadas empregando o operador de dominância. Os vetores de decisão não dominados são chamados ótimos de Pareto e são definidos como:

Definição 2 *Ótimo de Pareto: Um vetor $\mathbf{x}^*$ é ótimo de Pareto se não existe um vetor $\mathbf{x} \neq \mathbf{x}^* \in S_{dec}$ que o domine, isto é, $\nexists k : f_k(\mathbf{x}) < f_k(\mathbf{x}^*)$. Se $\mathbf{x}^*$ é ótimo de Pareto, então o vetor dos objetivos, $\mathbf{f}(\mathbf{x}^*)$, é também ótimo de Pareto*

O conjunto de todas as soluções ou vetores de decisão ótimos de Pareto é conhecido como Conjunto Ótimo de Pareto (POS), o qual representa as soluções ótimas do problema e está definido como:

Definição 3 *Conjunto Ótimo de Pareto POS^* : É formado pelo conjunto de todos os vetores de decisão ótimos de Pareto, isto é:*

$$POS^* = \{\mathbf{x}^* \in S_{dec} | \nexists \mathbf{x} \in F : \mathbf{x} \preceq \mathbf{x}^*\} \quad (2.2)$$

O conjunto de valores das funções objetivo das soluções do conjunto ótimo de Pareto é chamado Fronteira ou Frente Ótima de Pareto (POF). Ela é definida como:

Definição 4 *Fronteira Ótima de Pareto (POS^*): Para o vetor objetivo $\mathbf{f}(\mathbf{x})$, cujo conjunto ótimo de Pareto é POS^* , a Fronteira Ótima de Pareto $POF^* \subseteq S_{obj}$ é definida como:*

$$POF^*(t) = \{\mathbf{f}(\mathbf{x}^*) = (f_1(\mathbf{x}^*), f_2(\mathbf{x}^*), \dots, f_m(\mathbf{x}^*)) | \mathbf{x}^* \in POS^*\}, \quad (2.3)$$

Assim, o objetivo de um algoritmo para resolver DMOPs é rastrear a POF no tempo:

$$POF^*(t) = \{\mathbf{f}(\mathbf{x}^*, t) = (f_1(\mathbf{x}^*, t), f_2(\mathbf{x}^*, t), \dots, f_m(\mathbf{x}^*, t)) | \mathbf{x}^* \in POS^*(t)\}, \quad (2.4)$$

encontrando um conjunto bem distribuído de soluções na Frente de Pareto em cada mudança, com a maior eficiência computacional possível.

Os pesquisadores têm usados diferentes características para classificar os DMOPs, uma das mais usadas foi proposta por (FARINA; DEB; AMATO, 2004) que apontaram quatro tipos de acordo com as mudanças que afetam a frente ótima de Pareto (POF) e o conjunto

ótimo de Pareto (*POS*). Os problemas são de Tipo I quando o *POS* se modifica enquanto o *POF* permanece invariante; Tipo II quando ambos *POS* e *POF* modificam-se; Tipo III quando a *POF* modifica-se enquanto o *POS* permanece invariante; e Tipo IV quando ambos *POS* e *POF* permanecem invariantes, embora seja possível que a topologia da aptidão se modifique ao longo do tempo. A Tabela 1 resume os casos.

Tabela 1 – Tipos de DMOPs segundo (FARINA; DEB; AMATO, 2004).

DMOPs	Mudança da POS	Sem mudança na POS
Mudança na POF	Tipo II	Tipo III
Sem mudança na POF	Tipo I	Tipo IV

(TANTAR; TANTAR; BOUVRY, 2011) propuseram uma classificação baseada nas causas das mudanças dinâmicas:

- Primeira ordem: As variáveis de decisão mudam com o tempo.
- Segunda ordem: As funções objetivo mudam ao longo do tempo.
- Terceira ordem: Os valores atuais das variáveis de decisão ou das funções objetivo dependem de seus valores anteriores.
- Quarta ordem: Os ambientes inteiros ou parte deles mudam ao longo do tempo.

Recentemente, alguns estudos envolvendo EAs dinâmicos adaptaram MOPs para criar benchmarks, permitindo a avaliação e comparação de DMOEAs. Geralmente, estes são problemas artificiais irrestritos que possuem características particulares e, portanto, avaliam se um algoritmo pode superar dificuldades específicas. A Tabela 2 mostra um resumo dos DMOPs e os trabalhos nos que tem-se usados os mesmos. A quantidade de algoritmos testados nas funções de teste FDA (FARINA; DEB; AMATO, 2004) e dMOP(GOH; TAN, 2009) mostram uma tendência dos pesquisadores a utilizar estes problemas relativamente simples. Dada a variedade de funções que são usadas e que continuam se criando, é percebido também, que ainda não existe um conjunto de problemas padrões para avaliar os DMOEAs. Apesar de existirem vários DMOPs reais, poucos DMOEAs têm sido aplicados para resolvê-los (HELBIG; DEB; ENGELBRECHT, 2016). Geralmente, o grau de dificuldade de um problema real é desconhecido, e em muitos casos, a verdadeira *POF* deles não é conhecida.

Uma vez definidos os conceitos fundamentais dos DMOPs, e feita uma breve discussão das funções de teste mais usadas pelos pesquisadores, e dos problemas reais, será apresentado, na próxima seção, o estado da arte dos algoritmos evolucionários de múltiplos objetivos em ambientes dinâmicos.

Tabela 2 – Problemas de Otimização Multiobjetivo Dinâmicos e os estudos em que foram usados.

DMOPs	Estudos
<i>Moving Peaks</i> (BRANKE; BRANKE, 1999)	(BUI; ABBASS; BRANKE, 2005)
<i>Benchmark Generator</i> (JIN; SENDHOFF, 2004)	(LIU; WANG, 2007; LIU, 2010)
FDA (FARINA; DEB; AMATO, 2004)	(FARINA; DEB; AMATO, 2004; GARIS, 2006; MEHNEN; WAGNER; RUDOLPH, 2006; HATZAKIS; WALLACE, 2006; BINGUL, 2007; ZHENG, 2007; DEB; KARTHIK et al., 2007; CÁMARA; ORTEGA; TORO, 2007; ZHOU et al., 2007; LIU; WANG, 2007; WANG; DANG, 2008; GOH; TAN, 2009; KOO; GOH; TAN, 2010; CHEN; LI; CHEN, 2009; WANG; LI, 2009; WANG; LI, 2010; LIU, 2010; AZEVEDO; ARAÚJO, 2011b; GARROZI; ARAUJO, 2011; LIU; ZENG, 2012; MURUGANANTHAM et al., 2013; WAN; WANG, 2013; AZZOUZ; BECHIKH; Ben Said, 2014; ZHOU; JIN; ZHANG, 2014; LI et al., 2014; LIU et al., 2014; WU; JIN; LIU, 2015; LIU et al., 2015; PENG et al., 2015; CHEN et al., 2015; AZZOUZ; BECHIKH; SAID, 2015a; CHEN; ZHANG; ZENG, 2015; MURUGANANTHAM; TAN; VADAKKEPAT, 2016; JIANG; YANG, 2016; LIU; LIU, 2016; SAHMOUD; TOPCUOGLU, 2016b; GEE; TAN; ALIPPI, 2017; RUAN et al., 2017; FU; SUN, 2017; ZOU et al., 2017; CHEN et al., 2018)
DSW (MEHNEN; WAGNER; RUDOLPH, 2006)	(MEHNEN; WAGNER; RUDOLPH, 2006)
ZJZ (ZHOU et al., 2007)	(ZHOU et al., 2007; MURUGANANTHAM et al., 2013; CHEN et al., 2018)
$DTLZ_{Av}$ (AVDAGIĆ; KONJICIJA; OMANOVIĆ, 2009; JIANG; YANG, 2016)	(AVDAGIĆ; KONJICIJA; OMANOVIĆ, 2009)

dMOP (GOH; TAN, 2009)	(GOH; TAN, 2009; AZEVEDO; ARAÚJO, 2011b; WAN; WANG, 2013; AZZOUZ; BECHIKH; Ben Said, 2014; MURUGANANTHAM et al., 2013; ZHOU; JIN; ZHANG, 2014; PENG; ZHENG; ZOU, 2014; WU; JIN; LIU, 2015; LIU et al., 2015; PENG et al., 2015; CHEN et al., 2015; CHEN; ZHANG; ZENG, 2015; MURUGANANTHAM; TAN; VADAKKEPAT, 2016; JIANG; YANG, 2016; SAHMOUD; TOPCUOGLU, 2016b; RUAN et al., 2017; ZOU et al., 2017; CHEN et al., 2018)
DIMP (KOO; GOH; TAN, 2010)	(KOO; GOH; TAN, 2010; GARROZI; ARAUJO, 2011)
DMZDT (WANG; LI, 2010)	(WANG; LI, 2009; WANG; LI, 2010; AZZOUZ; BECHIKH; SAID, 2015a)
WYL (WANG; LI, 2010)	(WANG; LI, 2009; WANG; LI, 2010)
ZJZ2 (ZHOU; JIN; ZHANG, 2014)	(ZHOU; JIN; ZHANG, 2014; PENG; ZHENG; ZOU, 2014; WU; JIN; LIU, 2015; PENG et al., 2015; MURUGANANTHAM; TAN; VADAKKEPAT, 2016; RUAN et al., 2017; ZOU et al., 2017)
UDF (BISWAS et al., 2014)	(BISWAS et al., 2014; JIANG; YANG, 2016; FU; SUN, 2017; CHEN et al., 2018)
WYL2 (WU; JIN; LIU, 2015)	(WU; JIN; LIU, 2015)
MT (MURUGANANTHAM; TAN; VADAKKEPAT, 2016)	(MURUGANANTHAM; TAN; VADAKKEPAT, 2016)
CEC2015 (HELBIG; ENGELBRECHT, 2015)	(JIANG et al., 2017; JIANG et al., 2018)
GTA (GEE; TAN; ABBASS, 2017)	(GEE; TAN; ABBASS, 2017; GEE; TAN; ALIPPI, 2017)
JY (JIANG; YANG, 2017)	(JIANG; YANG, 2017)
CEC2018 (JIANG et al., 2018)	
CDMOPs	DMOEAs
DTF (MEHNEN; WAGNER; RUDOLPH, 2006)	(MEHNEN; WAGNER; RUDOLPH, 2006)
DCTP (AZZOUZ; BECHIKH; SAID, 2015b)	(AZZOUZ; BECHIKH; SAID, 2015b)
DMaOPs	DMOEAs
SJY (JIANG; YANG, 2014)	(SAHMOUD; TOPCUOGLU, 2016a; SAHMOUD; TOPCUOGLU, 2016b)

DMOPs reais	DMOEAs
Controle da combustão em um queimador de lixo	(FARINA; DEB; AMATO, 2004)
Agendamento dinâmico da energia hidrotermal	(DEB; KARTHIK et al., 2007)
Alocação de força ideal para simulação de combate	(BINGUL, 2007)
Identificação do conjunto ótimo de parâmetros para materiais de gradiente de usinagem	(ROY; MEHNEN, 2008)

2.2 TAXONOMIA DOS ALGORITMOS EVOLUCIONÁRIOS MULTIOBJETIVO DINÂMICOS

A evolução intrínseca aos EAs sugere que eles sejam aplicados aos problemas dinâmicos, realidade presente na literatura, em especial nas duas últimas décadas, quando a quantidade de publicações teve considerável aumento. Para resolver DMOPs, um algoritmo evolucionário multiobjetivo dinâmico é geralmente dividido nas seguintes etapas (PENG et al., 2015):

1. Definir os parâmetros do algoritmo;
2. Inicializar e avaliar a população;
3. Se as condições de parada não forem satisfeitas:
 - Executar um algoritmo de evolucionário;
 - Executar método de detecção de mudanças: Uma mudança de ambiente deve ser detectada para permitir uma resposta do algoritmo;
 - Executar mecanismo de reação da mudança: O mecanismo deve responder com rapidez e precisão a mudanças no ambiente, para que a população possa convergir para o novo conjunto de soluções ótimas.

Para projetar um DMOEA bem sucedido, quatro componentes principais podem ser alterados: 1) a estrutura do otimizador usado (MOEA), denominado na literatura como a abordagem para atribuição de aptidão; 2) os operadores de variação; 3) o mecanismo de detecção da mudança; 4) o mecanismo de reação à mudança (CRM).

2.2.1 Estrutura do Otimizador Usado

De acordo com a estrutura do MOEA usado, os DMOEAs podem ser baseados em Pareto, baseados em Indicadores ou baseados em Decomposição (TRIVEDI et al., 2017). Nos DMOEAs baseados em Pareto (MEHNEN; WAGNER; RUDOLPH, 2006), (GARIS, 2006), (HATZAKIS; WALLACE, 2006), (ZHOU et al., 2007), (DEB; KARTHIK et al., 2007), (CÁMARA; ORTEGA; TORO, 2007), (LIU; WANG, 2007), (KHALED; TALUKDER; KIRLEY, 2008), (WANG; DANG, 2008), (ROY; MEHNEN, 2008), (WANG; LI, 2009), (GOH; TAN, 2009), (LIU, 2010), (KOO; GOH; TAN, 2010), (AZEVEDO; ARAÚJO, 2011b), (GARROZI; ARAUJO, 2011), (LIU; ZENG, 2012), (WAN; WANG, 2013), (LI et al., 2014), (ZHOU; JIN; ZHANG, 2014), (PENG; ZHENG; ZOU, 2014), (AZZOUZ; BECHIKH; Ben Said, 2014), (LIU et al., 2014), (AZZOUZ; BECHIKH; SAID, 2015a), (AZZOUZ; BECHIKH; SAID, 2015b), (WU; JIN; LIU, 2015), (PENG et al., 2015), (JIANG; YANG, 2016), (SAHMOUD; TOPCUOGLU, 2016b), (GUERRERO-PEÑA; ARAÚJO, 2017), (JIANG et al., 2017), (RUAN et al., 2017), (ZOU et al., 2017), (JIANG et al., 2018), todos os objetivos são otimizados simultaneamente. Geralmente, os EAs baseados em Pareto são computacionalmente mais rápidos do que aqueles baseados em indicadores ou decomposição. Geralmente, esses métodos estão entre as abordagens mais simples com poucos parâmetros. O princípio de domínio de Pareto (DEB et al., 2002) é usado para atribuir aptidão a soluções. Se a estrutura baseada em Pareto for usada, a perda de diversidade pode ocorrer quando o processo evolucionário é executado. Portanto, um mecanismo de diversidade é necessário para selecionar soluções que tenham a mesma classificação, tanto no processo de seleção de pais quanto de sobreviventes. Os esquemas de preservação da diversidade evitam a concentração da exploração da população em alguns nichos do espaço objetivo, ou seja, a manutenção da diversidade no algoritmo baseado em Pareto ajuda na sua capacidade de pesquisa. Entre os mecanismos de diversidade mais amplamente utilizados, podem ser citados o Método de Compartilhamento de Aptidão, Método Baseado em Clustering, Método baseado na Distância de Aglomeração, Método de Estimativa de Densidade Baseado em Célula, Método Baseado no Vizinho Mais Próximo (VACHHANI; DABHI; PRAJAPATI, 2015).

Em algoritmos baseados em indicadores (CHEN; LI; CHEN, 2009; WANG; LI, 2010), um indicador de desempenho modificado é usado para atribuir a cada solução de população um valor de aptidão com base em sua contribuição para a convergência e a diversidade do algoritmo. No entanto, algumas dificuldades nessas abordagens são apresentados em (HE; YEN, 2016), como 1) o valor de adequação atribuído por um único indicador pode fornecer apenas algumas informações específicas do desempenho, mas incompletas; 2) alguns indicadores de desempenho violam o princípio de domínio de Pareto e podem atribuir uma pontuação melhor ao indivíduo dominado do que ao indivíduo não dominado; 3) este método pode não resolver as dificuldades relacionadas aos espaços de busca de grande dimensão. Portanto, a seleção do indicador não é trivial e é crucial para o algoritmo. Na literatura, embora existam muitos indicadores, a métrica de Hipervolume é a mais

utilizada. O cálculo de HV é computacionalmente caro e cresce exponencialmente com o aumento do número de objetivos. Esse problema torna os DMOEAs baseados no HV muito lentos. Pode ser por conta disso que poucos algoritmos foram desenvolvidos nesta categoria.

Os algoritmos baseados em decomposição (FARINA; DEB; AMATO, 2004), (MEHNEN; WAGNER; RUDOLPH, 2006), (BINGUL, 2007), (ZHENG, 2007), (MURUGANANTHAM et al., 2013; MURUGANANTHAM; TAN; VADAKKEPAT, 2016), (LIU et al., 2015), (CHEN; ZHANG; ZENG, 2015), (LIU; LIU, 2016), (GEE; TAN; ALIPPI, 2017), (FU; SUN, 2017), (CHEN; LI; YAO, 2018), (CHEN et al., 2018), decompõem o problema em vários subproblemas de objetivo único por meio de uma agregação ponderada linear ou não linear de múltiplos objetivos. Vetores de ponderação ou direções de busca são predefinidos para cada subproblema. Esses vetores agregam todos os valores objetivos e direcionam a convergência para a verdadeira frente de Pareto. A diversidade populacional é mantida usando uma distribuição uniforme nos vetores ponderados ou pontos de referência. No entanto, vetores e subproblemas de peso bem distribuídos não garantem que suas soluções ótimas correspondentes também sejam bem distribuídas (TRIVEDI et al., 2017). Neste tipo abordagem, uma solução só pode ser recombinaada com outras dentro de sua vizinhança. O tamanho da vizinhança tem um impacto no desempenho do algoritmo. Uma escolha adequada do tamanho da vizinhança pode ser complicada. O desempenho dos algoritmos também é bastante dependente do método de agregação escolhido e do número de vetores de peso. Se o número for muito grande, o esforço computacional também será alto (TRIVEDI et al., 2017).

2.2.2 Operadores de Variação

Os DMOEAs usam uma grande variedade de operadores de variação que mudam de acordo com o domínio ou a representação do problema a ser resolvido. Assim, para otimização contínua, dentre os operadores mais utilizados, pode-se citar os operadores variacionais tradicionais, como os Algoritmos Genéticos (GA) e os operadores de Evolução Diferencial (DE). Alguns MOEAs também usam modelos probabilísticos para gerar descendentes.

Os operadores de DE compartilham propriedades favoráveis tais como simplicidade de implementação, robustez e eficiência, e auto-adaptação e versatilidade. Os operadores típicos são a mutação diferencial (STORN; PRICE, 1997) com taxa de mutação $F \in [0, 2]$ e recombinação binária com taxa de crossover $CR \in [0, 1]$. Entre MOEAs típicos adotando DE, pode-se citar (MEHNEN; WAGNER; RUDOLPH, 2006), (LIU; ZENG, 2012), (WAN; WANG, 2013), (GUERRERO-PEÑA; ARAÚJO, 2017). A mutação polinomial (DEB; AGRAWAL, 1999) com o parâmetro index (η_m) e cruzamento binário simulado (SBX) com o índice de distribuição η_c são dos operadores mais usados. Os seguintes estudos que usam esses operadores podem ser mencionados: (MEHNEN; WAGNER; RUDOLPH, 2006), (DEB; KARTHIK et al., 2007), (ROY; MEHNEN, 2008), (WANG; LI, 2009), (AZEVEDO; ARAÚJO, 2011b), (AZZOUZ; BECHIKH; Ben Said, 2014) (LI et al., 2014), (LIU et al., 2014), (AZZOUZ; BECHIKH;

SAID, 2015a), (AZZOUZ; BECHIKH; SAID, 2015b), (LIU et al., 2015), (CHEN; ZHANG; ZENG, 2015), (LIU; LIU, 2016), (JIANG; YANG, 2016), (SAHMOUD; TOPCUOGLU, 2016b), (JIANG et al., 2017), (FU; SUN, 2017), (CHEN; LI; YAO, 2018).

Nos últimos anos, um conjunto de DMOEAs mostrou as vantagens de usar operadores de DE em combinação com a mutação polinomial. Desta forma, o desempenho dos algoritmos tem melhorado, o que é evidenciado nos resultados obtidos nas simulações das diversas pesquisas. Entre os algoritmos mais significativos que adotam essa combinação de operadores, encontram-se (MURUGANANTHAM et al., 2013; MURUGANANTHAM; TAN; VADAKKEPAT, 2016), (WU; JIN; LIU, 2015). Os operadores típicos do GA, o *One-Point Crossover* e o *Bit-flip Mutation*, também são freqüentemente usados pelos pesquisadores, como pode ser visto em (BINGUL, 2007), (CÁMARA; ORTEGA; TORO, 2007), (GOH; TAN, 2009). Em (FARINA; DEB; AMATO, 2004), (LIU; WANG, 2007), (WANG; LI, 2010), (GARROZI; ARAUJO, 2011), operadores como a mutação de Gaussiana e o crossover aritmético também foram usados.

Outros estudos como (GARIS, 2006), (HATZAKIS; WALLACE, 2006), (WANG; DANG, 2008), (YUPING; HAILIN; LEUNG, 2000), (WANG; LI, 2010) propuseram novos operadores para criar prole, evidenciando que os operadores de variação influenciam no desempenho dos DMOEAs. Estudos recentes usam um modelo probabilístico como um operador de variação para gerar descendentes ou usá-lo para produzir uma nova população (KHALLED; TALUKDER; KIRLEY, 2008), (CHENG; JIN; NARUKAWA, 2015), (JIANG et al., 2018). Modelos tais como de regressão linear, não-linear ou polinomial, e processos Gaussianos, visam manter a diversidade, extraindo características probabilísticas do espaço de decisão durante a pesquisa. Um modelo pode ser estimado usando indivíduos potencialmente aptos (soluções candidatas) ou soluções com muita incerteza. Vários estudos, como (ZHOU et al., 2007), (ZHOU; JIN; ZHANG, 2014), (PENG; ZHENG; ZOU, 2014), (PENG et al., 2015), (JIANG et al., 2017), (RUAN et al., 2017), e (ZOU et al., 2017), usam o operador de variação proposto em *Regularity Model-based Multi-objective Estimation of Distribution Algorithm* (RM-MEDA) (ZHANG; ZHOU; JIN, 2008) para gerar soluções candidatas. O RM-MEDA modela uma região promissora no espaço de decisão com uma distribuição de probabilidade em cada geração. Um modelo é construído usando o algoritmo *Local Principal Component Analysis* e, em seguida, novos indivíduos são criados amostrando do modelo. Estudos como (GEE; TAN; ALIPPI, 2017), empregam os fundamentos do *Inverse Modeling-based Multi-Objective Evolutionary Algorithm* (IM-MOEA) (CHENG et al., 2015) para gerar a prole. O IM-MOEA divide o espaço objetivo em sub-regiões usando vetores de referência predefinidos em cada geração. Para cada sub-região, é construído um modelo inverso, que mapeia soluções candidatas no espaço objetivo para o espaço de decisão.

2.2.3 Mecanismo de Detecção da Mudança

O mecanismo de detecção de mudanças é um dos componentes essenciais para projetar um DMOEA, embora nem todos os algoritmos o usem. Esta etapa detecta se uma alteração do ambiente ocorre e, se for o caso, aciona-se uma estratégia para lidar com o novo cenário. Alguns métodos também revelam os graus de similaridade entre o problema atual e os anteriores. De acordo com (SAHMOUD; TOPCUOGLU, 2016b), os esquemas de Detecção de Mudanças são classificados em dois grandes grupos: 1) estratégias de detecção baseada em sensores; e 2) estratégias de detecção baseada na população.

O primeiro mecanismo usa a reavaliação de um número fixo de sensores (soluções) durante o processo de busca. (DEB; KARTHIK et al., 2007) propôs a seleção aleatória de algumas soluções da população e sua reavaliação. Se qualquer uma das funções objetivo ou de restrições for alterado, haverá uma mudança no problema. Alguns dos estudos que empregam esse mecanismo são (DEB; KARTHIK et al., 2007), (ZHENG, 2007), (ZHOU et al., 2007), (GOH; TAN, 2009), (CHEN; LI; CHEN, 2009), (WANG; LI, 2009), (WANG; LI, 2010), (KOO; GOH; TAN, 2010), (LIU, 2010), (GARROZI; ARAÚJO, 2011), (LIU; ZENG, 2012), (LI et al., 2014), (ZHOU; JIN; ZHANG, 2014), (PENG; ZHENG; ZOU, 2014), (WU; JIN; LIU, 2015), (AZZOUZ; BECHIKH; SAID, 2015a), (AZZOUZ; BECHIKH; SAID, 2015b), (MURUGANANTHAM et al., 2013; MURUGANANTHAM; TAN; VADAKKEPAT, 2016), (PENG et al., 2015), (GUERRERO-PEÑA; ARAÚJO, 2017), (RUAN et al., 2017), (ZOU et al., 2017), (CHEN et al., 2018). Este método pode ignorar alguns pontos fortemente influenciados pelo ambiente. Além disso, quando há indivíduos idênticos na população, os mesmos indivíduos podem ser selecionados, fazendo com que o impacto da paisagem em alguns pontos seja ignorado. Por esse motivo, em (LIU et al., 2015), alguns indivíduos uniformemente distribuídos na Frente de Pareto são selecionados para detectar a mudança do ambiente, equilibrando o impacto do novo cenário. Outras adaptações desse mecanismo foram propostas em (JIANG; YANG, 2016), (SAHMOUD; TOPCUOGLU, 2016b), (GEE; TAN; ALIPPI, 2017) para incrementar a eficiência da detecção da mudança.

A determinação de um esquema de posicionamento de sensores e o número de sensores necessários são os principais problemas de uma estratégia deste tipo. Esta abordagem é fácil de implementar, mas assume que não há ruído nas avaliações de função, ou seja, não é robusto. Assim, o método pode não ser adequado para detectar alterações em ambientes ruidosos, o que é uma característica de problemas reais como o controle de processos.

As estratégias baseadas na população detectam a mudança considerando as informações de comportamento da população e não requer avaliações de função adicionais. Os métodos deste grupo podem detectar mudanças em ambientes ruidosos; entretanto, os algoritmos podem precisar de alguns outros parâmetros dependentes de problemas (ZHOU; JIN; ZHANG, 2014). Além disso, esse mecanismo pode causar falsos positivos, induzindo os algoritmos a reagir de forma exagerada quando nenhuma alteração ocorre (JIANG; YANG, 2016). Assim, podem ser mencionados os trabalhos de (FARINA; DEB; AMATO, 2004) que

propôs um mecanismo baseado nas mudanças das funções objetivo. A maioria dos estudos existentes se concentra em saber se uma mudança ocorreu. Em (LIU et al., 2014), (LIU; LIU, 2016), os autores estimam a extensão da mudança ambiental, que é definida como o grau de desvio entre as populações antes e depois da mudança. Com base no grau de mudança, esta pesquisa apresenta um método para ajustar dinamicamente a proporção de introdução da diversidade, em vez de ser desenhada e fixada manualmente.

Como ambas as classes de métodos escolhem uma pequena proporção de membros da população como detectores, eles não podem garantir que as mudanças vão ser detectadas. Isso só ocorre se todos os indivíduos do espaço de solução factível forem selecionados como detectores, o que será computacionalmente caro. Portanto, um método de detecção adequado deve encontrar um equilíbrio entre a capacidade de detecção e a eficiência (JIANG; YANG, 2016).

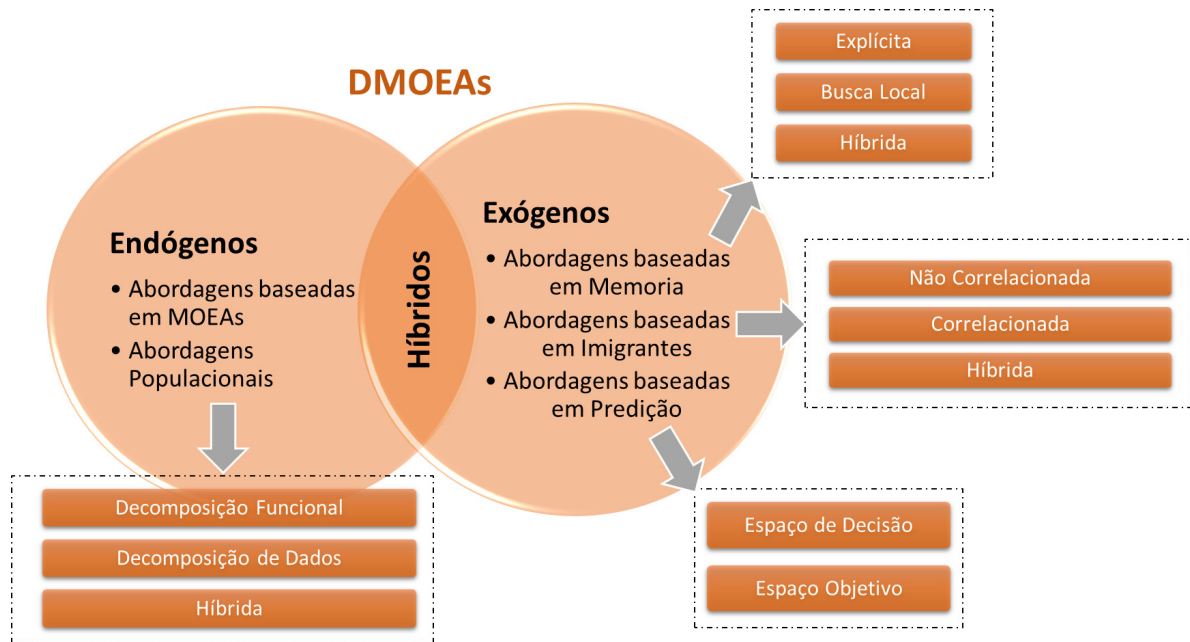
2.2.4 Mecanismo de Reação à Mudança (CRM)

Devido à complexidade do espaço de busca, a ocorrência de mudanças em um ambiente tende a mudar seus ótimos de lugar. Por outro lado, a probabilidade dos indivíduos estarem nas regiões que mudaram aumenta à medida que o tamanho da população se aproxima do total de pontos existentes no espaço de soluções. Contudo, torna-se inviável manter populações muito grande nas heurísticas. Portanto, após a ocorrência de uma mudança no ambiente (de maneira análoga, nos objetivos do problema) o EA deverá responder a estas mudanças. Os indivíduos provenientes do espaço de busca anterior podem permanecer sem mudanças ou serem adaptados após ocorrências de mudanças. Por exemplo, mudanças no ambiente podem causar mudanças na sua quantidade de genes.

Nos problemas dinâmicos, a ocorrência de modificações no ambiente ao longo do tempo acarreta a necessidade de aumentar a exploração e diminuir a exploração (COBB, 1990). O aumento da exploração pode ser realizado através da modificação dos parâmetros (BINGUL, 2007), da preservação da diversidade durante o processo de evolução, do emprego de várias populações, de abordagens baseadas em predição, de abordagens baseadas em memória, ou ainda, da inserção de novos indivíduos na população (ZHENG, 2007). Quando a exploração do EA aumenta, a diversidade genética também se eleva; portanto, as buscas devem se tornar mais eficientes visando atingir os novos ótimos que foram gerados pelas mudanças ocorridas recentemente no ambiente.

Vários estudos desenvolveram diferentes CRMs que aumentam a exploração preservando a diversidade ao longo do processo evolucionário, usando múltiplas populações, abordagens baseadas em previsão e técnicas baseadas em memória, ou até mesmo inserindo novos indivíduos na população (ZHENG, 2007). Considerando como os algoritmos dinâmicos aumentam a diversidade, uma nova taxonomia para os DMOEAs é proposta nesta tese (Figura 2). Algoritmos evolucionários multiobjetivos dinâmicos podem ser divididos em DMOEAs Endógenos, DMOEAs Exógenos e DMOEAs Híbridos.

Figura 2 – Nova Taxonomia dos DMOEAs



Fonte: Autor

Nos **DMOEAs Endógenos**, a diversidade é mantida durante todo o processo evolucionário (JIN; BRANKE, 2005). Em geral, estes DMOEAs não usam mecanismo de detecção de alterações, tornando-os adequados para problemas com mudanças difíceis de detectar. Um DMOEA endógeno pode ser classificado em abordagens dinâmicas baseadas em MOEAs e abordagens populacionais.

Abordagens dinâmicas baseadas em MOEAs são algoritmos que introduzem diversidade por meio da aplicação de operadores de variação (VO) ou busca local (LS), criando novos indivíduos em regiões não exploradas, podendo incluir as regiões onde ocorreram mudanças recentes. A principal diferença em ambos os métodos é que, no segundo, uma dada condição atendida aciona um método de busca local. A presente abordagem tende a apresentar melhor desempenho quando há um grau significativo de similaridade entre o antigo e o novo cenário. Nesta categoria, pode-se mencionar os seguintes DMOEAs: (GARIS, 2006), (BINGUL, 2007), (ROY; MEHNEN, 2008), (KHALED; TALUKDER; KIRLEY, 2008), (AZZOUZ; BECHIKH; Ben Said, 2014), (GEE; TAN; ALIPPI, 2017).

Abordagens Populacionais se caracterizam por dividir uma população em subpopulações que exploram diferentes regiões na paisagem de aptidão. As diferentes subpopulações contêm informações sobre diferentes áreas promissoras do espaço de busca. Buscam-se diferentes soluções ótimas para cada subpopulação. No processo de seleção, os indivíduos são escolhidos de acordo com informações globais considerando todas as subpopulações. O processamento paralelo permite acelerar a reação do algoritmo às mudanças, reduzindo o tempo de processamento e alcançando um conjunto de soluções não dominadas próximas à Frente de Pareto. Assim, DMOPs com uma taxa rápida de mudança também podem ser resolvidos (CÁMARA; ORTEGA; TORO, 2007).

Estas abordagens podem ser particionadas em (CÁMARA; ORTEGA; TORO, 2007): decomposição funcional, decomposição de dados e métodos híbridos nos quais cada processador avalia um subconjunto de funções objetivo para um subconjunto da população. Os métodos de decomposição funcional (WANG; DANG, 2008) executam tarefas separadamente de maneira paralela. Os métodos de decomposição de dados executam tarefas diferentes em dados diferentes, isto é, os indivíduos da população. Entre os métodos de decomposição de dados, vários esquemas de paralelização foram desenvolvidos, baseados na estrutura mestre-trabalhador (*master-worker*) e no modelo de ilha (CÁMARA; ORTEGA; TORO, 2007). Nos métodos híbridos (GOH; TAN, 2009) combinam mecanismos competitivos e cooperativos observados na natureza. Cada subpopulação de espécies competirá para representar um determinado subcomponente de problema de múltiplos objetivos, enquanto os vencedores irão cooperar para desenvolver as melhores soluções.

Nos **DMOEAs Exógenos**, ações explícitas são tomadas para aumentar a diversidade em resposta à mudança do ambiente. Os aumentos de exploração são gerados externamente à população por transferência de indivíduos de outras populações, imigrantes, repositórios ou ilhas. Neste caso, a população passa a agregar o material genético de indivíduos externos à população considerada durante o processo de evolução e, portanto, o aumento da exploração não é realizado pelos operadores de variação. O controle exógeno de diversidade é realizado por abordagens baseadas em memória, abordagens baseadas em imigrantes e abordagens baseadas em predição.

Abordagens Baseadas em Memória: usam um arquivo para armazenar as melhores soluções passadas que serão empregadas caso o ótimo retorne aos locais anteriores. Quando certos aspectos do problema demonstram algum tipo de comportamento periódico, estas soluções podem ser utilizadas para guiar as buscas na sua vizinhança e reduzir o tempo computacional (YOUNES, 2006). No entanto, a memória usada para armazenar soluções anteriores de alta qualidade deve ser usada com cuidado, pois também pode ter o efeito adverso de buscar soluções em regiões não promissoras ao invés de explorar novas regiões. Isto foi confirmado por (BRANKE; BRANKE, 1999) que concluiu que a memória não pode atuar sozinha, deve ser combinada com um mecanismo de diversificação para continuar explorando o espaço de busca por novas áreas promissórias.

As abordagens de memória podem ser classificadas como memória explícita, memória de busca local e memória híbrida. Os esquemas de memória explícita (CHEN; LI; CHEN, 2009; CHEN; ZHANG; ZENG, 2015), selecionam aleatoriamente uma solução da memória ou soluções geradas aleatoriamente. A principal limitação deste esquema é que ele pode induzir a erro na busca quando o passo da mudança não é pequeno o suficiente. As abordagens de memória de Busca Local (WANG; LI, 2010) também selecionam aleatoriamente uma solução da memória ou soluções geradas aleatoriamente. Cada solução escolhida do arquivo fornece uma solução de vizinhança na nova população, por meio de uma estratégia de busca local Gaussiana com média de 0 e desvio padrão σ . No entanto, um valor

σ grande pode causar a perda das informações anteriores, enquanto um σ pequeno pode não lidar com um passo de mudança grande. Finalmente, o esquema de memória híbrida mistura esquemas de memória explícita e de busca local, para superar as limitações das duas abordagens anteriores. A estratégia seleciona algumas soluções de memória explícita com probabilidade p_m , escolhe soluções de memória de busca local com probabilidade p_e , ou então gera uma proporção de soluções aleatórias (WANG; LI, 2009).

Abordagens baseadas em Imigrantes: Neste modelo exógeno, a exploração do ambiente é realizada com o auxílio de novos indivíduos aleatórios que são gerados externamente ao EA e inseridos na população. Em (AZEVEDO; ARAÚJO, 2011b), uma taxonomia estendida para abordagens baseadas em imigrantes foi proposta. De acordo com a dependência dos imigrantes (conjunto Y) com a população em evolução (conjunto Z), essas abordagens podem ser divididas em: (1) não correlacionadas quando Y e Z são estatisticamente independentes; (2) correlacionada quando Y apresenta dependências estatísticas em relação a Z; e (3) híbrido.

Esquemas de imigração não correlacionados (GREFENSTETTE et al., 1992; LIU; WANG, 2007; DEB; KARTHIK et al., 2007), utilizam fortemente a distribuição uniforme. Esquemas de imigração correlacionados (FARINA; DEB; AMATO, 2004; DEB; KARTHIK et al., 2007; LIU; ZENG, 2012; AZZOUZ; BECHIKH; SAID, 2015b; ZHENG, 2007; GARROZI; ARAUJO, 2011) podem gerar imigrantes por mutações em soluções anteriores. Normalmente, os indivíduos mais aptos do instante precedente de otimização são usados na geração de imigrantes. Esquemas de Imigração Híbridos (WAN; WANG, 2013; AZZOUZ; BECHIKH; SAID, 2015b; AZZOUZ; BECHIKH; SAID, 2015a; LIU; LIU, 2016; CHEN; LI; YAO, 2018) adicionam subpopulações de imigrantes ao modelo, cada um deles com uma taxa de imigração associada. O esquema híbrido requer pelo menos duas subpopulações de imigrantes com e sem a presença de dependências estatísticas sobre Z. As estratégias de imigração aumentam a diversidade da população sem uma orientação correta, e o algoritmo pode não convergir ao lidar com DMOPs mais complexos (PENG; ZHENG; ZOU, 2014).

Abordagens baseadas em Predição: Estas abordagens tentam aprender padrões a partir da experiência de buscas passadas para prever as mudanças no futuro com base nesses padrões. Como rastrear e prever as mudanças na paisagem é a questão que muitos estudos atuais tentam abordar. Idealmente, um modelo de previsão poderia prever a localização exata da fronteira de Pareto, dos melhores locais para a reinicialização dos indivíduos quando ocorre uma mudança no ambiente, prever o momento em que a próxima mudança ocorrerá e como deve ser o ambiente novo (HATZAKIS; WALLACE, 2006). Os métodos baseados em previsão fornecem orientação para a evolução da população por meio de um mecanismo de previsão que ajuda o algoritmo a responder rapidamente às novas mudanças (PENG et al., 2015). Dada uma estimativa da próxima localização do POS/POF, esses métodos podem criar um conjunto de indivíduos dentro e próximo dessa estimativa, para antecipar a mudança no ambiente. Dependendo da qualidade da previsão, os indivíduos

podem ajudar na descoberta da região promissora (HATZAKIS; WALLACE, 2006).

A maioria das abordagens (HATZAKIS; WALLACE, 2006; ZHOU et al., 2007; KOO; GOH; TAN, 2010; LIU, 2010; ZHOU; JIN; ZHANG, 2014; PENG; ZHENG; ZOU, 2014; LI et al., 2014; MURUGANANTHAM et al., 2013; MURUGANANTHAM; TAN; VADAKKEPAT, 2016; LIU et al., 2015; PENG et al., 2015; RUAN et al., 2017; FU; SUN, 2017; ZOU et al., 2017; CHEN et al., 2018; JIANG et al., 2018) fazem a previsão no espaço de decisão, para estimar a próxima localização do conjunto ótimo de Pareto. A abordagem mais simples emprega um único indivíduo em cada local previsto. Outros métodos podem incluir os pontos extremos da frente de Pareto ou outras soluções além desses extremos e rastreá-los. Estratégias mais elaboradas, dividem a frente em hipercubos de indivíduos, podem fornecer informações sobre a cobertura do conjunto previsto, o erro e as margens de confiança da previsão (HATZAKIS; WALLACE, 2006). Por outro lado, muito poucos algoritmos (JIANG et al., 2017) fizeram a previsão da POF para criar um conjunto de previsão.

Métodos baseados no modelo de predição parecem promissores para acelerar a convergência do algoritmo. No entanto, a previsão pode não ser bem-sucedida, pois pode não haver previsibilidade no comportamento dinâmico da paisagem ou o padrão pode não ser identificável pelo método de previsão. Se a previsão não for precisa, ela pode encaminhar a população para regiões não promissoras dificultando a convergência do algoritmo. (HATZAKIS; WALLACE, 2006) definiram duas fontes primárias de erro: 1) Precisão do histórico da solução ótima; 2) Precisão do modelo de previsão. O primeiro erro é produzido devido à não convergência correta do EA, o que faz com que as séries temporais das melhores soluções encontradas não coincidam com as séries temporais do POF/POS real. O segundo erro é produzido por uma previsão imprecisa e depende da natureza do problema e do tipo e qualidade do modelo de previsão. Até agora, como projetar um modelo de previsão mais preciso ainda é a principal dificuldade para os pesquisadores.

Os **DMOEAs Híbridos** levam em consideração as vantagens das duas classificações de CRM mencionadas acima. Assim, novos algoritmos misturaram as formas endógena e exógena para manter a diversidade através do processo evolucionário, bem como aumentá-la quando ocorre uma mudança. Pode-se mencionar os seguintes trabalhos nesta categoria: (AZEVEDO; ARAÚJO, 2011b; WU; JIN; LIU, 2015; JIANG; YANG, 2016; GUERRERO-PEÑA; ARAÚJO, 2017).

Para avaliar o desempenho dos DMOEAs na resolução de DMOPs, são empregadas medidas e métricas de desempenho. Na próxima seção definem-se as mesmas, bem como os problemas que podem surgir com a seleção errada deles para a avaliação do DMOEA ou de comparar uns com outros.

2.3 INDICADORES DE DESEMPENHO

A definição dos indicadores é essencial para a validação de um algoritmo e para compará-lo com outros (COELLO, 2005). Contudo, ao lidar com um algoritmo multiobjetivo, a verificação qualitativa dos resultados torna-se difícil. Devido ao fato dos algoritmos evolucionários gerarem diversas soluções de compromisso e de sua natureza estocástica; diversas execuções serão necessárias para avaliar a sua performance e, portanto, os resultados devem ser validados através de ferramentas de análise estatística.

Muitas das medidas e métricas usadas para avaliar o desempenho dos MOEAs, foram adaptadas para avaliar os DMOEAs. (HELBIG; ENGELBRECHT, 2013b) classificam as medidas de desempenho dinâmicas de múltiplos objetivos em:

Medidas de Precisão: Estas medidas avaliam a convergência do algoritmo, a precisão do algoritmo no rastreamento da *POF/POS* em DMOPs. As medidas de precisão podem estar baseadas 1) na distância entre a frente ótima aproximada (POF^*) e a verdadeira *POF* (POF'), 2) na percentagem ou razão de soluções ótimas encontradas que pertencem ao *POF/POS* real.

Medidas de Diversidade: As medidas de diversidade avaliam quão espalhadas estão as soluções da POF^* ou sua extensão.

Medidas Combinadas: Estas medidas podem comparar os algoritmos simultaneamente no sentido da convergência e da diversidade.

Medidas de Robustez: As medidas de robustez avaliam a capacidade de recuperação de um algoritmo depois que as mudanças no ambiente ocorrem.

Geralmente as medidas e métricas são calculadas para todas as mudanças que ocorrem no ambiente. Para dar o valor da avaliação do DMOEA, é calculado o indicador como sua média ($\bar{\theta}$) no tempo:

$$\bar{\theta} = \frac{1}{num_{change}} \sum_{i=1}^{num_{change}} \theta_i \quad (2.5)$$

onde θ é o indicador de desempenho usado (medida ou métrica), num_{change} é o número de mudanças que ocorreram no ambiente, e θ_i é o valor de θ calculado para a geração imediatamente antes que a $(i + 1)$ -ésima mudança ocorra.

A Tabela 3 mostra um resumo das métricas e medidas mais usadas. O sentido da seta na última coluna da tabela indica quando é melhor o indicador. Dentre do grupo de medidas, as mais utilizadas pelos pesquisadores pertencem ao grupo de medidas de precisão, e neste conjunto, a mais utilizada é medida de Distância Geracional Invertida (IGD). Nas medidas de diversidade, a mais usada é o Espalhamento Máximo (MS), e nas medidas combinadas são mais aplicadas o Hipervolume (HV) e a Razão de Hipervolume (HVR).

Tabela 3 – Indicadores de desempenho.

Medidas de Precisão		
Indicador	Equação	
Medida de Desempenho de Convergência (e_x , e_f) (FARINA; DEB; AMATO, 2004)	$e_x = \frac{1}{N} \sum_{j=1}^N \min_{i=1: POF' } \left\ \frac{POF'_i(t) - POF_i^*(t)}{R(t) - U(t)} \right\ $ $e_f = \frac{1}{N} \sum_{j=1}^N \min_{i=1: POS' } \ POS'_i(t) - POS_i^*(t)\ $ onde $U(t)$ é o ponto utópico e $R(t)$ é o ponto Nadir dependente do tempo (MIETTINEN, 2012)	↓
Indicador de Desempenho baseado em Distância ($D(P)$) (ZHANG; ZHOU; JIN, 2006)	$D(P) = \frac{1}{ POF^* } \sum_{x \in POF^*} \ x - y(x)\ _2$ $y(x) = \arg \min_{y \in POF'} \ x - y\ _2$	↓
Taxa de Sucesso (SC_τ) (MEHNEN; WAGNER; RUDOLPH, 2006)	$SC_\tau = \frac{ \{x f(x) \in POF'\} }{ POF^* }$	↑
Distância Geracional (GD) (VELDHUIZEN; A.; LAMONT, 1998)	$GD_t = \sqrt{\frac{\sum_{i=1}^N d_i^2}{N}}$ onde d_i são as distâncias Euclidianas mínimas do valor dos objetivos de uma solução para a POF	↓
Relação Convergente (CR) (ZHANG, 2008)	$CR = \frac{1}{KT} \sum_{i=1}^T \sum_{j=1}^{K-1} \frac{1}{K-j} \sum_{l=j+1}^K C(A_{ij}, A_{il})$ onde K é o número de vezes que um algoritmo A é executado no i-ésimo ambiente, com $1 \leq i \leq T$, resultando em K conjuntos de soluções não dominadas $\{A_{ij}\}_{j \geq 1}^K$.	↓
Distância Geracional do Espaço Variável (VD) (GOH; TAN, 2009)	$VD(t) = \sqrt{\frac{N \sum_{i=1}^N d_i^2}{N}}$ onde d_i é a distância Euclidiana entre a i-ésima solução do POS^* e o membro mais próximo do POS'.	↓
Métrica de Convergência (λ) (CHEN; LI; CHEN, 2009)	$\lambda = \frac{1}{N} \sum_{j=1}^N \min \ POF^* - POF'\ $	↓
Distância Média de Hausdorff (Δ_p) (SCHÜTZE et al., 2012)	$\Delta_p(X, Y) = \max\{GD_p(X, Y), IGD_p(X, Y)\} = \max\left(\left(\frac{1}{N} \sum_{i=1}^N dist(x_i, Y)^p\right)^{\frac{1}{p}}, \left(\frac{1}{M} \sum_{i=1}^M dist(y_i, X)^p\right)^{\frac{1}{p}}\right)$ onde $X = x_1, \dots, x_N$, $Y = y_1, \dots, y_M \subset \mathbb{R}^m$ são conjuntos finitos não vazios, GD_p e IGD_p são as médias das métricas GD e IGD, respectivamente, usando a norma p (p-norm).	↓

Medidas de Diversidade		
Indicador	Equação	
Métrica de Espaçamento de Deb (Δ) (DEB et al., 2002)	$\Delta = \frac{\sum_{k=1}^m d_k^e + \sum_{i=1}^N d_i - \bar{d} }{\sum_{k=1}^m d_k^e + Nd}$ onde d_i é qualquer distância entre soluções vizinhas, $\bar{d}$ é a distância média e d_k^e é a distância entre as soluções extremas de POF^* e POF'. É usado apenas para problemas bi-objetivos.	↓
Métrica PL (PL) (MEHNEN; WAGNER; RUDOLPH, 2006)	$PL := \frac{\sum_{f(x_i) \in POF'} \ln(\xi_{x_i})}{L_{POF'}}$ onde $\xi = L(\gamma, f(x_i), f(x_{i+1})) + 1$ é cada subseção entre os pontos vizinhos dispostos na frente de Pareto real, adicionando 1 para garantir que novas soluções aumentem o valor da métrica. $L(\gamma, a, b) := \int_a^b \dot{\gamma} dt = \int_a^b \sqrt{\dot{\gamma}_1^2 + \dots + \dot{\gamma}_m^2}$ é o tamanho do caminho entre duas soluções $[a, b]$. $\gamma(t) : \subseteq \mathbb{R} \rightarrow \mathbb{R}^m$ é uma função paramétrica contínua.	↑
Densidade Média (AD) (ZHANG, 2008)	$AD = \frac{1}{KT} \sum_{i=1}^T \sum_{j=1}^K \sqrt{\frac{1}{N-1} \sum_{l=1}^N (\bar{d}_{ij} - d_{ijl})^2}$ com $d_{ijl} = \min_{r \neq l, 1 \leq r \leq N} \{\ \mathbf{x}_l - \mathbf{x}_r\ \mathbf{x}_l, \mathbf{x}_r \in X_{ij}\}$, e $\bar{d}_{ij} = \frac{1}{N} \sum_{i=1}^N d_{ijl}$	↓
Métrica C (C) (WANG; DANG, 2008)	$C(A_i, B_i) = \frac{ \{b \in B_i \exists a \in A_i : b \succeq a\} }{ B_i }$ onde (A_i, B_i) são dois conjuntos de soluções não dominadas encontradas por dois algoritmos. Se $C(A_i, B_i) = 1$, B_i é fracamente dominado por A_i. Se $C(A_i, B_i) > C(B_i, A_i)$, o algoritmo A tem melhor desempenho do que o algoritmo B. O operador de dominância não é simétrico. Para analisar quão perto está a POF^* da POF', presume-se que o conjunto A é formado pelas soluções de POF', e o conjunto B é formado a partir do POF^*.	↑
Taxa de Cobertura (Co) (ZHANG, 2008)	$Co(A, B) = \frac{1}{4TK^2} \sum_{i=1}^T \sum_{j=1}^K \sum_{r=1}^K C(A_{ij}, B_{ir})$ Se $Co(A, B) > Co(B, A)$, o algoritmo A é melhor do que o B. Para analisar quão perto está a POF^* da POF', presume-se que o conjunto A é formado pelas soluções de POF', e o conjunto B é formado a partir do POF^*.	↑

Medida U (U) (WANG; DANG, 2008)	$U_m(POF^*) = d_{std} = \sqrt{\frac{1}{2mN-1} \sum_{g \in POF^*} \sum_{r=1}^{2m} (d_r - d_{mean})^2}$ com $d_{mean} = \frac{1}{2mN} \sum_{g \in POF^*} \sum_{r=1}^m (d_{2r-1} - d_{2r})$ onde $r \in \{1, 2, \dots, m\}$ é usado para classificar o POF^* em subconjuntos $POF_{1r}^*; POF_{2r}^*; \dots; POF_{kr}^*$ de acordo com o r -ésimo valor objetivo dos pontos em POF^* . $d_r = \frac{1}{2 T - \{T_l^r, T_k^r\} } \sum_{g \in T - \{T_l^r, T_k^r\}} (d_{2r-1} - d_{2r})$, para qualquer $g \in T_i^r \subset T$, $d_{2r-1} = d_{2r-1} + d_r$ se $i = 1$ e $d_{2r} = d_{2r} + d_r$, se $i = k$. d_{2r-1} e d_{2r} são as distâncias de dois vizinhos de um ponto $g \in POF^*$ ao ponto g .	↓
Métrica de Diversidade γ (γ) (CHEN; LI; CHEN, 2009)	$\gamma = \frac{1}{NP} \sum_{j=1}^{NP} \frac{1}{L} \sum_{l=1}^L \left[- \sum_{k=1}^S P_{lk} \log(P_{lk}) \right]$ onde NP é o tamanho da população, L é o comprimento do cromossomo, S é a cardinalidade dos alelos genotípicos, e P_{lk} é de k -ésimo allele genotype rate in the l -ésimo location.	↑
Espalhamento Máximo (MS) (ZITZLER, 1999)	$MS(t) = \sqrt{\sum_{k=1}^m (\overline{POF_k^*(t)} - \underline{POF_k^*(t)})^2}$ onde $\overline{POF_k^*(t)}$ e $\underline{POF_k^*(t)}$ são os valores máximo e mínimo do k -ésimo objetivo em POF^* , respectivamente.	↑
Adaptação do Espalhamento Máximo (MS') (GOH; TAN, 2009)	$MS'(t) = \sqrt{\frac{1}{m} \sum_{k=1}^m \left[\frac{ms(t)}{\overline{POF_k^*(t)} - \underline{POF_k^*(t)}} \right]^2}$ $ms(t) = \min \left[\overline{POF_k^*(t)}, \overline{POF_k^*(t)} \right] - \max \left[\underline{POF_k^*(t)}, \underline{POF_k^*(t)} \right]$	↑
Medida do escopo de cobertura (CS) (ZHANG; QIAN, 2011)	$CS = \frac{1}{N} \sum_{i=1}^N \max \{ \ f(\mathbf{x}_i) - f(\mathbf{x}_j) \ \}$ $\mathbf{x}_i, \mathbf{x}_j \in POF^*, i \geq 1$ and $j \leq N$	↑
Métrica de Espaçamento de Schott (S) (SCHOTT, 1995)	$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\bar{d} - d_i)^2}$ onde $d_i = \min_{j=1, \dots, N} \{ \sum_{k=1}^m f_{ki}(\mathbf{x}) - f_{kj}(\mathbf{x}) \}$ e $\bar{d}$ é a média de todos os valores d_i .	↓

Medidas Combinadas

Indicador	Equação	
Hipervolume (HV) (ZITZLER; BROCKHOFF; THIELE, 2007; ZITZLER; THIELE, 1998)	$HV = Leb \bigcup_{x \in POF} [f_1(x), r_1] \times \dots \times [f_m(x), r_m]$ $r = (r_1, r_2, \dots, r_m)$ é um vetor de referência. Esta métrica é usada principalmente quando o POF' é desconhecido.	↑
Relação de Hipervolume (HVR) (VELDHUIZEN, 1999)	$HVR(t) = \frac{HV(POF^*(t))}{HV(POF'(t))}$	↑

Diferença de Hipervolume (HVD) (FONSECA et al., 2005)	$HVD = HV(POF') - HV(POF^*)$	↓
Distância Geracional Invertida (IGD) (WANG; LI, 2009)	$IGD_j = \frac{\sum_{i=1}^N d_i}{N}$ com $d_i = \min_{k=1}^{ POF' } \sqrt{\sum_{j=1}^m (f_j^{*(i)} - f_j'^{(k)})^2}$	↓
Versão melhorada do IGD ($IGD+$) (ISHIBUCHI; MASUDA; NOJIMA, 2015)	$IGD_j = \frac{\sum_{i=1}^N d_i}{N}$ com $d_i = \min_{k=1}^{ POF' } \sqrt{\sum_{j=1}^m (\max\{f_j^{*(i)} - f_j'^{(k)}, 0\})^2}$	↓
Medida de Precisão (acc) (CÁMARA, 2010)	$acc(t) = \frac{HV_{max}(t)}{HV(POF^*(t))}$	↑
Medida de Precisão Alternativa (acc_{alt}) (CÁMARA, 2010)	$acc_{alt}(t) = HV(POF'(t)) - HV(POF^*(t)) $	↑
Métrica de cobertura do Conjunto (η) (GUAN; CHEN; MO, 2005)	$\eta = \frac{D(POF^*, POF')}{HV(POF')} + \frac{D(POF', POF^*)}{HV(POF')}$ onde $D(POF^*, POF') = HV(POF^* + POF') - HV(POF')$ é a métrica D (ZITZLER, 1999).	↓

Medidas de Robustez

Indicador	Equação	
Estabilidade ($stab$) (CÁMARA, 2010)	$stab(t) = \max\{0, acc(t-1) - acc(t)\}$ É usado para quantificar os efeitos das mudanças ambientais na precisão do algoritmo (acc).	↓
Reatividade ($react$) (CÁMARA, 2010)	$react(t, \epsilon) = \min \{t' - t t < t' < \tau_{max}, t' \in \mathbb{N}, \frac{acc(t')}{acc(t)} \geq (1 - \epsilon)\}$ onde τ_{max} é o número máximo de iterações. Este indicador calcula quanto tempo leva para um algoritmo se recuperar após uma mudança de ambiente.	

Da pesquisa bibliográfica do tema, é percebido que ainda não existe um conjunto padrão de medidas para avaliar os DMOEAs. Ao avaliar e comparar o desempenho de um algoritmo com outros, é importante escolher os indicadores de desempenho adequados.

2.3.1 Limitações dos Indicadores de Desempenho

Em (HELBIG; ENGELBRECHT, 2013b; HELBIG; ENGELBRECHT, 2013a) são identificadas limitações de uso de medidas e métricas de desempenho para avaliar DMOEAs, quando (1) o algoritmo perde o rastro da POF ; (2) existem *outliers* nas soluções da POF encontrada; (3) há violações das restrições; (4) as medidas de desempenho são calculadas no espaço de decisão; (5) são comparados vários algoritmos.

Quando o algoritmo perde o rastro da POF , as medidas baseadas no Hipervolume podem dar resultados errados. Por isso, (HELBIG; ENGELBRECHT, 2013b; HELBIG; ENGELBRECHT, 2013a) propõem o uso da *alternative accuracy measure* (acc_{alt}) desenvolvida em (CÁMARA, 2010). A desvantagem de usar acc_{alt} , é que só pode ser utilizada quando a verdadeira POF é conhecida. Para frente de Pareto desconhecida, (HELBIG; ENGELBRECHT, 2013b; HELBIG; ENGELBRECHT, 2013a) propõem calcular medidas baseadas no HV .

Os *outliers* da POF , isto é, as soluções fora da frente de Pareto encontradas na última iteração do algoritmo, distorcem os valores das medidas baseadas em distâncias, de espalhamento das soluções e de Hipervolume. Com o aumento dos objetivos ou com as mudanças no ambiente, mais soluções *outlier* podem estar presente na POF . (HELBIG; ENGELBRECHT, 2013b) propõem duas alternativas para lidar com este problema: (1) Remover os *outliers*, processo complexo por não existir consenso para identificar uma solução não-dominada como *outlier*; (2) empregar a medida de distância de Hansdorff (SCHÜTZE et al., 2012), onde a influência dos *outliers* pode ser minimizada com a escolha do parâmetro p que indica a norma que vai se empregar no cálculo da distância.

A maioria dos DMOPs sem restrições têm limites no espaço de busca. Contudo, se o algoritmo não leva em conta as restrições, os limites podem ser violados, podendo aparecer soluções não factíveis na POF^* . Isto pode levar à remoção das melhores soluções da POF^* e a aparição de valores errados nas medidas de desempenho.

As medidas baseadas em distância podem ser calculadas no espaço de decisão ou no espaço objetivo. No espaço de decisão, as métricas medem a distância entre o POS' e POS^* , isto é, a precisão das soluções encontradas. No espaço objetivo as medidas calculam a distância entre POF' e POF^* . Acontece que pequenas mudanças na POS podem causar grandes mudanças na POF , isto é, pode ocorrer que o POS^* esteja perto do POS' , mas a POF^* esteja muito longe da POF' . Em (HELBIG; ENGELBRECHT, 2013b) aponta-se que medidas calculadas no espaço de decisão são apropriadas só para DMOPs de Tipo I, onde o POS muda com o tempo, mas a POF permanece estática.

Quando se comparam algoritmos dinâmicos multiobjetivos, eles podem responder bem a alguns indicadores de desempenho e o valor médio da medida pode não fornecer informações com relação ao desempenho do algoritmo para os vários ambientes durante a execução. Usualmente, o método empregado para comparar algoritmos, é fazer o *rank* deles levando em conta o desempenho com a medida utilizada. Contudo, este método é ineficiente, pois se for escolhida uma medida de desempenho inadequada para um problema, os resultados são incorretamente ordenados. (HELBIG; ENGELBRECHT, 2013b; HELBIG; ENGELBRECHT, 2013a) propõem uma abordagem de ganho-perda para lidar com este problema. Para cada instante de tempo, imediatamente antes que a mudança ocorra no ambiente, o indicador de desempenho é calculado. Então, quando vão ser comparados vários DMOEAs, o seguinte processo é feito:

1. O teste de Kruskal-Wallis (MCKIGHT; NAJAB, 2010) é aplicado para saber se há

diferenças estatísticas significativas no desempenho dos algoritmos.

2. Se existirem diferenças significativas, para cada par de DMOEAs:

- a) É aplicado Mann-Whitney (MCKNIGHT; NAJAB, 2010) para saber quais são, dois a dois, os algoritmos com diferenças estatísticas significativas.
- b) Se existir diferenças significativas entre par de algoritmos, para cada instante de tempo imediatamente antes da mudança do ambiente, é comparada a média do indicador de desempenho. Ganha aquele DMOEA com melhor valor no indicador de desempenho.

3. Para cada algoritmo, é calculada a diferença entre os ganhos e as perdas ($Diff = \#ganho - \#perdas$), determinados como acima.

4. É feito o *Rank* dos algoritmos baseados na *Diff*.

O algoritmo com melhor desempenho, tem o maior valor da *Diff*, e aquele com pior desempenho tem a *Diff* menor.

2.4 CONCLUSÕES DO CAPÍTULO

A pesquisa feita produziu muitas evidências de que os EAs podem ser aplicados a problemas dinâmicos. Um número significativo de DMOEAs foi desenvolvido para solucionar os DMOPs. Um DMOEA tem que lidar com as dificuldades relacionadas ao espaço de busca, como a paisagem descontínua, a forma da POF e do POS, entre outras. Portanto, um DMOEA bem-sucedido deve lidar com as dificuldades mencionadas para convergir antes que um ambiente sofra novas alterações que possam invalidar o processo evolucionário simulado.

A convergência rápida pode causar a perda de diversidade durante o processo de otimização, dificultando o rastreamento da POF. Assim, se uma paisagem muda, um novo problema é estabelecido; consequentemente, a diversidade é essencial para encontrar novas soluções. Por esta razão, um DMOEA deve introduzir um certo grau de diversidade através de um mecanismo de reação à mudança para lidar com o novo ambiente após sua mudança. Nos últimos anos, um número de DMOEAs foi desenvolvido para resolver os DMOPs, melhorando este mecanismo.

Os métodos exógenos parecem ser promissores para aumentar a diversidade e portanto, a velocidade da convergência, isto é, diminuir o número de gerações que o algoritmo demora para determinar soluções de qualidade. Entre eles, os métodos baseados em previsão têm se mostrado promissores em acelerar a convergência do algoritmo. Eles fornecem orientação para a evolução da população por meio de um mecanismo de previsão que ajuda o algoritmo a responder rapidamente às novas mudanças. No entanto, várias questões devem ser levadas em conta na aplicação destes algoritmos, tais como maneiras

de incorporar técnicas de previsão, funções de teste que serão usadas, e escolher a medida de desempenho adequada.

Do estudo realizado pode-se perceber uma tendência a usar mais os problemas FDA e dMOP do que outros com características mais complexas. Não existe um conjunto de problemas padrões definidos para avaliar este tipo de algoritmo. A análise também mostra que apesar de existirem muitos DMOPs do mundo real, muito poucos DMOEAs foram avaliados sobre este tipo de problema. Da análise dos indicadores de desempenho atuais, a mais utilizada por pesquisadores pertencem ao grupo de medidas de precisão, aquelas que são baseados na distância entre a frente de Pareto aproximada e a verdadeira. A tendência é usar a medida Distância Geracional Invertida *IGD*. No caso das medidas de diversidade, a mais usada é o Espalhamento Máximo *MS*, e nas medidas combinadas são mais aplicadas o Hipervolume (*HV*) e a Razão do Hipervolume (*HVR*) do que as outras. Pode-se notar que não há, ainda, medidas padrões definidas para avaliar os DMOEAs. Mostrou-se também que, na avaliação do desempenho do algoritmo e a comparação deles, é necessária a escolha certa do indicador de desempenho.

Algumas hipóteses podem ser levantadas acerca do problema e do algoritmo evolucionário utilizado para otimizá-lo. A pesquisa sobre aprendizagem de máquina estatística aborda este desafio mediante o desenvolvimento da área de modelagem probabilística, um campo que oferece uma abordagem formal de desenvolver novas metodologias para análise de dados (MURPHY, 2012). No próximo capítulo este tema vai ser abordado em detalhes.

3 MODELAGEM PROBABILÍSTICA

Novas possibilidades de pesquisa para os EAs podem ser realizadas através do análise de dados, considerando um outro tipo de conhecimento ou extraindo o comportamento probabilístico da população e da prole gerada, com representações dos indivíduos no espaço probabilístico. Esta experiência pode ser relevante para compreender a dinâmica dos algoritmos nas regiões de exploração e de exploração, para melhorar a qualidade das soluções encontradas pelos EAs, assim como para acelerar a convergência dos mesmos.

Problemas de análise de dados podem ser resolvidos utilizando modelos de variáveis latentes (LVM), que pressupõem que um complexo conjunto de dados observados exibe padrões mais simples, mas não observados (BLEI, 2014). A investigação em aprendizagem de máquina estatística aborda este desafio na área de modelagem probabilística, que oferece uma abordagem formal para desenvolver novas metodologias de análise de dados ((PEARL, 1988), (JORDAN, 1999), (BISHOP, 2007), (KOLLER; FRIEDMAN, 2009), (MURPHY, 2012)).

Na modelagem probabilística são usadas variáveis ocultas para codificar a estrutura oculta em dados observados. É articulada a relação entre as variáveis escondidas e as observadas com uma distribuição de probabilidade fatorizada, isto é, um modelo gráfico. Em particular, os modelos gráficos probabilísticos proveem uma linguagem visual para expressar suposições sobre os dados e sua estrutura escondida.

A modelagem probabilística usa algoritmos de inferência para estimar a distribuição posterior: a distribuição condicional da estrutura oculta dadas as observações. Os correspondentes algoritmos de inferência posterior permitem analisar os dados sob essas premissas: inferir a estrutura oculta que melhor explica as observações (HOFFMAN et al., 2013). Para muitos modelos de interesse, a posterior não é tratável computacionalmente e empregam-se métodos aproximados.

Neste capítulo é feita uma revisão da literatura na área de modelagem probabilística. Na próxima seção são apresentados os Modelos de Variáveis Latentes (LVM), assim como sua classificação. Os métodos para determinar os parâmetros dos modelos, vão ser resumidos e organizados na Seção 3.2.

3.1 MODELO DE VARIÁVEIS LATENTES

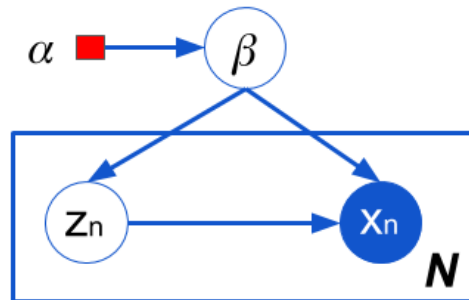
As variáveis latentes (ou variáveis ocultas) são variáveis que não são observadas diretamente, mas são inferidas de outras variáveis observadas, através de um modelo matemático. Os modelos matemáticos que tentam explicar as variáveis observadas em termos de variáveis latentes são chamados Modelos de Variáveis Latentes.

Quando se constrói um modelo de variáveis latentes, imaginam-se os tipos de quanti-

dades ocultas que poderiam ser usados para descrever os dados de interesse, e se codifica essa relação em uma distribuição de probabilidade conjunta de variáveis aleatórias e observadas. Em seguida, para um conjunto de dados observados, descobre-se as quantidades particulares ocultas que o descrevem através da probabilidade a posteriori, o que representa a distribuição condicional das variáveis ocultas dadas as observações.

Formalmente, um modelo consiste em três tipos de variáveis (Figura 3). O primeiro tipo é a observação que representa um ponto dos dados. Denota-se N pontos de dados (observações) como $x = x_{1:N}$. O segundo tipo são as variáveis ocultas, que codificam quantidades escondidas que são usadas para governar a distribuição das observações. Definem-se variáveis ocultas locais e globais: o vetor de variáveis ocultas globais é β ; as N variáveis locais escondidas são $z = z_{1:N}$. A distinção entre as variáveis escondidas globais e locais é determinada pelas dependências condicionais. Em particular, a n -ésima observação x_n e a n -ésima variável local z_n são condicionalmente independentes, dadas as variáveis globais β , para todas as outras observações e variáveis locais ocultas (HOFFMAN et al., 2013). O terceiro tipo é o hiperparâmetro, que é uma quantidade fixa não aleatória denotada por α .

Figura 3 – Representação Gráfica de um Modelo de Variáveis Latentes.



Fonte: Autor

Um modelo é uma distribuição conjunta de x e β , $p(\beta, x|\alpha) = p(\beta|\alpha)p(x|\beta)$, que descreve formalmente como as variáveis escondidas e as observações interagem em uma distribuição de probabilidade. Depois de observar os dados, busca-se a distribuição condicional das variáveis ocultas, $p(\beta|x, \alpha) \propto p(\beta, x|\alpha)$.

Segundo (BLEI, 2014), um modelo pode ser especificado de três maneiras:

- **Processo Gerador Probabilístico.** Descreve como os dados seriam gerados pelo modelo. Embora o modelo raramente seja o real, o processo gerador ajuda a tornar clara a forma como as variáveis latentes interagem para governar a distribuição das observações. A distribuição a posteriori pode ser considerada para reverter esse processo: fornecidos os dados, qual é a distribuição da estrutura oculta que provavelmente os gera? Este processo torna os hiper-parâmetros explícitos e também ajuda a identificar as variáveis ocultas locais e globais.

- **Distribuição Conjunta.** A maneira tradicional de representar um modelo é com a distribuição conjunta fatorizada de suas variáveis escondidas e observadas. Esta fatoração vem diretamente do processo gerador e para cada termo substitui-se a função de densidade apropriada. A distribuição conjunta viabiliza o cálculo da distribuição a posteriori para examinar a estrutura escondida particular que se manifesta nos dados observados.
- **Modelo Gráfico Dirigido.** Pode-se gerar um modelo de variável latente como um modelo gráfico probabilístico, outra representação que deriva do processo gerador. O processo gerador indica um padrão de dependência entre as variáveis aleatórias que podem ser codificadas em grafo cujos nós representam variáveis aleatórias e as arestas denotam as dependências entre eles. Este é um modelo gráfico que ilustra a estrutura da distribuição conjunta fatorada e o fluxo do processo gerador. Formalmente, um modelo gráfico representa a família de distribuições que respeite as independências que a mesma implica.

3.1.1 Tipos de Modelos

Muitos dos modelos descritos nesta seção foram descobertos em áreas específicas de pesquisa. Também foram fornecidos por algum algoritmo para inferir a probabilidade a posteriori, modelos por vezes desenvolvidos sem visar encontrar modelagem probabilística (BLEI, 2014). Eles têm aplicações úteis e uma história nas áreas de estatística e de aprendizagem de máquinas, destacando-se o seu uso como componentes em modelos mais complexos para problemas mais difíceis. Esta é a vantagem de a estrutura de modelagem probabilística.

A Figura 4 representa a taxonomia dos LVMs. Estes podem ser agrupados em modelos Bayesianos paramétricos ou não-paramétricos, e modelos adaptados de outros existentes. No caso dos dois primeiros, a principal diferença consiste em que o tamanho dos modelos não-paramétricos cresce com a quantidade dos dados, permitindo que estes determinem a estrutura dos modelos. A figura mostra também, para cada uma das categorias, exemplos de modelos usados nas pesquisas.

Modelos Bayesianos Paramétricos. Em um modelo Bayesiano paramétrico, além da especificação de uma distribuição de probabilidade conjunta para o vetor de dados condicional a um vetor de parâmetros $\theta\{\beta, z\}$, $f(x|\theta)$, há que considerar a especificação de uma distribuição de probabilidade para o próprio vetor de parâmetros. Esta distribuição é denominada distribuição *a priori* de θ , representada por $h(\theta)$, e expressa de alguma forma o conhecimento prévio à observação do conjunto de dados, do estatístico ou investigador, sobre os possíveis valores de θ . Esses modelos são relativamente simples, mas a principal desvantagem é a dificuldade de atribuir uma

Figura 4 – Tipos de Modelos de Variáveis Latentes.



Fonte: Autor

distribuição *a priori* para o vetor de parâmetros. Entre os modelos paramétricos mais usados encontram-se:

Modelos de Misturas. É um dos modelos mais simples de variáveis latentes. Um modelo de mistura pressupõe que os dados são agrupados e que cada amostra é extraída de uma distribuição associada com o seu agrupamento designado. As variáveis ocultas do modelo são as atribuições de *cluster* e os parâmetros para as distribuições por grupo. Com os dados observados, o modelo de mistura posterior é uma distribuição condicional sobre os agrupamentos e os parâmetros. Esta condicional identifica um grupo provável para os dados e as características de cada grupo (BLEI, 2014).

Modelos de Fatores Lineares. Mapeiam dados observados de alta dimensão em um espaço de baixa dimensionalidade. Estes modelos têm sido um pilar no domínio da estatística por quase um século; análise de componentes principais, análise fatorial e análise de correlação canônica podem ser interpretadas desta forma, embora a perspectiva probabilística sobre eles seja mais recente (ROWEIS, 1998; TIPPING; BISHOP, 1999; COLLINS; DASGUPTA; SCHAPIRE, 2001). Em um modelo de fatores, há um conjunto de componentes ocultos, e cada ponto de dados está associado a um vetor de pesos escondidos, com um peso para cada componente. Os dados surgem a partir de uma distribuição, cujos parâmetros combinam as componentes globais com os pesos por ponto dos dados. Tradicionalmente, os dados são extraídos de uma gaussiana (TIPPING; BISHOP, 1999), mas as extensões para fatores lineares têm considerado famílias exponenciais (COLLINS; DASGUPTA; SCHAPIRE, 2001; MOHAMED; GHAMRANI; HELLER, 2009).

Modelos Mistos de Pertinência. São utilizados para a análise não-supervisionada

de dados agrupados, vários conjuntos de observações que estão estatisticamente relacionados entre si. Em um modelo misto de pertinência, cada grupo é modelado a partir de uma mistura, na qual as proporções da mistura são únicas para o grupo e os componentes da mistura são compartilhados entre os grupos. A distribuição a posteriori do modelo misto de pertinência revela os padrões recorrentes nos dados e a proporção com que eles ocorrem em cada grupo. Contrasta esta situação com a de modelos de mistura simples, na qual cada grupo de dados está associado com um único componente de mistura. Os modelos mistos de pertinência produzem melhores previsões do que os modelos de mistura e fornecem uma estrutura exploratória mais interessante (BLEI, 2014).

Modelos de Matriz de Fatorização. Muitos conjuntos de dados são organizados em uma matriz, na qual cada observação é indexada por uma linha e uma coluna. Cada célula observada é modelada por uma distribuição cujos parâmetros são uma combinação linear da linha e da coluna. Condicionando uma matriz observada, a distribuição posterior fornece representações de baixa dimensão de suas linhas e colunas. Estas representações podem ser utilizadas para formar as previsões sobre células ocultas e para explorar a estrutura escondida nos dados. Este modelo está intimamente relacionado com o modelo de fator linear, exceto que a distribuição de cada uma das células é determinada por variáveis ocultas que dependem da linha e da coluna da célula.

Modelos de Séries Temporais. Muitas observações são sequenciais: elas são indexadas pelo tempo ou posição, e se quer aproveitar esta estrutura ao fazer inferências. Em estatística, econometria, matemática aplicada e processamento de sinais, uma série temporal é uma coleção de observações feitas sequencialmente ao longo do tempo, onde a ordem dos dados é fundamental. Uma característica muito importante deste tipo de dados é que as observações vizinhas são dependentes e o interesse é analisar e modelar esta dependência (WOOLDRIDGE, 2015).

Dois modelos de séries temporais paramétricos são o modelo escondido de Markov (HMM) e o filtro de Kalman. Cada um teve aplicações científicas e de engenharia importantes, e cada um demonstra como pode se usar modelos de variáveis latentes simples para construir modelos mais complexos.

Modelos Escondidos de Markov. Em um HMM, cada observação das séries temporais é projetada a partir de uma de mistura de componentes não observadas, e esse componente é projetado condicionado à componente de mistura da observação anterior. A posteriori é semelhante a um modelo de mistura, onde o modelo assume uma estrutura Markoviana na sequência de tarefas de mistura. HMMs foram usados com sucesso em muitas aplicações, especialmente no reconhecimento de fala (RABINER, 1989) e biologia computacional (DURBIN et al., 1998).

Filtro de Kalman. Um filtro de Kalman é uma adaptação da série temporal de um modelo de fator linear com gaussianas (KALMAN, 1960), utilizando o modelo de espaço de estado. Os Filtros de Kalman são influentes em atividades industriais (AUGER et al., 2013), rastreamento de alvos e de radar (BAR-SHALOM; LI; KIRUBARAJAN, 2004), robótica (CHEN, 2011) e atividades aeroespaciais (GREWAL; ANDREWS, 2010). As inferências a partir desse modelo ajudam a monitorar a verdadeira posição de um objeto e prever a sua próxima posição. Embora as distribuições subjacentes sejam diferentes, no filtro de Kalman as cadeias ocultas de variáveis são uma sequência de variáveis contínuas, ao invés de discretas.

Modelos Bayesianos Não-Paramétricos. Os modelos não-paramétricos constituem uma abordagem para seleção e adaptação de modelos, onde os tamanhos dos mesmos crescem com o tamanho de dados. Esses modelos permitem que os dados determinem a estrutura das variáveis latentes.

Mais precisamente, um modelo Bayesiano não-paramétrico é caracterizado por (1) estar imerso em um espaço de parâmetros de dimensão infinita e (2) ser avaliado em uma amostra finita de maneira que use apenas um subconjunto finito dos parâmetros disponíveis para explicar a amostra. O espaço de parâmetros tipicamente consiste de funções ou de medidas, enquanto a avaliação geralmente desconsidera as dimensões excedentes. Os modelos Bayesianos não-paramétricos foram aplicados a uma variedade de problemas de aprendizagem de máquina, incluindo regressão, classificação, agrupamento, modelagem de variável latente, modelagem sequencial, segmentação de imagens e indução gramatical (ORBANZ; TEH, 2010)

Processo de Dirichlet. O processo de Dirichlet (DP) foi introduzido por (FERGUSON, 1973) e é definido por um parâmetro de concentração, $c > 0$, e uma distribuição de referência sobre o espaço amostral mensurável, G_0 . Tal processo pode ser visto como uma distribuição sobre medidas de probabilidade, isto é, cada realização de um processo de Dirichlet é, por si só, uma distribuição de probabilidade. O processo de Dirichlet requer a especificação de uma distribuição de referência, G_0 . (ANTONIAK, 1974) sugere centrar o processo de Dirichlet em uma família de distribuições paramétricas, com o objetivo de incorporar a família paramétrica na ampla classe de modelos para G_0 . Surge, então, o modelo de mistura de processos de Dirichlet (Mixture of Dirichlet Processes). A simplicidade das suas propriedades e a facilidade de amostrar de um processo de Dirichlet, fizeram com que o modelo se tornasse atrativo e alvo de uma forte investigação nestas últimas décadas. No entanto, o processo de Dirichlet ou uma mistura de processos de Dirichlet gera quase certamente distribuições discretas (FERGUSON, 1973), limitando a sua aplicação a muitos dos problemas estatísticos de modelagem de dados contínuos.

Árvores de Pólya. As distribuições árvores de Pólya definem uma outra distribuição

a priori não-paramétrica, isto é, formam uma classe de distribuições para a medida de probabilidade aleatória G . Foi inicialmente discutida por (FREEDMAN, 1963), (FABIUS et al., 1964) e (FERGUSON, 1974). Estas distribuições são uma generalização do processo de Dirichlet. Em particular, permitem a modelação de distribuições contínuas ou absolutamente contínuas, contornando o problema da discretização do processo de Dirichlet. (LAVINE, 1992); (LAVINE, 1994) e (MAULDIN; SUDDERTH; WILLIAMS, 1992) desenvolveram e catalogaram detalhadamente a base teórica das distribuições árvores de Pólya, apresentando várias aplicações práticas.

Modelos de Características Latentes. Apresentam um conjunto de objetos em termos de um conjunto de características latentes, cada um representando um grau independente da variação apresentada pelos dados. Tal representação dos dados é muitas vezes referida como uma representação distribuída, e pelo o processo estocástico envolvido, ele é conhecido como processo beta (*Beta Process* (BP)). Em analogia com modelos DP com um número desconhecido de *clusters*, o modelo permite um número desconhecido de características latentes (TEH; JORDAN, 2010).

Modelos Hierárquicos. Os modelos hierárquicos permitem que os dados exibam grupos complexos e sobrepostos. A ideia básica é que os parâmetros sejam dotados de distribuições que podem introduzir novos parâmetros, isto é, uma construção recursiva. Um tema comum em modelagem hierárquica é a independência condicional da hierarquia, na qual um conjunto de parâmetros são acoplados ao fazer as suas distribuições depender de um parâmetro subjacente compartilhado. Essas distribuições são muitas vezes tomadas para ser idênticas, com base em uma afirmação de permutabilidade (TEH; JORDAN, 2010). Exemplos populares desses modelos incluem:

Processo Hierárquico de Dirichlet. O processo de Dirichlet hierárquico (HDP) é útil em problemas em que há vários grupos de dados, onde o modelo para cada grupo incorpora uma variável discreta de cardinalidade desconhecida, e onde busca-se vincular estas variáveis entre os grupos (TEH et al., 2006). Por exemplo, o modelo de mistura HDP permite compartilhamento de grupos através de vários problemas de agrupamento.

Modelos Escondidos de Markov Infinitos. É uma abordagem Baysiana não-paramétrica para HMMs, onde o número de estados é ilimitado e permite o crescimento com o tamanho da sequência. A ideia é misturar o HDP com o HMM. O modelo resultante é, portanto, referido como o modelo Processo Hierárquico Dirichlet com Modelo Escondido de Markov (HDP-HMM). A dificuldade com o HDP-HMM é que ele tende a ser pobre em capturar a persistência do estado; ele tem uma tendência para criar estados redundantes e comutar rapidamente entre eles. Isto pode não ser problemático para aplicações em que os estados são variáveis de perturbação e o que importa

é a probabilidade de previsão global, mas pode ser problemático para aplicações de segmentação ou de análise na qual os estados são objeto de inferência e quando é esperada a persistência do estado (TEH; JORDAN, 2010).

Adaptação de Modelos Existentes. A modelagem probabilística pode ser usada como uma linguagem de hipóteses para adaptar os modelos de variáveis latentes aos dados. Existem diferentes maneiras de se adaptar e desenvolver novos modelos de variáveis latentes (BLEI, 2014):

Mudando a distribuição geradora de dados do modelo existente. O processo probabilístico gera uma observação condicionada na estrutura latente. A modificação da distribuição pode levar a novos modelos de variáveis latentes. Poder-se-ia empregar distribuições ordinais para os dados ordenados, distribuições gama e gaussianas truncadas para dados positivos, ou distribuições discretas para dados categóricos.

Mudando a suposição da distribuição nas variáveis latentes. Poder-se-ia substituir uma gaussiana por uma distribuição gama para impor positividade. Alternativamente, poderia se substituir distribuições simples por distribuições mais complexas com estruturas latentes.

Mistura de componentes de diferentes modelos. Poder-se-ia misturar e combinar componentes de diferentes modelos para criar novas técnicas.

(BLEI, 2014) propõe um conjunto de etapas para usar modelos de variáveis latentes para resolver problemas de análise de dados. Primeiramente, é formulado um modelo simples baseado nos tipos de estruturas ocultas que pode existir nos dados. Então, dado o conjunto de dados, é usado um algoritmo de inferência para aproximar a distribuição posterior que aponta para os padrões escondidos específicos que exibem os dados. Finalmente, é utilizado a posterior para testar o modelo, identificando os sucessos e falhas. Se o modelo representa os dados, é usado para resolver o problema se não, deve-se achar um novo modelo. Este processo aplicado à análise exploratória de dados permite resumir, constatar e formular hipóteses sobre os dados observados.

Uma variedade de abordagens foi proposta para resolver o problema de determinar os parâmetros do modelo relativo a observações realizadas. Os principais métodos são tratados na próxima seção.

3.2 INFERÊNCIA DA PROBABILIDADE A POSTERIORI

As tarefas de predição e análises dos dados, usando modelos de variáveis latentes, dependem de calcular a distribuição a posteriori: a distribuição condicional das variáveis ocultas dadas as observações. Calcular a distribuição a posteriori é o problema algorítmico central na modelagem probabilística. Este é o problema de inferência da distribuição a posteriori

(BLEI, 2014) calculada pela Equação 3.1:

$$p(\vartheta|x) = \frac{p(x|\vartheta)p(\vartheta)}{\int p(x|\vartheta)p(\vartheta)d\vartheta} \quad (3.1)$$

onde $p(x|\vartheta)$ é o modelo, e $p(\vartheta)$ é a distribuição *a priori* dos parâmetros do modelo.

Por vezes, esta distribuição (Equação 3.1) não pode ser obtida analiticamente, principalmente por depender de integrais que não tem solução analítica, ou cuja solução é custosa computacionalmente. Nestes casos, há a necessidade da utilização de métodos aproximados de inferência, podendo estes ser determinísticos ou de simulação, em particular, de simulação estocástica (RESENDE, 2011). As diferentes alternativas são esquematizadas na Figura 5.

Figura 5 – Métodos de Inferência para determinar os parâmetros dos modelos.



Fonte: Autor

As aproximações determinísticas foram desenvolvidas principalmente na década de 1980. Até então, os métodos computacionalmente intensivos estavam no início de seu desenvolvimento. A complexidade das técnicas determinísticas cresce à medida que a dimensão do espaço paramétrico cresce.

Os métodos de inferência baseados em simulação estocástica utilizam amostras da distribuição a posteriori para sumarizar informação. Esses métodos são utilizados para obter uma aproximação da distribuição a posteriori quando não é possível obtê-la analiticamente ou como alternativa aos métodos numéricos, devido a altas dimensões paramétricas. A partir da década de 1990, com o avanço das técnicas computacionais e com o trabalho de (GELFAND; SMITH, 1990), esses métodos começaram a ser extensamente explorados. Esta metodologia de aproximação pode ser excessivamente custosa computacionalmente em determinadas situações, o que foi uma das motivações para o estudo de métodos de aproximação alternativos.

Nas seções que seguem, uma breve revisão de literatura é feita sobre os métodos de inferência, com o objetivo de entender o contexto histórico ao qual o desenvolvimento dos métodos aproximados de inferência se inserem.

3.2.1 Métodos Determinísticos

Desde a década de 1980, muitos trabalhos surgiram na literatura abordando aproximações numéricas, (NAYLOR; SMITH, 1982), (TIERNEY; KADANE, 1986), (SMITH et al., 1987), (TIERNEY; KASS; KADANE, 1989). Métodos determinísticos vêm sendo explorados na literatura, principalmente na literatura de aprendizagem de máquina. Entre os métodos determinísticos mais conhecidos encontram-se:

Método de Laplace. Dá uma solução ao aproximar a integral do denominador da Equação (3.1) à posteriori, e a variância dada pelo inverso da segunda derivada do logaritmo da distribuição da posteriori à sua moda. Além da aproximação feita no denominador, (TIERNEY; KADANE, 1986) propuseram localizar a moda do numerador, achar a segunda derivada nesta moda e então aproximar esta integral através de uma segunda aplicação do método de Laplace. Assim, aplicando a mesma técnica de aproximação no numerador e no denominador, os erros associados às aproximações serão similares e a razão dessas aproximações fará com que parte desses erros se cancelem. O método é relativamente rápido e apropriado quando o integrando é unimodal e o máximo pode ser obtido facilmente. O tempo computacional necessário para obter essa aproximação cresce à medida que a dimensão do espaço paramétrico cresce (RESENDE, 2011).

(RUE; MARTINO; CHOPIN, 2009) propuseram uma melhora ao método de Laplace, conhecido como método de aproximação INLA (*Integrated Nested Laplace Approximation* (INLA)). O método foi usado para realizar inferência Bayesiana de forma aproximada, em uma subclasse de modelos de regressão estruturados aditivos chamados modelos latentes gaussianos. O INLA aproxima de forma direta as distribuições a posteriori de interesse. Com isso, é possível realizar inferência com um custo computacional consideravelmente inferior além de não ser preciso lidar com problemas de convergência.

Variacional de Bayes. O método *Variational Bayes* (VB) é uma forma de realizar inferência aproximada através da transformação de problemas de inferência complexas em problemas de otimização de alta dimensão ((JORDAN et al., 1999), (WAINWRIGHT; JORDAN, 2008)). Como o método é aproximado, o custo computacional é significativamente inferior às abordagens via MCMC e é também inferior à aproximação de Laplace, já que não é necessário o cálculo da matriz hessiana.

Variacional Estocástico. É baseado na Inferência Variacional (*Variational Inference* (VI)), mas ele é um algoritmo mais eficiente que utiliza otimização estocástica (ROBBINS; MONRO, 1951), uma técnica que segue estimativas ruidosas do gradiente do objetivo. O método itera entre a subamostragem dos dados e o ajuste da estrutura escondida com base apenas na subamostra, o que faz ao método muito mais

competente para analisar grandes conjuntos de dados com modelos probabilísticos complexos (HOFFMAN et al., 2013). O algoritmo usa a otimização estocástica para melhorar os métodos de inferência posterior aproximada. Pode-se também considerar dados cuja distribuição muda ao longo do tempo, como quando se quer modelar um fluxo infinito de dados, mas esquecer os dados do passado distante na estimativa atual do modelo.

Propagação da Expectativa. Outro esquema de aproximação comum em aplicações de aprendizagem de máquina é a abordagem *Expectation Propagation* (EP), proposto por (MINKA, 2001). O método EP, similarmente ao VB, minimiza iterativamente a pseudo-distância entre $p(x, \theta | \alpha)$ e a aproximação $q(x, \theta)$, sujeito a uma fatorização de $q(x, \theta)$ em que cada um dos fatores paramétricos envolva apenas um único componente de (x, θ) . Neste caso, deseja-se achar a aproximação $q(x, \theta)$ que minimiza a divergência de Kullback-Leibler de $q(x, \theta)$ com respeito a $p(x, \theta | \alpha)$ e não o contrário como é feito no método VB. Como visto em (BISHOP, 2007) e (RUE; MARTINO; CHOPIN, 2009), por conta disso, o método EP, em geral, superestima a variância a posteriori.

Maximização da Expectativa. O algoritmo de maximização da expectativa (DEMPS-TER; LAIRD; RUBIN, 1977) (*Expectation Maximization* (EM)) é um método iterativo para encontrar estimativas da máxima verossimilhança ou máxima a posteriori (maximum a posteriori MAP), em modelos estatísticos que dependem de variáveis latentes não observadas. A iteração do EM alterna entre executar uma etapa de expectativa (E), que cria uma função para a expectativa da log-verossimilhança avaliada usando a estimativa atual dos parâmetros, e uma etapa de maximização (M), que calcula os parâmetros maximizando a log-verossimilhança encontrada no passo E. Essas estimativas de parâmetros são então usadas para determinar a distribuição das variáveis latentes no passo E seguinte. Um dos aspectos negativos do algoritmo EM é a sua convergência lenta. Para aumentar a sua rapidez de convergência utilizam-se principalmente aproximações de Newton, incluindo os métodos quasi-Newton. Recentemente outros algoritmos, tais como o algoritmo EM Incremental (IEM), o algoritmo Sparse EM (SPEM), têm sido propostos para aumentar a rapidez de convergência do algoritmo EM, preservando a sua simplicidade (FARIA, 2006).

3.2.2 Métodos de Simulação Estocástica

Os métodos de simulação estocástica utilizam amostras da distribuição a posteriori para sumarizar informação quando não é possível obter a distribuição a posteriori de forma analítica. (GAMERMAN; LOPES, 2006) apresentam uma descrição detalhada dos métodos de simulação estocástica, em especial, dos métodos de Monte Carlo via cadeias de Markov

(MCMC). Estes são métodos de simulação estocástica que consistem na construção de uma cadeia de Markov que, por meio de escolhas adequadas de núcleos de transição, tem como distribuição estacionária a distribuição de interesse. Uma vez que a convergência da cadeia tenha sido atingida, as amostras estarão sendo geradas da distribuição estacionária. Para aproximar a distribuição a posteriori, utilizam-se amostras suficientemente grandes da distribuição estacionária.

Dentre os métodos mais utilizados para a construção da cadeia de Markov está o Amostrador de Gibbs proposto por (GEMAN; GEMAN, 1984), sendo subsequentemente popularizado por (GELFAND; SMITH, 1990), e o método de Metropolis-Hastings proposto por (METROPOLIS et al., 1953) e estendido por (HASTINGS, 1970).

Algoritmo de Metropolis-Hastings. Utiliza-se este método para gerar amostras de uma distribuição de interesse a qual não é possível obter analiticamente. Em sua aplicação, utiliza-se uma distribuição auxiliar, chamada de distribuição proposta, para gerar um valor proposto desta distribuição e aceitá-lo ou não de acordo com uma probabilidade. Este mecanismo de correção garante a convergência da cadeia para a distribuição de equilíbrio.

Amostrador de Gibbs. O amostrador de Gibbs é um caso especial do algoritmo de Metropolis-Hastings. Este método gera uma sequência de amostras de uma distribuição conjunta de probabilidades de duas ou mais variáveis aleatórias de interesse $p(\Psi)$. O núcleo de transição da cadeia é formado pelas distribuições condicionais das componentes Ψ_j de um vetor paramétrico $\Psi, \forall j$. Para cada parâmetro, a distribuição condicional é a distribuição deste parâmetro condicional à informação de todos os outros parâmetros.

Apesar de sua popularidade em inferência Bayesiana, relativamente poucos trabalhos têm se concentrado no desenvolvimento de algoritmos MCMC que podem ser dimensionados para grandes conjuntos de dados. Uma exceção é o método Monte Carlo Sequencial, embora este geralmente não tenha fortes garantias de convergência. Outro método é o gradiente Langevin estocástico de (WELLING; TEH, 2011) que oferece garantias de convergência assintótica e também tira proveito da otimização estocástica. Finalmente, na modelagem do tópico, os pesquisadores desenvolveram várias abordagens para paralelizar o MCMC (NEWMAN et al., 2009), (SMOLA; NARAYANAMURTHY, 2010), (AHMED et al., 2012).

Além da dificuldade no diagnóstico da convergência e o alto custo computacional, os métodos MCMC apresentam outras desvantagem do ponto de vista computacional devido à dependência dos elementos do campo latente, e da dependência existente entre os hiperparâmetros e os parâmetros do campo latente (RESENDE, 2011).

3.3 DISTRIBUIÇÃO PREDITIVA

Normalmente, dois tipos de tarefas lançam mão de um modelo: exploração e previsão. Na exploração, as inferências são feitas sobre as variáveis ocultas - geralmente através de expectativas posteriores aproximadas - para resumir, visualizar ou analisar os dados, ou seja, dividi-lo em grupos com estruturas ditadas pelas inferências. Exemplos de tarefas de exploração incluem o uso de modelos de tópicos para navegar grandes coleções de documentos e utilização de modelos de *clustering* em dados de microarranjos para sugerir grupos de genes relacionados (BLEI, 2014).

Na previsão, os dados futuros são previstos através da distribuição preditiva posterior:

$$p(x_{new}|x) = \int p(\beta, x) \left(\int p(z_{new}|\beta) p(x_{new}|z_{new}, \beta) dz_{new} \right) d\beta \quad (3.2)$$

A distribuição preditiva posterior ou distribuição de probabilidade preditiva, é a distribuição de dados preditos pelo modelo estatístico. Ela consiste na curva da função de densidade preditiva do modelo, um critério gráfico para avaliar a precisão do ajuste do modelo (POLETO, 2011).

Normalmente, a distribuição a posteriori $p(\beta|x)$ usada para a predição não está disponível, portanto, é substituída por uma aproximação $q(\beta)$, determinada por um algoritmo de inferência aproximada como MCMC ou inferência variacional.

A distribuição de probabilidade preditiva, permite que se façam previsões em modelos de regressão e em estudos que envolvem um componente temporal (KINAS; ANDRADE, 2014). Exemplos de tarefas preditivas incluem o uso de uma matriz de fatoração para prever quais itens um usuário irá comprar ou o uso de um modelo de séries temporais para prever preços futuros das ações com base em suas histórias (BLEI, 2014). Porém, essa forma de inferência também é importante para estudar a validade dos modelos probabilísticos utilizados nas análises Bayesianas. Como novos valores x_{new} devem ser observados com probabilidades $p(x_{new}/x)$, pode-se inferir quais são os valores prováveis. Se isso não se confirmar na prática, há indícios de que pelo menos um dos componentes do modelo (verossimilhança ou *a priori*) pode estar especificado inadequadamente (KINAS; ANDRADE, 2014).

3.4 CONCLUSÕES DO CAPÍTULO

O estudo realizado sugere que os modelos paramétricos de mistura podem oferecer uma base probabilística para a análise das populações de um EA. Especificamente, pode ser usado o Modelo de Mistura Gaussiana (*Gaussian Mixture Model* (GMM)) baseado em que um conjunto de gaussianas pode aproximar qualquer distribuição de um conjunto de dados. Além disso, o custo computacional na estimação dos parâmetros é menor neste modelo.

Os modelos Bayesianos não-paramétricos diferem dos modelos paramétricos porque a estrutura daqueles não precisa ser especificada a priori, podendo ser determinada com base nos dados. O termo "não-paramétrico" rotula modelos nos quais o número e a natureza dos parâmetros são flexíveis e não são fixados antecipadamente. Por sua vez, esta flexibilidade aumenta a complexidade computacional da análise. consequentemente, o desenvolvimento dos modelos Bayesianos não-paramétricos nas últimas duas décadas se deve, em parte, ao grande desenvolvimento das recursos computacionais.

Nos modelos de séries temporais as observações são feitas ao longo do tempo, onde a ordem dos dados é fundamental. Esta característica pode ser facilmente aplicada aos EAs nos quais as populações podem ser vistas como pontos indexadas pelo tempo. Uma outra característica muito importante é que as observações vizinhas no tempo são dependentes e o interesse é analisar e modelar esta dependência. Isto permitiria realizar algum tipo de previsão nos algoritmos.

Na Seção 3.2 foram mencionadas as técnicas mais conhecidas para aproximar a distribuição posterior, isto é, para determinar os parâmetros do modelo dadas as observações. Delas, as duas estratégias mais empregadas em estatística e aprendizagem de máquina são MCMC e Inferência Variacional. Enquanto amostragem MCMC dá garantias teóricas de precisão, a Inferência Variacional possibilita rápida e determinística aproximação para probabilidades a posteriori. Além disso, ambos os métodos são paradigmas computacionais que possuem ampla variedade de abordagens algorítmicas específicas que permite *trade-off* entre velocidade e precisão da resposta além de facilidade de implementação por diferentes maneiras.

Este trabalho emprega Inferência Variacional para determinar os parâmetros de um modelo escolhido. Como se viu neste capítulo, a Inferência Variacional calcula uma probabilidade marginal ou condicional em termos de um problema de otimização. O método permite também considerar dados cuja distribuição muda ao longo do tempo, como quando se quer modelar um fluxo infinito de dados, ao mesmo tempo que permite desprezar dados do passado distante na estimativa atual de um modelo. Os métodos de Inferência Variacional têm sido desenvolvidos, principalmente, no contexto da família exponencial, onde as propriedades de convexidade do espaço de parâmetros natural e a função cumulativa compõem um formalismo variacional elegante. Outro motivo para a escolha da Inferência Variacional é o custo computacional significativamente inferior àqueles das abordagens que empregam MCMC e é também inferior à aproximação de Laplace, já que não é necessário o cálculo da matriz hessiana. Além disso, o algoritmo é mais robusto à inicialização dos parâmetros quando comparado com outros métodos como Maximização da Expectativa (EM).

Por último, pode-se enfatizar que o conhecimento da distribuição preditiva é suficiente para realizar inferências preditivas, ou até tomar decisões com base no que se espera do futuro.

4 ALGORITMO EVOLUCIONÁRIO MULTIOBJETIVOS DINÂMICO COM PREDIÇÃO NO ESPAÇO OBJETIVO

Em problemas dinâmicos deve ocorrer algum tipo de adaptação no ambiente e, de acordo com o problema, podem haver restrições de tempo para estas adaptações. Se em um dado problema ocorrerem mudanças de baixa intensidade, isto é, após a modificação do problema as novas soluções encontram-se próximas às anteriores, então o conhecimento prévio do espaço de decisões pode ser utilizado na resposta de um EA à mudança do ambiente. Se a mudança for grande, a solução anterior pode não ser útil pois sua adaptação poderia ser equivalente ao reinício da evolução.

Assim, os problemas dinâmicos de otimização que envolvem mudanças de porte não precisariam de algoritmos dinâmicos para resolvê-las, simplesmente, recorrendo a reinicialização do EA. Logo, para problemas cujas mudanças não são de grande porte ou apenas parte do ambiente sofre modificações, o conhecimento prévio do espaço de decisão do problema pode acelerar o processo de adaptação e, em alguns casos, obter soluções que não seriam alcançadas através de métodos estáticos.

Em geral, os EAs baseados em predição executam o procedimento de previsão de movimentação dos indivíduos no espaço de decisão, um espaço predominantemente com dimensões significativamente mais altas do que o espaço objetivo. Não obstante o espaço de decisão ser usado para predição, as informações extraídas do espaço objetivo tendem a ser mais consistentes para estabelecer as tendências de movimentação da frente. Estes movimentos no espaço objetivo tendem a ocorrer em direção à frente ótima de Pareto através da inserção de novas soluções não dominadas. Portanto, movimentos populacionais promissores no espaço de decisão podem ser previstos se forem determinados indivíduos que levem a pontos em uma frente prevista no espaço objeto, frente esta que pode ser determinada por sua tendência de progressão à frente ótima. Em essência, um deslocamento da população baseado em movimentos anteriores de indivíduos no espaço de decisão pode levar o processo a regiões não promissoras de espaço objetivo, prejudicando a convergência do MOEA. Por outro lado, se o deslocamento for baseado na progressão da frente, ele pode determinar com maior precisão movimentos populacionais propícios no espaço de decisão.

Tendo em mente esse pressuposto, nesta tese optou-se por estimar o progresso da frente no espaço objetivo e, em seguida, encontrar indivíduos no espaço de decisão que se encaixam na frente prevista. Um passo de melhoria da frente é determinado, e a população é alterada para alcançá-lo.

Para rastrear e prever as mudanças da paisagem, bem como para acelerar a convergência de um algoritmo antes de uma nova modificação do ambiente, este capítulo propõe o algoritmo chamado Evolução Diferencial de Ordenação Não Dominada Dinâmica com

Previsão no Espaço Objetivo (DOSP-NSDE). O algoritmo proposto fornece orientação guiada para a evolução da população através de uma nova estratégia de previsão no espaço objetivo chamada (*Objective Space Prediction strategy* (OSP)). Este é usado tanto no mecanismo de reação à mudança (CRM) para responder a um novo cenário, quanto no processo evolucionário para acelerar a convergência. Se o OSP for aplicado como CRM, um modelo será construído usando a sequência anterior de locais da Frente de Pareto aproximada determinadas pelo algoritmo durante as etapas de tempo anteriores. Esta sequência passada pode ser vista como uma série temporal estocástica. Quando o OSP é usado no processo evolucionário, ele considera informações úteis das gerações passadas por meio de métodos estatísticos para auxiliar na estimativa de uma região promissora. A estratégia extrai informações tanto do espaço de decisão quanto do espaço objetivo. As informações estatísticas obtidas no espaço objetivo estimam com mais precisão o provável comportamento do movimento da frente e, assim, aumentam a eficácia e eficiência do algoritmo. O DOSP-NSDE emprega um critério baseado na métrica do Hipervolume Aproximado (HV) para desencadear o OSP em gerações adequadas. Se a condição baseada em HV não for satisfeita, o algoritmo executa os operadores de variação da evolução diferencial e o operador de mutação polinomial, além do mecanismo de diversidade baseado em k vizinhos mais próximos (ZITZLER; LAUMANN; THIELE, 2001) para equilibrar exploração e exploração.

4.1 FUNDAMENTOS PROBABILÍSTICOS PARA O MODELO PROPOSTO

Nesta Secção são apresentados os fundamentos do novo DMOEA: 1) o modelo de previsão *AutoRegressive Model* (AR) para estimar regiões promissoras para o processo evolucionário e para a nova paisagem; 2) os Modelos de Mistura Gaussiana (GMM), e 3) o método Inferência Variacional (VI), ferramentas necessárias para gerar novos indivíduos em áreas promissoras nos dois espaços de trabalho.

4.1.1 Modelo AutoRegressivo (AR)

Um processo evolucionário pode ser caracterizado como uma série temporal, isto é, como uma coleção de observações de amostras (indivíduos na população) X_t sequencialmente tomadas ao longo do tempo t que muda de acordo com um determinado conjunto de premissas de movimento (RINCÓN, 2012). Se a população evolui como consequência de algum mecanismo aleatório, então X_t é uma variável aleatória para cada valor do índice t ; assim, $\{X_t : t \in T\}$ é um processo estocástico. Muitas vezes, variáveis aleatórias dependentes formam esse processo. Portanto, a análise e a modelagem da dependência possibilitam representar a evolução probabilística de um sistema ao longo do tempo.

O modelo AutoRegressivo (AR) especifica que a variável de saída depende linearmente de seus próprios valores anteriores e de um termo estocástico (um termo previsivelmente

imperfeito). O modelo tem a forma de uma equação de diferença estocástica (AKAIKE, 1969).

O modelo AR de ordem p , é matematicamente definido como:

$$\begin{aligned} X_t &= c + \sum_{i=1}^p \Phi_i X_{t-i} + \varepsilon_t \\ X_t &= \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \cdots + \Phi_p X_{t-p} + \varepsilon_t \end{aligned} \quad (4.1)$$

onde p é um inteiro não negativo, $\Phi_1, \dots, \Phi_p$ são os parâmetros do modelo, c é uma constante, e ε_t é ruído branco, $\varepsilon_t \sim N(0, \sigma^2)$, e ε_t não está correlacionado com X_s para cada $s < t$ (AKAIKE, 1969).

Os modelos auto-regressivos são flexíveis para lidar com ampla gama de diferentes padrões de séries temporais, alterando os parâmetros $\Phi_1, \dots, \Phi_p$ (AKAIKE, 1969). O modelo AR pode ser usado para prever um número arbitrário de instantes de tempo no futuro. A variável de interesse é prevista por autorregressão que usa uma combinação linear de seus valores anteriores (AKAIKE, 1969):

$$\begin{aligned} X_{T+k} &= \Phi^k X_T + \Phi^{k-1} \varepsilon_{T+1} + \cdots + \Phi \varepsilon_{T+k-1} + \varepsilon_{T+k} \\ X_{T+k,T} &= E(X_{T+k} | I_T) = \Phi^k X_T \end{aligned} \quad (4.2)$$

onde $I_T = \{X_1, X_2, \dots, X_T; \varepsilon_1, \varepsilon_2, \dots, \varepsilon_T\}$ (AKAIKE, 1969).

Outros modelos probabilísticos têm sido utilizados para extrair o comportamento probabilístico da população (GUERRERO-PEÑA; ARAÚJO, 2017), para gerar novas soluções em regiões promissoras do espaço de busca. Um dos modelos mais utilizados e simples é o Modelo de Mistura Gaussiana GMM.

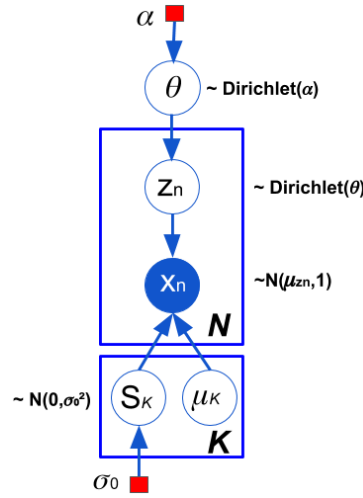
4.1.2 Modelo de Mistura Gaussiana (GMM)

O Modelo de Mistura Gaussiana (GMM) é um modelo probabilístico que assume que os dados são agrupados e que cada ponto de dados é extraído de uma distribuição normal associada ao seu grupo atribuído. A Figura 6 tem a representação gráfica do GMM.

O GMM é um dos modelos de variáveis latentes paramétricos mais simples, Equação 4.3. Este tem como variáveis ocultas os K componentes da mistura; as médias $\mu_{1:K}$ e S_k que é a precisão ou inversa da matriz de covariância da distribuição gaussiana para cada grupo; o conjunto de pesos de mistura θ que satisfaz $0 \leq \theta \leq 1$ e $\sum_{k=1}^K \theta_k = 1$; e as atribuições de mistura para cada ponto de dados $z_{1:N}$. A atribuição de componente z_n é escolhida a partir das proporções da mistura. Com ele pode ser amostrado o ponto de dados x_n da Gaussiana correspondente: $x_i | z_i \sim N(\mu_{z_i}, 1)$. Os hiperparâmetros fixos η são os parâmetros da distribuição dos parâmetros do modelo (BLEI, 2014).

$$p(x|\vartheta) = \sum_{k=1}^K p(z_n = k|\theta) p(x_n | z_n = k, \vartheta) = \sum_{k=1}^K \theta_k N(x | \mu_k, S_k) \quad (4.3)$$

Figura 6 – Representação Gráfica do Modelo de Mistura Gaussiana.



Fonte: Autor

onde $\vartheta = \{\theta_k, \mu_k, S_k\}$

Os parâmetros do GMM vão ser determinados estimando a distribuição a posteriori (Equação 3.1), dos componentes da mistura, as proporções da mistura e a forma como os dados são agrupados $p(\mu_{1:K}, \theta, z_{1:N} | x_{1:N})$. A distribuição a posteriori revela a estrutura escondida nos dados, agrupa-os em K grupos e descreve a localização (a média) de cada grupo (BLEI, 2014). A distribuição posterior vai ser aproximada usando a Inferência Variacional (VI).

4.1.3 Inferência Variacional (VI)

Em muitos problemas estatísticos, a distribuição posterior não pode ser calculada diretamente porque a constante de normalização é intratável, então métodos para aproximá-la devem ser usados, sendo a Inferência Variacional (BLEI, 2014) uma alternativa adequada. VI é uma das técnicas mais conhecidas e rápidas para aproximar as distribuições a posteriori.

A Inferência Variacional usa uma família de modelos onde cada variável escondida é independente, indexada por parâmetros livres; encontrando-se a configuração dos parâmetros (o membro da família) que for mais próximo à distribuição a posteriori. Assim, resolve-se o problema de inferência através da resolução da otimização da divergência de Kullback-Leibler (KULLBACK; LEIBLER, 1951).

O método aproxima uma distribuição de probabilidade intratável, p , com uma tratável, q , tornando-as tão próximas quanto possível. Uma distribuição a posteriori é tratável se a integral puder ser estimada. Uma maneira de conseguir isso é fazendo cada ϑ_i independente:

$$q(\vartheta | \lambda) = \prod_{i=1} q_i(\vartheta_i | \lambda_i) \quad (4.4)$$

para algumas distribuições marginais q_i parametrizadas por λ_i .

O método usa a medida de divergência de Kullback-Leibler que mede a discrepância entre $p(x, \theta)$ e $q(x, \theta)$. Duas distribuições são próximas uma da outra no sentido de Kullback-Leibler:

$$KL(q||p) \equiv \int q(\vartheta) \log \frac{q(\vartheta)}{p(\vartheta|x)} d\vartheta = E_q \left(\log \frac{q(\vartheta)}{p(\vartheta|x)} \right) \quad (4.5)$$

O objetivo, então, é achar a aproximação $q(x, \theta)$ que minimiza $KL(q||p)$, isto é, a divergência de Kullback-Leibler de $p(x, \theta)$ com respeito a $q(x, \theta)$. A minimização é feita com restrição e, em geral, assume-se que os estados e os hiperparâmetros são independentes, isto é, $q(x, \theta) = q_x(x)q_\theta(\theta)$. A fim de minimizar a divergência KL, $KL(q||p)$, basta maximizar o limite inferior da evidência (ELBO) (BLEI, 2014):

$$ELBO = E_q \left(\log \frac{p(x, \vartheta)}{q(\vartheta)} \right) \quad (4.6)$$

Uma maneira de otimizar sobre a escolha da Equação 4.4 é considerar o ELBO com respeito a algum $q_j(\vartheta_j)$, separando-o da expectativa com relação a todas as outras variáveis, $E_{-q_j}(\cdot)$, onde $\int_{-q_j}$ é a integral sobre todos os ϑ_i com $i \neq j$:

$$ELBO(q_j) = \int \prod_{i=1} q_i(\vartheta_i) (\log p(x, \vartheta) - \sum_k \log_k(\vartheta_k)) \quad (4.7)$$

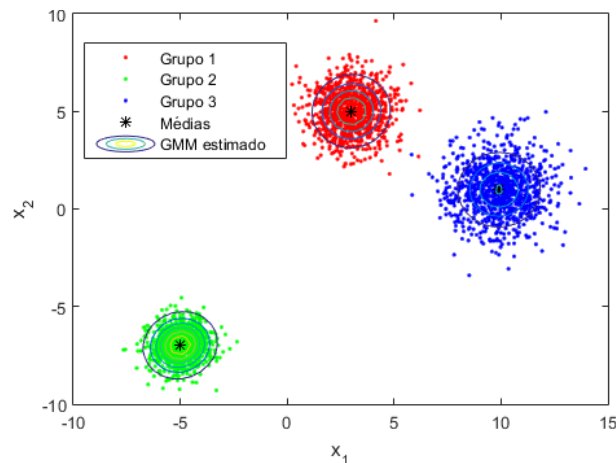
$$= -KL(q_j||\tilde{q}_j) + constant \quad (4.8)$$

onde

$$\tilde{q}_j \propto \exp \left(E_{-q_j} [\log p(x, \vartheta)] \right)$$

Isso motiva um esquema de atualização particular: iterar sobre as distribuições marginais q_j , maximizando o ELBO em cada etapa com $\tilde{q}_j$, e repetir. Esta forma irá então definir os parâmetros variacionais λ_i , e o algoritmo iterativo fornecerá uma maneira para calculá-los. A Figura 7 mostra um exemplo da estimação do GMM usando a Inferência Variacional.

Figura 7 – GMM estimado usando VI. O modelo está formado por três gaussianas.

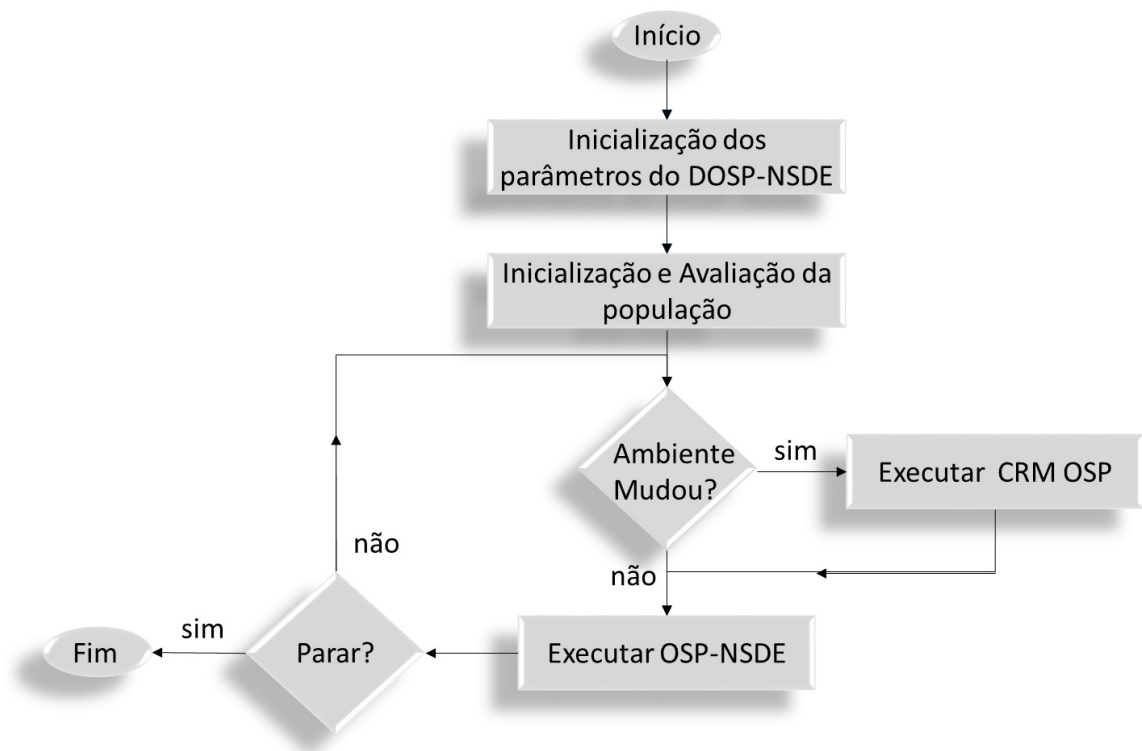


Fonte: Autor

4.2 EVOLUÇÃO DIFERENCIAL DE ORDENAÇÃO NÃO DOMINADA MELHORADA DINÂMICA COM A ESTRATÉGIA DE PREDIÇÃO DO ESPAÇO OBJETIVO, DOSP-NSDE

Esta seção descreve detalhadamente o algoritmo proposto, DOSP-NSDE (Figura 8). O algoritmo 1 descreve o DOSP-NSDE que recebe como entradas os parâmetros da evolução diferencial (F e CR), os parâmetros de mutação polinomial (p_m e η_m), a percentagem de variação do Hipervolume λ e o horizonte de previsão inicial^[1] $p_{inicial}$.

Figura 8 – Evolução Diferencial de Ordenação Não Dominada Melhorada Dinâmica com a Estratégia de Predição do Espaço Objetivo.



Fonte: Autor

Primeiramente, uma população inicial P^1 (de tamanho NP) é gerada e avaliada (linhas 2-3). Na linha 4, a primeira geração (t_{inic}) do conjunto de gerações consecutivas (Z), usada na estratégia do OSP no processo estático (dentro do MOEA), é definida. O conjunto de gerações consecutivas G que contém as gerações da mudança é definido na linha 5. Esse conjunto G será usado no mecanismo de reação à mudança (Algoritmo 3). Enquanto o critério de parada não for satisfeito, as linhas 6-16 serão executadas. Se uma mudança ambiental for detectada (linha 7), o CRM é acionado (linha 12) usando as informações de mudanças passadas. Nas linhas 9-11, o conjunto G e os parâmetros p e t_{inic} são atualizados. Então, a nova população é ordenada usando o conceito de dominância de Pareto (linha 14).

^[1] O horizonte de previsão é o período de tempo no futuro para o qual as previsões devem ser preparadas (STATISTICS, 2010). Também é chamado de período de previsão.

Algorithm 1: ALGORITMO GERAL DO DOSP-NSDE

```

1  $t \leftarrow 1$ ;
2  $P^t \leftarrow InicializacaoAleatoria(NP)$ ;
3  $(\phi_P)^t \leftarrow Avaliar(P^t)$ ;
4  $t_{inic} \leftarrow 2$ ;
5  $G \leftarrow \emptyset$ ;
6 while Critério de parada não atingido do
    /* Mecanismo de detecção de mudança */
7      $mudanca \leftarrow DeteccaoMudanca(P_t, (\phi_P)^t)$ ;
    /* Mecanismo de reação à mudança */
8     if  $mudanca = sim$  then
9          $G \leftarrow G \cup t$ ;
10         $p \leftarrow p_{inicial}$ ;
11         $t_{inic} \leftarrow t + 1$ ;
12         $[P^t, (\phi_P)^t] \leftarrow OSPCRM(\bigcup_{k \in G} P^k, \bigcup_{k \in G} (\phi_P)^k, \bigcup_{k \in G} (\rho_P)^k)$  (Algoritmo 3);
13    end
14     $(\rho_P)^t \leftarrow OrdenacaoNaoDominancia((\phi_P)^t)$ ;
    /* Algoritmo Evolucionário Multiobjetivo */
15     $[P^{t+1}, (\phi_P)^{t+1}] \leftarrow OSP - NSDE(P^t, (\phi_P)^t, (\rho_P)^t, p, \lambda, t_{inic}, F, CR, p_m, \eta_m)$ 
        (Algoritmo 4);
16 end
17 return  $P^T$ 

```

A seguir, o MOEA (OSP-NSDE) é aplicado no estágio estacionário (linha 15), Algoritmo 4. Neste MOEA, uma condição baseada na métrica de Hipervolume Aproximado (Equação 4.9) seleciona o método para gerar a nova população. Se o critério for atingido, o método OSP emprega as tendências de movimentos de soluções anteriores entre mudanças, no espaço objetivo, para gerar uma nova população. Sempre que o critério de disparo não for satisfeito (na grande maioria das gerações), o algoritmo emprega os operadores de variação para gerar a descendência. Em seguida, o processo de seleção ambiental ocorre para manter soluções não dominadas e esparsas. A saída P^T do algoritmo consiste nas soluções obtidas na última geração T do processo evolucionário.

A seguir, são discutidos os detalhes das etapas do DOSP-NSDE.

4.2.1 Mecanismo de Detecção da Mudança

O mecanismo de detecção de mudanças detecta se uma alteração ocorreu no ambiente e, em caso afirmativo, aciona uma estratégia para lidar com o novo cenário. Nesta tese, emprega-se a estratégia de detecção baseada em sensores proposta por (DEB; KARTHIK et al., 2007): reavaliação de um número fixo de sensores de paisagem durante o processo de busca. Em cada geração, (DEB; KARTHIK et al., 2007) propuseram a seleção aleatória de algumas soluções da população de potenciais pais para serem reavaliados. Se qualquer uma das funções de aptidão e/ou restrição for alterada, a modificação deve ser detectada pela reavaliação. Essa abordagem é fácil de implementar e amplamente usada na área.

4.2.2 Estratégia de Previsão no Espaço Objetivo

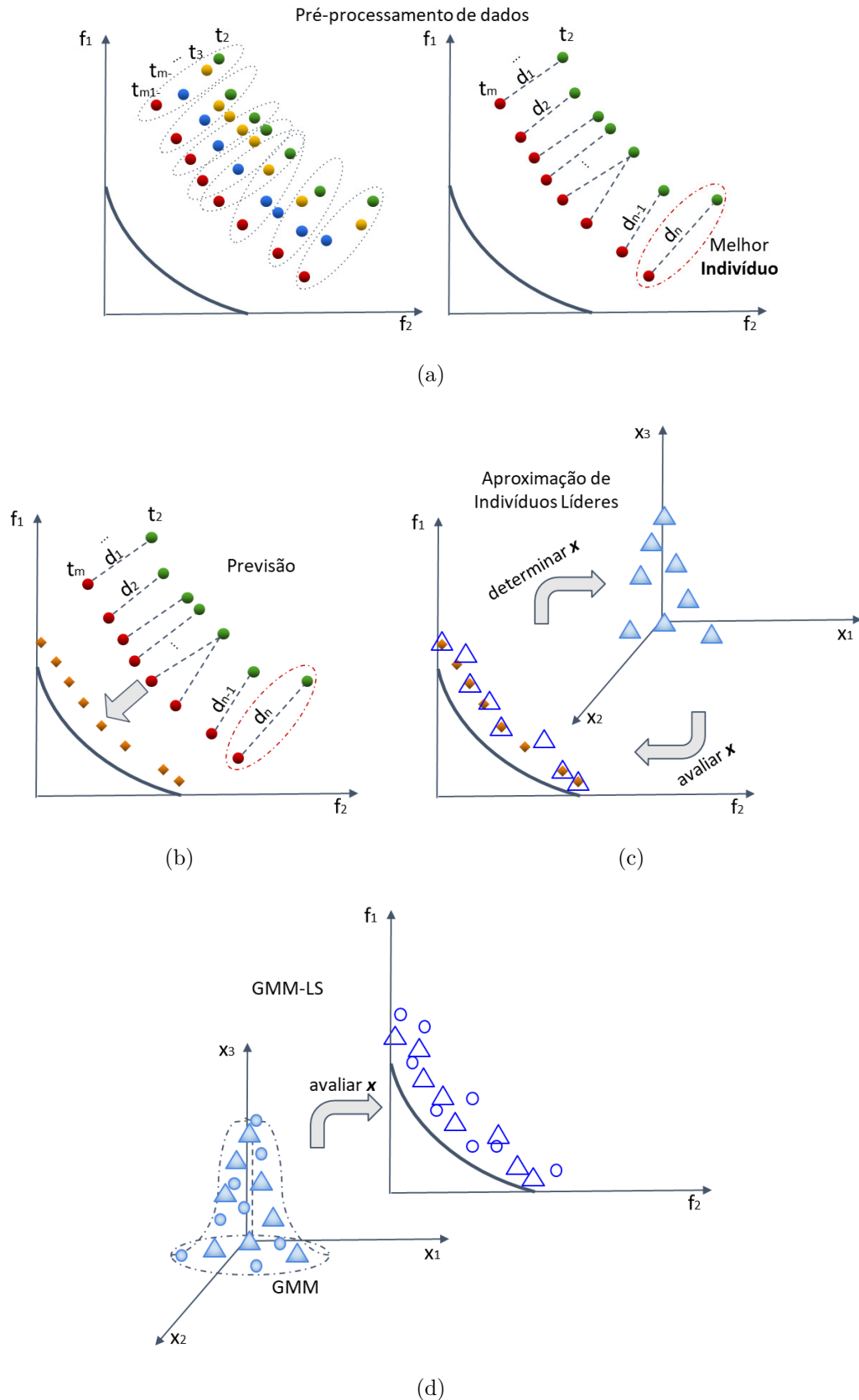
A estratégia de Previsão no Espaço Objetivo (OSP) visa acelerar a convergência do algoritmo, melhorando esperançosamente a qualidade da Frente de Pareto aproximada em termos de precisão e disseminação. O OSP consiste em quatro blocos principais (Ver Figura 9): 1) Pré-processamento de dados; 2) Previsão do movimento dos indivíduos não dominados usando modelos de séries temporais; 3) Busca de soluções no espaço de decisão que possuam valores de aptidão no espaço objetivo o mais próximo possível dos da frente prevista; 4) Completar a nova população usando modelos probabilísticos.

As etapas da estratégia do OSP são apresentadas no Algoritmo 2. Nele, sempre que um conjunto tiver uma letra subscrita, ou seja, i , ele se refere ao i -ésimo elemento deste conjunto. Quando o conjunto tem dois índices, (i, j) , representa o i -ésimo elemento da dimensão j . Enquanto isso, o sobrescrito indica a geração à qual uma variável de dados pertence. Este algoritmo recebe como entradas o horizonte de previsão p ; o tamanho da população NP ; o conjunto de indivíduos P^Z ; os valores das funções objetivo $(\phi_P)^Z$ dos indivíduos de P^Z ; e o $rank(\rho_P)^Z$ dos indivíduos de P^Z . Se a estratégia do OSP for incluída no MOEA, o conjunto Z é formado pelas gerações consecutivas $Z = \{t_{inic}, t_{inic+1}, \dots, t\}$. Por outro lado, se o procedimento OSP for usado como CRM, o conjunto Z será formado pelas gerações anteriores, onde aconteceram as mudanças do ambiente.

O Algoritmo 2 começa com o pré-processamento dos dados (linhas 1-14) que visa ordenar os dados para identificar melhor as tendências de movimento das soluções não dominadas. Na Figura 9 (a), os pontos com cores diferentes representam as gerações. Assim, o ponto verde e o ponto vermelho representam a geração inicial t_{inic} e a geração atual t , respectivamente. Inicialmente, os indivíduos não-dominados S e seus valores de aptidão (funções objetivo) ϕ_S , na geração atual t , são obtidos na linha 1. Para cada geração em Z , os valores das funções objetivo das soluções não dominadas são selecionadas e inseridas no conjunto ϕ (linha 3). Se ϕ tem um número diferente de elementos que ϕ_S (linha 4), os $|\phi_S|$ elementos de ϕ que estão mais perto de cada componente do conjunto ϕ_S , são escolhido usando a distância euclidiana (linhas 5-6), e armazenados no conjunto ϕ^j , com $1 < j < |Z|$. Caso contrário, todos os objetos no conjunto ϕ serão selecionados e estocados em ϕ^j (linha 8). Em seguida, todos os elementos de ϕ^j , em cada geração, são classificados em ordem crescente, considerando um dos objetivos selecionados aleatoriamente (linha 10). Na linha 12, o conjunto ϕ_{data} é formado pelos conjuntos ordenados.

Em seguida, o "indivíduo mais apto" do conjunto ϕ_{data} (linhas 13-14) - definido como a solução com o melhor movimento no espaço objetivo - é determinado. Os valores objetivos de cada solução ordenada da primeira e última geração são subtraídos. Os indivíduos Δ_{index} , cujo valor é positivo em todas as dimensões no resultado da subtração, são selecionados (linha 13). Para esses indivíduos escolhidos, a distância euclidiana entre a primeira e a última geração de ϕ_{data} é calculada. Finalmente, o indivíduo idx com a distância maior é selecionado (linha 14). Esse indivíduo é o padrão para criar o modelo de previsão. Caso

Figura 9 – Etapas da Estratégia de Predição no Espaço Objetivo. (a) Pré-processamento de dados; (b) Previsão do movimento dos indivíduos não dominados usando modelos de séries temporais; (c) Busca de soluções no espaço de decisão que possuam valores de aptidão no espaço objetivo o mais próximo possível dos da frente prevista; (d) Completar a nova população usando modelos probabilísticos.



Algorithm 2: ESTRATÉGIA DE PREVISÃO NO ESPAÇO OBJETIVO

```

/* Bloco 1: Pré-processamento de dados */
1  $[S, \phi_S] \leftarrow \text{OrdenacaoNaoDominancia}(P^t, (\phi_P)^t, (\rho_P)^t);$ 
2 for  $j \leftarrow 1$  to  $|Z| - 1$  do
3    $\phi \leftarrow \text{OrdenacaoNaoDominancia}((\phi_P)^{z_j}, (\rho_P)^{z_j});$ 
4   if  $|\phi| \neq |\phi_S|$  then
5      $\text{index} \leftarrow \underset{l \in [1, |\phi|]}{\text{argmin}} \|(\phi_S)_k - (\phi)_l\|_2^2, 1 \leq k \leq |\phi_S|;$ 
6      $\phi^j \leftarrow (\phi)_{\text{index}};$ 
7   else
8      $\phi^j \leftarrow \phi;$ 
9   end
10   $\phi'^j \leftarrow \text{Ordenar}(\phi^j, \text{rand}(m));$ 
11 end
12  $(\phi_{\text{data}})^Z \leftarrow \bigcup_{j \in Z} \phi'^j;$ 
   /* Determinação do "melhor indivíduo" */
13  $\Delta_{\text{index}} \leftarrow \{i \mid \sum_{j=1}^m [((\phi_{\text{data}})_{i,j}^1 - (\phi_{\text{data}})_{i,j}^{|Z|}) > 0] = m, 1 \leq i \leq |\phi_{\text{data}}|\};$ 
14  $\text{idx} \leftarrow \underset{l \in \Delta_{\text{index}}}{\text{argmax}} \|(\phi_{\text{data}})_l^1 - (\phi_{\text{data}})_l^{|Z|}\|_2^2;$ 

   /* Bloco 2: Previsão com o modelo AR */
   /* Dados da série temporal do melhor indivíduo */
15  $\phi_{\text{idx}} \leftarrow \{y_i \mid y_i \in (\phi_{\text{data}})^Z, i = \text{idx}\};$ 
   /* Estimar os parâmetros do modelo AR para o melhor indivíduo */
16  $\Phi \leftarrow \underset{\Phi}{\text{argmin}} \{\sum_{i=1}^{n_p} \Phi_i(\phi_{\text{idx}})^{n_p-i} + \varepsilon_{n_p}\} \text{ (Eq. 4.1), } n_p = I_{|\phi_{\text{data}}|};$ 
   /* Previsão com o modelo AR encontrado */
17  $(pf)_i \leftarrow (\Phi)^p (\phi_{\text{data}})^{n_i}, 1 \leq i \leq |\phi_{\text{data}}|; \text{ (Eq. 4.2)}$ 
18  $\phi_f \leftarrow \{(pf)_1, (pf)_2, \dots, (pf)_{|\phi_{\text{data}}|}\};$ 

   /* Bloco 3: Aproximação de soluções líderes */
19  $x_i \leftarrow \underset{x}{\text{argmin}} \|f(x) - (\phi_f)_i\|_2^2, \text{ with } x_0 = S_i, 1 \leq i \leq |\phi_f|;$ 
20  $X \leftarrow \{x_1, x_2, \dots, x_{|\phi_f|}\};$ 

   /* Bloco 4: Completar a população usando GMM-LS */
21  $k_s \leftarrow NP - |X|;$ 
   /* Determinar o GMM usando Inferência Variacional */
22  $p(x|\vartheta) \leftarrow \sum_{k=1}^K \theta_k N(X|\mu_k, S_k);$ 
23  $x_{1:k_s}^* \sim p(x|\vartheta);$ 
   /* Reparar indivíduos não factíveis */
24 if  $\text{lower}_{\text{bound}} > x_i^* > \text{upper}_{\text{bound}}$  then
25    $x_{i,d}^* \leftarrow \text{lower}_{\text{bound}};$ 
26    $x_{i,d}^* \leftarrow \text{upper}_{\text{bound}};$ 
27 end
28  $L \leftarrow x_{1:k_s,d}^*;$ 
   /* Próxima geração */
29  $P^t \leftarrow X \cup L;$ 
30  $(\phi_P)^t \leftarrow \text{Avaliar}(P^t);$ 
31 return  $[P^t, (\phi_P)^t]$ 

```

esse indivíduo não exista, um modelo AR é construído para cada indivíduo usando os seus movimentos nas gerações.

No segundo bloco (Figura 9 (b)), um modelo AR realiza um procedimento de previsão sobre o espaço objetivo (linhas 15-18). Para cada geração considerada, os dados referentes ao melhor indivíduo ϕ_{idx} são selecionados, onde $\phi_{idx} \subseteq \phi_{data}$ (linha 15). Em seguida, o movimento no espaço objetivo do idx é usado para estimar os parâmetros Φ do modelo AR (linha 16). Este modelo é usado para prever cada elemento de ϕ_{data} , onde os pontos da frente preditiva (pontos laranjas) $\phi_f = \{y_i | y_i \in S_{obj}, 1 \leq i \leq |\phi_S|\}$ são determinados p instantes à frente (linhas 17-18). Semelhante à Figura 9 (a), os pontos verdes e os pontos vermelhos representam a geração inicial t_{inic} e a geração atual t , respectivamente.

Uma vez obtida a frente prevista ϕ_f (pontos laranjas), uma otimização intermediária (Figura 9 (c)) é executada no espaço objetivo para determinar os "indivíduos líderes" ou "soluções líderes" ($X \leftarrow \{x_1, x_2, \dots, x_{|\phi_f|}\}$). Tais indivíduos (triângulo azul preenchido) são definidos como as soluções x_i que alcançam ou aproximam seus valores das funções objetivo $f(x_i)$ (triângulo azul não preenchido) como aqueles obtidos em ϕ_f , com $x \in [(x_{min})_d, (x_{max})_d]$ (linha 19). As soluções iniciais para o algoritmo de otimização são os indivíduos não dominados da geração atual. É importante observar que nem todos os indivíduos líderes encontrados têm os mesmos valores objetivos da frente prevista. Neste estudo, foi utilizado o método de programação quadrática sequencial (SQP) (NOCEDAL; WRIGHT, 2006) para a otimização de cada ponto previsto. O método SQP é um método determinístico baseado em gradientes que decompõe um problema não-linear complexo em uma série de problemas de programação quadrática pelo método de Lagrange-Newton. Com base nas características do gradiente descendente, o SQP apresenta velocidade de convergência rápida e alta eficiência. Durante cada processo de iteração, o SQP resolve um subproblema de programação quadrática, determina o comprimento do passo de acordo com uma função de mérito e repete essas etapas até que a solução do problema determinado seja determinada ou o número máximo de iterações seja atingido (ZHAO et al., 2018).

Finalmente, no bloco 4 (Figura 9 (d)), com os indivíduos $X \leftarrow \{x_1, x_2, \dots, x_{|\phi_f|}\}$ encontrados anteriormente (triângulo azul preenchido), uma versão melhorada do método de busca local *Gaussian Mixture Model based Local Search* (GMM-LS) (GUERRERO-PEÑA; ARAÚJO, 2017), gera $k_s = NP - |\phi_f|$ novos indivíduos (linha 21). Os parâmetros do GMM são determinados usando a Inferência Variacional (linha 22). k_s novas soluções são amostradas da distribuição do GMM para completar uma nova população (linha 23). Em seguida, soluções além dos limites do espaço de busca (soluções não factíveis) são reparadas (linhas 24-27). O método cria um conjunto de indivíduos (L) (linha 28) que são distribuídos esparsamente e compartilham recursos com os indivíduos previstos. Esses indivíduos são representados pelo círculo azul preenchido na figura. Os triângulos azuis não preenchidos e os círculos azuis não preenchidos representam os valores das funções

objetivo dos principais indivíduos e as soluções recém-criadas, respectivamente. O GMM-LS), é um esquema de Imigração Correlacionada para busca local guiada. Dada uma população X e um conjunto de imigrantes Y , a inserção de imigrantes é correlacionada quando Y apresenta dependência estatística com X , por exemplo os imigrantes são gerados por uma série de mutações de soluções anteriores (AZEVEDO; ARAÚJO, 2011b).

As saídas do algoritmo são a nova geração P^t , formada pela união do conjunto L e o conjunto de soluções líderes (X) (linha 29), e os valores dos objetivos $(\phi_P)^t$ (linha 30) das soluções P^t .

4.2.3 Mecanismo de Reação de Mudança (CRM)

O Mecanismo de Reação de Mudança deve responder com rapidez e precisão às mudanças no ambiente, para que a população possa convergir para um novo conjunto de soluções ótimas. O CRM proposto usa o método OSP para reiniciar a população na nova paisagem, baseando-se nas tendências dos movimentos da frente aproximada (PF) encontrada pelo algoritmo antes de que aconteça uma mudança no ambiente. O Algoritmo 3 apresenta as etapas do CRM proposto. Esse algoritmo recebe como entradas os indivíduos das frentes aproximadas encontrada pelo algoritmo ($\bigcup_{k \in G} P^k$), seus valores objetivos ($\bigcup_{k \in G} (\phi_P)^k$), e sua classificação ($\bigcup_{k \in G} (\rho_P)^k$) em cada mudança no conjunto G .

Algorithm 3: MECANISMO DE REAÇÃO DE MUDANÇA USANDO OSP

```

1 if  $|G| < 3$  then
2    $P^t \leftarrow \text{ReinicializacaoAleatoria}(NP)$ ;
3    $(\phi_P)^t \leftarrow \text{Avaliar}(P^t)$ ;
4 else
5    $[P^t, (\phi_P)^t] \leftarrow \text{OSP}(\bigcup_{k \in G} P^k, \bigcup_{k \in G} (\phi_P)^k, \bigcup_{k \in G} (\rho_P)^k)$  (Algoritmo 2);
6 end
7 return  $[P^t, (\phi_P)^t]$ 

```

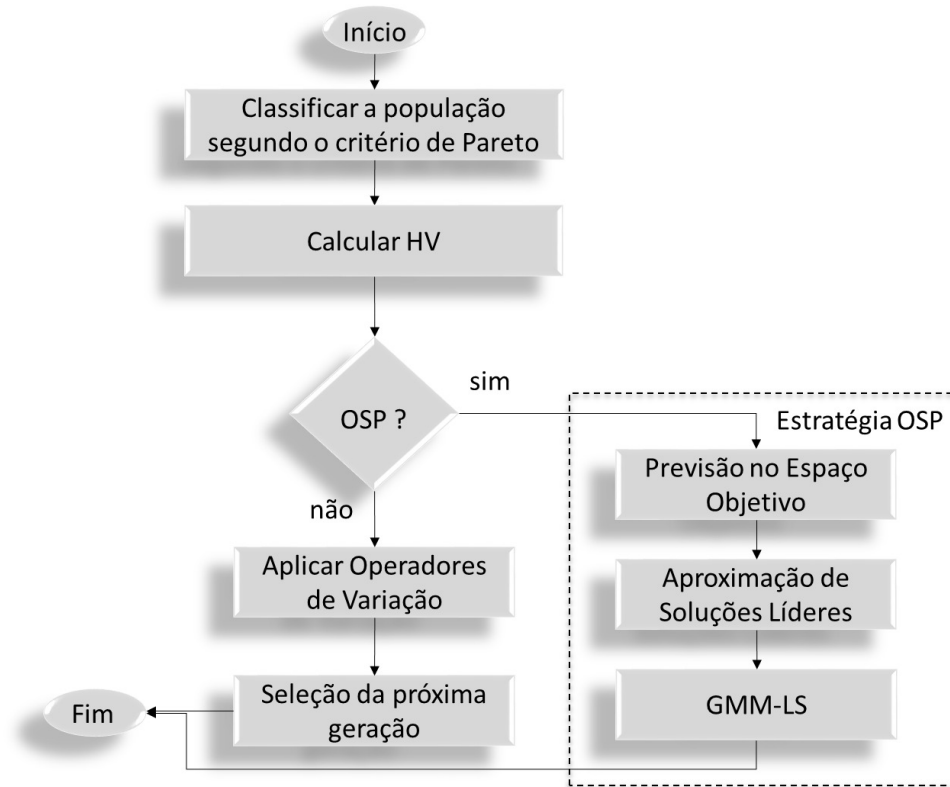
Quando o número de mudanças $|G|$ é menor que 3 (linha 1), indivíduos da população atual aleatoriamente escolhidos são substituídos por soluções geradas aleatoriamente (linha 2), e toda a população é avaliada (linha 3). Caso contrário ($|G| \geq 3$), a estratégia OSP (Algoritmo 2) é acionada para reiniciar a população (linha 5) usando a POF e o POS aproximados encontrados pelo OSP. Nesse caso, o horizonte de previsão (p) é definido como 1. A saída do algoritmo consiste nas novas soluções de população P^t e seus valores objetivos $(\phi_P)^t$.

4.2.4 Evolução Diferencial com Ordenação por Não Dominância com Previsão no Espaço Objetivo, OSP-NSDE

Quando a mudança de paisagem não é detectada, o *Non-dominated Sorting Differential Evolution improvement with Prediction in the Objective Space* (OSP-NSDE) é executado

(Figura 10). O OSP-NSDE pode ser caracterizado como um algoritmo baseado em Pareto e um algoritmo baseado em aprendizagem.

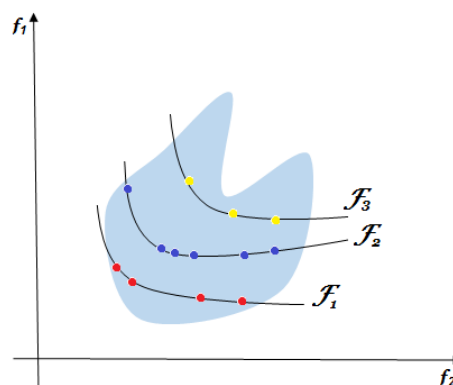
Figura 10 – Evolução Diferencial com Ordenação por Não Dominância com Previsão no Espaço Objetivo.



Fonte: Autor

O OSP-NSDE usa a ordenação elitista por dominância (*Pareto ranking*) (DEB et al., 2002). O procedimento de ordenação por dominância consiste em classificar as soluções de um conjunto M em diversas fronteiras $F_1, F_2, \dots, F_k$ conforme o grau de dominância de tais soluções. A Figura 11 ilustra este procedimento aplicado a soluções que minimizam duas funções objetivo (f_1 e f_2).

Figura 11 – Ordenação por dominância de Pareto.



Fonte: Adaptado de (DEB et al., 2002)

A aprendizagem do OSP-NSDE ocorre pela inclusão da estratégia OSP no algoritmo. O OSP insere informações úteis do passado no processo evolucionário para detectar regiões promissoras nos espaços de objetivo e de decisão. Na Figura 10, as linhas tracejadas delineiam os estágios da estratégia OSP.

Para acionar a estratégia OSP em determinadas gerações entre as mudanças, OSP-NSDE usa uma condição de previsão, com base no Hipervolume (WHILE et al., 2006). O procedimento do OSP é acionado se todas as condições a seguir forem atendidas:

1. A diferença entre as gerações t e t_{inic} é maior ou igual a 3. Isso ocorre porque pelo menos três gerações são necessárias (determinadas heurísticamente) para estabelecer o caminho do processo evolucionário com o modelo AR.
2. O número de soluções não dominadas ($|\phi_S|$) é menor que uma porcentagem β do tamanho da população.
3. O $(hv)_t$ na geração t deve aumentar em uma porcentagem λ em relação ao HV da geração inicial $(hv)_{t_{inic}}$.

O critério para acionar o método OSP é:

$$g = \begin{cases} 1; & (t - t_{inic}) \geq 3 \ \& \ |\phi_S| < \beta \cdot NP \ \& \ (hv)_t \geq \alpha \\ 0; & \text{caso contrário} \end{cases} \quad (4.9)$$

onde:

$$\alpha \leftarrow (hv)_{t_{inic}} + \lambda \cdot (hv)_{t_{inic}} \quad (4.10)$$

Para cada geração, o ponto de referência para o cálculo do HV é escolhido como o valor máximo possível de cada objetivo. Embora o valor do HV seja sensível a esse ponto de referência, neste caso a precisão não é tão relevante pois o cálculo HV visa estimar o progresso processo evolucionário.

Depois que a regressão é executada, os parâmetros são atualizados da seguinte forma:

- Geração inicial para a próxima previsão: $t_{inic} = t$,
- Limiar α ajustado de acordo com a Equação 4.10,
- Horizonte de previsão: $p = p/2$.

O parâmetro p varia conforme o OSP é executado, pois o processo evolucionário tende a se aproximar da frente real a cada geração. Portanto, a previsão é feita para poucas gerações à frente, para gerar indivíduos próximos da frente real.

O Algoritmo 4 apresenta os passos do OSP-NSDE. Esse algoritmo recebe como entradas os indivíduos da geração atual (P^t), seus valores objetivos $((\phi_P)^t)$, seus *ranks* $((\rho_P)^t)$, os parâmetros dos operadores de variação do DE (F e CR), os parâmetros de mutação

Algorithm 4: ALGORITMO EVOLUCIONÁRIO MULTI-OBJETIVO: OSP-NSDE

```

1  $\phi_S \leftarrow SolNaoDominadas((\phi_P)^t, (\rho_P)^t);$ 
2  $(hv)_t \leftarrow CalculoHV(\phi_S);$ 
   /* Cálculo do limiar do critério de previsão */
3 if  $t = t_{inic}$  then
4   |  $\alpha \leftarrow (hv)_{t_{inic}} + \lambda \cdot (hv)_{t_{inic}}$  (Eq. 4.10);
5 end
6 if  $(t - t_{inic}) \geq 3$  &  $|\phi_S| < 0.9 \cdot NP$  &  $(hv)_t \geq \alpha$  (Eq. 4.9) then
7   |  $Z \leftarrow \{t_{inic}, t_{inic+1}, \dots, t_{end-1}, t\};$ 
   /* Executar a estratégia OSP */
8   |  $[P^{t+1}, (\phi_P)^{t+1}] \leftarrow EstrategiaOSP(\bigcup_{k \in Z} P^k, \bigcup_{k \in Z} (\phi_P)^k, \bigcup_{k \in Z} (\rho_P)^k, p)$ 
   (Algoritmo 2);
   /* Atualizar */
9   |  $t_{inic} \leftarrow t;$ 
10  |  $p \leftarrow p/2;$ 
11 else
12  |  $Q \leftarrow OperadoresVariacao(F, CR, p_m, \eta_m, P^t);$ 
13  |  $\phi_Q \leftarrow Avaliar(Q);$ 
14  |  $[P^{t+1}, (\phi_P)^{t+1}] \leftarrow SelecaoAmbiental(P^t \cup Q, (\phi_P)^t \cup \phi_Q, NP);$ 
15 end
16 return  $[P^{t+1}, (\phi_P)^{t+1}]$ 

```

polinomial (p_m e η_m), o percentual de variação do Hipervolume λ , a primeira geração da previsão t_{inic} e o horizonte de previsão p .

Inicialmente, o valor do HV é calculado (linha 2) para o conjunto de soluções não dominadas (ϕ_S) da geração atual t (linha 1). Se a geração (t) atingir t_{inic} , o limiar α do critério de previsão será calculado (linhas 3-5). Se o critério de disparo de previsão (linha 6) é satisfeito, as linhas 7-10 são executadas. Dado o conjunto de gerações consecutivas Z (linha 7), o Algoritmo 2 (linha 8) é executado considerando todos os indivíduos de Z , seus valores de funções objetivo e seus *ranks*. Em seguida, os parâmetros do critério do OPS são atualizados (linha 9-10).

Se a condição de previsão não for acionada (linha 11), toda a população é selecionada como pais. Os descendentes são gerados (linha 12) usando os operadores de variação do DE (STORN; PRICE, 1997): a mutação diferencial (Equação 4.11) e a recombinação binária (Equação 4.12). A operação de mutação gera novos vetores de parâmetros através da adição da diferença ponderada entre dois vetores de parâmetros a um terceiro indivíduo. Para cada indivíduo x_i , o vetor mutado v_i é gerado por meio da seguinte relação:

$$v_i = x_{r1} + F(x_{r3} - x_{r2}) \quad (4.11)$$

onde x_{r1} , x_{r2} e x_{r3} são três indivíduos diferentes selecionados aleatoriamente e diferentes do indivíduo x_i . Na mutação, a diferença entre os indivíduos é usada como o passo de mutação, enquanto o fator F determina o tamanho do passo. F é uma constante real $F \in [0, 2]$.

Os vetores de parâmetros mutados criam uma população V que é combinada com os vetores da população X . Esta combinação de parâmetros é referida como cruzamento em DE e pode ser binária ou exponencial (LI et al., 2015), com a finalidade de aumentar a diversidade dos vetores de parâmetros mutados. A operação de recombinação de cada par de indivíduos (v_i, x_i) para gerar um novo indivíduo p_i é:

$$p_{ij} = \begin{cases} v_{ij}, & \text{se } rand(0, 1) \leq CR | j = jrand \\ x_{ij}, & \text{caso contrário, para } j = 1, \dots, D \end{cases} \quad (4.12)$$

onde $rand(0, 1)$ é um número aleatório uniformemente distribuído no intervalo $[0, 1]$, D é o número de dimensões, $CR \in [0, 1]$ é a taxa de cruzamento, a qual controla quantas dimensões do novo indivíduo gerado pelo cruzamento são do vetor mutado v_i , e $jrand$ é um índice selecionado aleatoriamente para que pelo menos 1-D do vetor mutado esteja no novo indivíduo gerado.

Na próxima etapa, o operador de mutação polinomial (DEB; AGRAWAL, 1999), com parâmetro de índice (η_m), é aplicado. Nesse operador, uma distribuição de probabilidade polinomial é usada para perturbar uma solução na vizinhança de um dos pais. A distribuição de probabilidade na esquerda e na direita de um valor variável é ajustada para que nenhum valor fora do intervalo especificado $[a, b]$ seja criado pelo operador de mutação. Para uma dada solução pai $p \in [a, b]$, a solução mutada p' para uma variável específica é criada para um número aleatório u criado dentro de $[0, 1]$, como segue:

$$p' = \begin{cases} p + \bar{\delta}_L(p - x_{min}^i), & \text{para } u \leq 0.5, \\ p + \bar{\delta}_R(x_{max}^i - p), & \text{para } u > 0.5 \end{cases} \quad (4.13)$$

onde:

$$\bar{\delta}_L = (2u)^{1/(1+\eta_m)} - 1 \quad (4.14)$$

e

$$\bar{\delta}_R = 1 - (2(1 - u))^{1/(1+\eta_m)} \quad (4.15)$$

Os indivíduos p' criam o conjunto de filhos Q , que são avaliados na linha 13. Ambas as populações, P^t e Q , são combinadas e usadas no processo de seleção de nova população que é executado (linha 14). Este processo cria a próxima geração P_{t+1} usando o princípio de dominância de Pareto, e o método de truncamento de arquivo proposto em (ZITZLER; LAUMANN; THIELE, 2001). Este mecanismo de diversidade baseado nos k -ésimos vizinhos mais próximos, remove iterativamente os indivíduos de P_{t+1} até que $|P_{t+1}| = NP$. Em cada iteração, o indivíduo i é escolhido para remoção se $i \leq_d j$ para todo $j \in P_{t+1}$ com

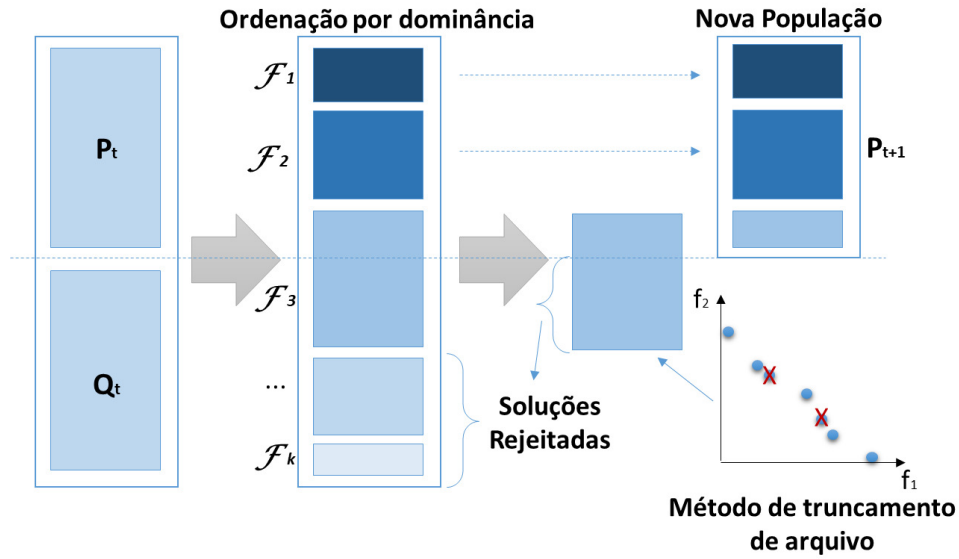
$$\begin{aligned} i \leq_d j : \iff \quad & \forall 0 < k < |P_{t+1}| : \sigma_i^k = \sigma_j^k \quad \vee \\ & \exists 0 < k < |P_{t+1}| : [(\forall 0 < l < k : \sigma_i^l = \sigma_j^l) \wedge \sigma_i^k < \sigma_j^k] \end{aligned} \quad (4.16)$$

onde σ_i^k denota a distância do indivíduo i ao seu k -ésimo vizinho mais próximo em P_{t+1} . Em outras palavras, o indivíduo que tem a distância mínima para outro indivíduo é

escolhido em cada iteração; se houver vários indivíduos com distância mínima, o empate será quebrado considerando-se as segundas menores distâncias e assim por diante. Tal método permite que os mais diversos indivíduos ocupem as últimas posições classificadas, espalhando o conjunto de soluções.

A saída do algoritmo consiste nas soluções da nova população P^{t+1} e seus valores objetivos $(\phi_P)^{t+1}$. O processo de seleção da próxima geração é ilustrada na Figura 12.

Figura 12 – Seleção da próxima geração



Fonte: Adaptado de (DEB et al., 2002)

4.2.5 Complexidade Computacional do DOSP-NSDE

Para uma geração do DOSP-NSDE, os principais custos computacionais estão envolvidos na estratégia do OSP (linha 5 do Algoritmo 3 e linha 8 do Algoritmo 4). Os outros procedimentos têm um custo computacional menor. Dados o número de funções objetivo M , a dimensão do indivíduo D e o tamanho da população N , as operações básicas e suas complexidades são:

1. A complexidade do processo de detecção de mudanças é $O(N)$.
2. O mecanismo de reação à mudança gasta $O(N^3)$ cálculos. A complexidade é dada pelo procedimento de Aproximação de Soluções Líderes na estratégia OSP.
3. A complexidade computacional total do OSP-NSDE ($O(N^3)$) também é dada pelo procedimento de Aproximação de Soluções Líderes.
4. O operador de truncamento tem complexidade $O(N^2 \log N)$, já que os indivíduos diferem em relação ao segundo ou terceiro vizinho mais próximo e, portanto, a classificação das distâncias rege a complexidade geral (ZITZLER; LAUMANN; THIELE, 2001).

5. A complexidade dos procedimentos da estratégia OSP (Algorithm 2) é a seguinte:

- Pré-processamento de dados: $O(|Z| \cdot M \cdot N)$, onde $|Z|$ é o número de gerações usadas no método OSP.
- Determinação do "melhor indivíduo": $O(N \cdot M)$.
- Previsão com o modelo AR: A estimativa dos parâmetros do modelo AR gasta $O(N \cdot M^2)$ cálculos no método dos mínimos quadrados, e a previsão gasta $O(N)$ cálculos.
- Aproximação de Soluções Líderes: Esta otimização intermediária usando o método SQP tem alta complexidade ($O(I \cdot N^3)$). Nele, a maior complexidade computacional do SQP é o número de passos (I) necessárias multiplicado pelo custo de resolver o QP convexo ($O(N^3)$). Outros métodos menos complexos podem ser empregados neste processo. No entanto, mais investigações são necessárias para analisar o desempenho do OSP-NSDE com esses outros métodos.
- Completando a população usando o GMM-LS: Determinar os parâmetros do GMM usando a VI requer $O(L \cdot D)$ cálculos, e reparar as soluções não viáveis exige $O(N \cdot D)$ cálculos.

É importante notar que a estratégia de OSP é pouco usada no EA multiobjetivo devido ao critério baseado em HV, uma vez que permite poucas execuções desta estratégia. Por conta disso, o pior caso raramente ocorre. Em suma, o OSP é executado apenas quando pode ser mais útil para o processo evolucionário.

5 RESULTADOS EXPERIMENTAIS

O DOSP-NSDE foi validado através de simulações, usando um número de benchmarks muito utilizado na literatura, e foi comparado com vários DMOEAs baseados em predição para um conjunto de DMOPs selecionados. A primeira parte do capítulo está dedicada a analisar o comportamento do DOSP-NSDE ante às mudanças dos seus parâmetros (análise da sensibilidade) e a analisar a influência da frequência e da severidade das mudanças da paisagem no desempenho do algoritmo proposto. Todas as abordagens foram independentemente executadas 20 vezes. O tamanho da população (NP) foi definido para 100 indivíduos de dimensão $D = 20$. O teste de Friedman (FRIEDMAN, 1937) e o pós-teste de Nemenyi (NEMENYI, 1962) foram realizados para indicar significância entre diferentes resultados no nível de significância de 0,05.

Para estes experimentos foram escolhidos nove DMOPs bi-objetivos (DF1-DF9) e cinco DMOPs de três objetivos (DF10-DF14) do conjunto de testes de competição CEC2018 (JIANG et al., 2018). Essas funções têm características que também podem ser frequentemente encontrados em problemas do mundo real, alguns deles são PF/PS geometrias dependentes do tempo, formas de PF irregulares, como desconexão e joelho (*knee*). Além disso, funções como FDA1-FDA3 e dMOP2 (DMOPs bi-objetivos) foram selecionadas do conjunto de testes FDA (FARINA; DEB; AMATO, 2004) e dMOP (GOH; TAN, 2009), respectivamente. Além disso, UDF1, UDF2, UDF4 e UDF5 foram escolhidos do conjunto de testes UDF (BISWAS et al., 2014), onde para esses dois últimos problemas de teste há uma relação de correlação não-linear entre as variáveis de decisão. Para avaliar o desempenho DOSP-NSDE foram escolhidas várias métricas amplamente usadas. Assim, se calcularam os valores médios da Distância Geracional Invertida (IGD) (WANG; LI, 2009), HiperVolume (HV) (ZITZLER; BROCKHOFF; THIELE, 2007) e da métrica de Espaçamento de Schotts (SP) (SCHOTT, 1995) para a comparação do algoritmo proposto com os outros algoritmos.

Outros experimentos foram realizados para compreender melhor o método OSP. Assim, neste capítulo apresenta-se uma seção dedicada a análise da estratégia quando ela é empregada como CRM, e a influência do OSP quando é usada como parte do otimizador multiobjetivo (OSP-NSDE). Nesta última parte, o OSP-NSDE foi validado através de simulações usando alguns problemas de teste MOPs bem conhecidos. Foram escolhidos sete MOPs de dois objetivos (UF1-UF7) e três MOPs de três objetivos (UF8-UF10) do conjunto de testes de UF. Os problemas de Zitzler-Deb-Thiele (ZDT), consistindo de cinco MOPs de dois objetivos (ZDT1-ZDT4 e ZDT6), também foram empregados no estudo. Por último, os experimentos foram realizados em dois outros conjuntos de testes conhecidos com três objetivos: DTLZ1-7 e WFG1-9. OSP-NSDE foi comparado com onze algoritmos de última geração sobre MOPs selecionados. Os experimentos foram execu-

tados usando PlatEMO, uma plataforma livre e baseada em MATLAB para otimização multiobjetivo evolucionária (TIAN et al., 2017). Três métricas de desempenho, Distância Geracional Invertida (IGD), HiperVolume (HV) e *Performance Metrics Ensemble* (HE; YEN, 2013) foram empregadas para avaliar o OSP-NSDE e os outros algoritmos. No caso do *Performance Metrics Ensemble*, nove indicadores^[1] codificados em PlatEMO são usados para calcular a classificação dos algoritmos. Todas as referências para os problemas de teste, métricas de desempenho e algoritmos usados em experimentos podem ser encontradas em (TIAN et al., 2017). Todas as abordagens foram executadas 30 vezes independentes para um número máximo de 300000 avaliações de funções. O tamanho da população foi definido para 100 indivíduos. O teste da soma de postos de Wilcoxon foi escolhido para comparar os resultados obtidos pelo OSP-NSDE e os outros algoritmos ao nível de significância de 0,05.

Finalmente, o capítulo apresenta os resultados dos experimentos em problemas reais. Todos os experimentos foram executados em um notebook com CPU de 2,5GHz, 8GB de RAM e sistema operacional Windows10. A Tabela 4 resume os diferentes experimentos e os problemas, métricas e configuração paramétrica usados neste capítulo.

Tabela 4 – Condições Experimentais.

Experimento	Secção	Problemas	Métricas	Configuração Experimental	Teste estatístico
*Análise da sensibilidade	5.1	DF1-14	IGD	Execuções: 20	
*Influência da frequência e da severidade das mudanças da paisagem	5.2	FDA1-3 dMOP2 UDF1,2,4,5	HV SP	NP: 100 D: 20	Teste de Friedman e Nemenyi
*OSP como CRM	5.3.1				
*Comparação do DOSP-NSDE com DMOEAs baseados em predição	5.4				
*OSP dentre do otimizador (OSP-NSDE)	5.3.2	UF1-7 ZDT1-4,6 DTLZ1-7 WFG1-9	IGD HV <i>Ensemble</i>	Execuções: 30 NP: 100 D: muda com o problema	<i>Wilcoxon rank sum test</i>

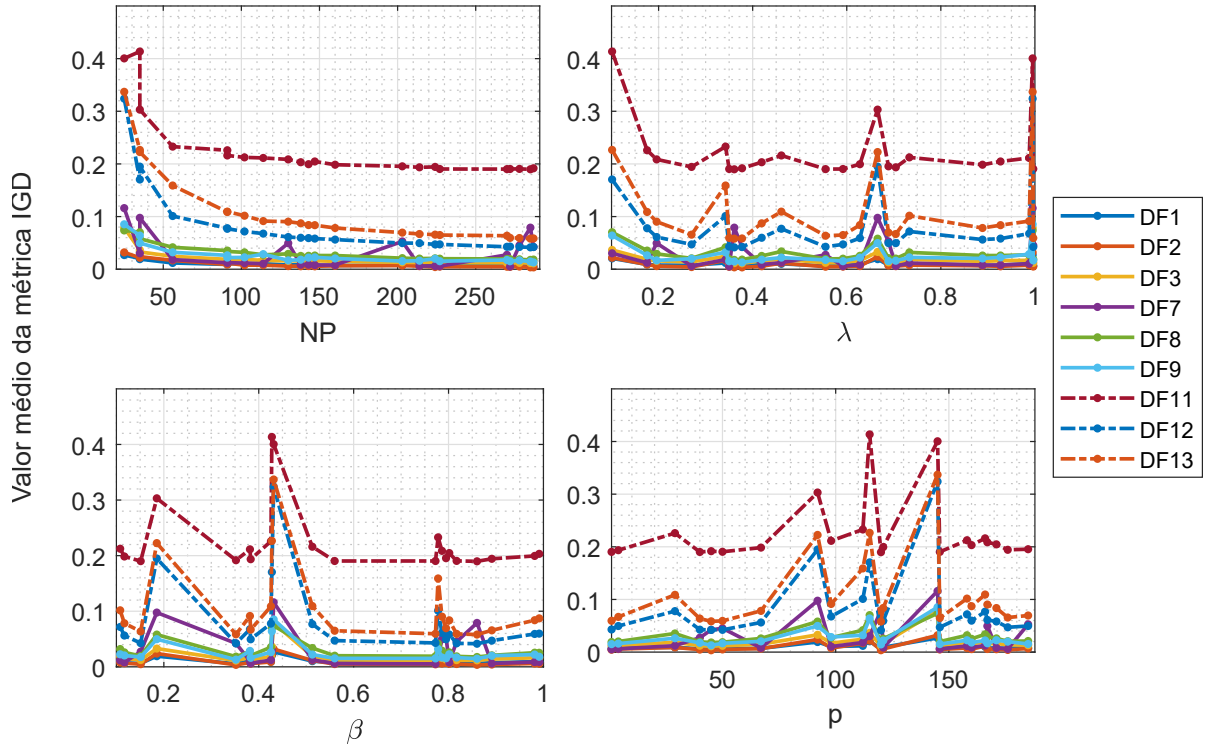
[1] Os indicadores de desempenho são Métrica de Cobertura, Métrica para Diversidade, Distância Geracional, HiperVolume, Distância Geracional Invertida, HiperVolume Normalizado, Diversidade Pura, Espaçoamento e Spread.

5.1 ANÁLISE DA SENSIBILIDADE DOS PARÂMETROS DO DOSP-NSDE

O primeiro experimento teve como objetivo avaliar a influência dos parâmetros do DOSP-NSDE na convergência do algoritmo. Para este estudo, foi usado o *benchmark* DF e foi analisado o desempenho do algoritmo alterando quatro parâmetros: 1) O tamanho da população NP ; 2) a percentagem de variação λ do Hipervolume; 3) a percentagem de variação β , ambas na condição baseada em Hipervolume que aciona o mecanismo OSP no processo estacionário; e 4) o horizonte inicial de previsão p . A análise paramétrica envolve os seguintes ranges: $NP \in [20; \dots; 300]$, $\lambda \in [0, 1; \dots; 1]$, $\beta \in [0, 1; \dots; 1]$ e $p \in [1; \dots; 200]$. Os valores dos parâmetros foram determinados usando *Latin Hypercube* (MCKAY; BECKMAN; CONOVER, 1979).

Cada simulação tem como objetivo responder a $N_c = 30$ mudanças ambientais. O DOSP-NSDE evolui por 50 gerações antes da primeira mudança. A frequência de mudança τ_t está definida para 10. A Figura 13 mostra o valor médio da métrica IGD para cada valor dos parâmetros. As linhas tracejadas representam as funções tri-objetivas, e as linhas sólidas mostram as funções com dois objetivos.

Figura 13 – Análise da sensibilidade dos parâmetros do DOSP-NSDE ($n_t = 10$, $\tau_t = 10$)



Fonte: Autor

A análise da variação de NP mostra que o desempenho do DOSP-NSDE cresce dentro do intervalo de 1 a 100. Além desse valor, a métrica IGD mantém pouca variabilidade para a maioria das funções quando o tamanho da população aumenta (Figura 13a). Devido ao custo computacional, foi selecionado $NP = 100$ como tamanho da população para todas as funções.

O valor do parâmetro $\lambda = 0,6$ parece alcançar os melhores desempenhos considerando o seu range (Figura 13b). Para os valores extremos $(0, 1; 1)$, o algoritmo proposto alcança os maiores valores de IGD. No caso de um valor de $0,1$, a estratégia do OSP é executada com frequência, mesmo nas últimas gerações. Como consequência, a diversidade é aumentada significativamente, aumentando assim a probabilidade da presença de soluções não dominadas de menor qualidade. Se $\lambda = 1$, o OSP raramente é executado; portanto, o mecanismo de diversidade não é acionado, causando o crescimento dos valores de IGD. Em suma, muita diversidade ou falta dela pode prejudicar o desempenho do OSP.

O algoritmo apresenta baixa sensibilidade ao parâmetro $\beta = 0,9$ para a maioria das funções de teste (Figura 13c). Para valores pequenos de β , a estratégia do OSP pode não ser acionada porque é fácil alcançar essa percentagem de soluções não dominadas. Vale ressaltar que o desempenho do algoritmo diminui para a vizinhança do valor máximo deste parâmetro. Assim, o valor de $\beta = 90\%NP$ parece ser uma escolha adequada. À medida que o processo evolucionário avança, os avanços da frente se tornam mais curtos, de modo que o processo de regressão pode se tornar desnecessário.

A análise do parâmetro p mostra que o aumento de p não indica nenhuma tendência regular (Figura 13d). As melhores opções ocorrem para valores próximos aos extremos. Os valores baixos de p causam um passo pequeno na previsão e, portanto, a velocidade de convergência pode diminuir. No caso de alto valor de p , a frente prevista pode ser estimada longe da frente ótima. Percebe-se que $p = 40$ alcança o menor valor IGD para a maioria das funções. Esse valor para o parâmetro p parece ser útil para definir um equilíbrio entre a velocidade de convergência e a etapa de previsão.

Com base nessa análise, definimos $NP = 100$, $\lambda = 0,6$, $\beta = 0,9$ e $p = 40$ para os experimentos a serem descritos.

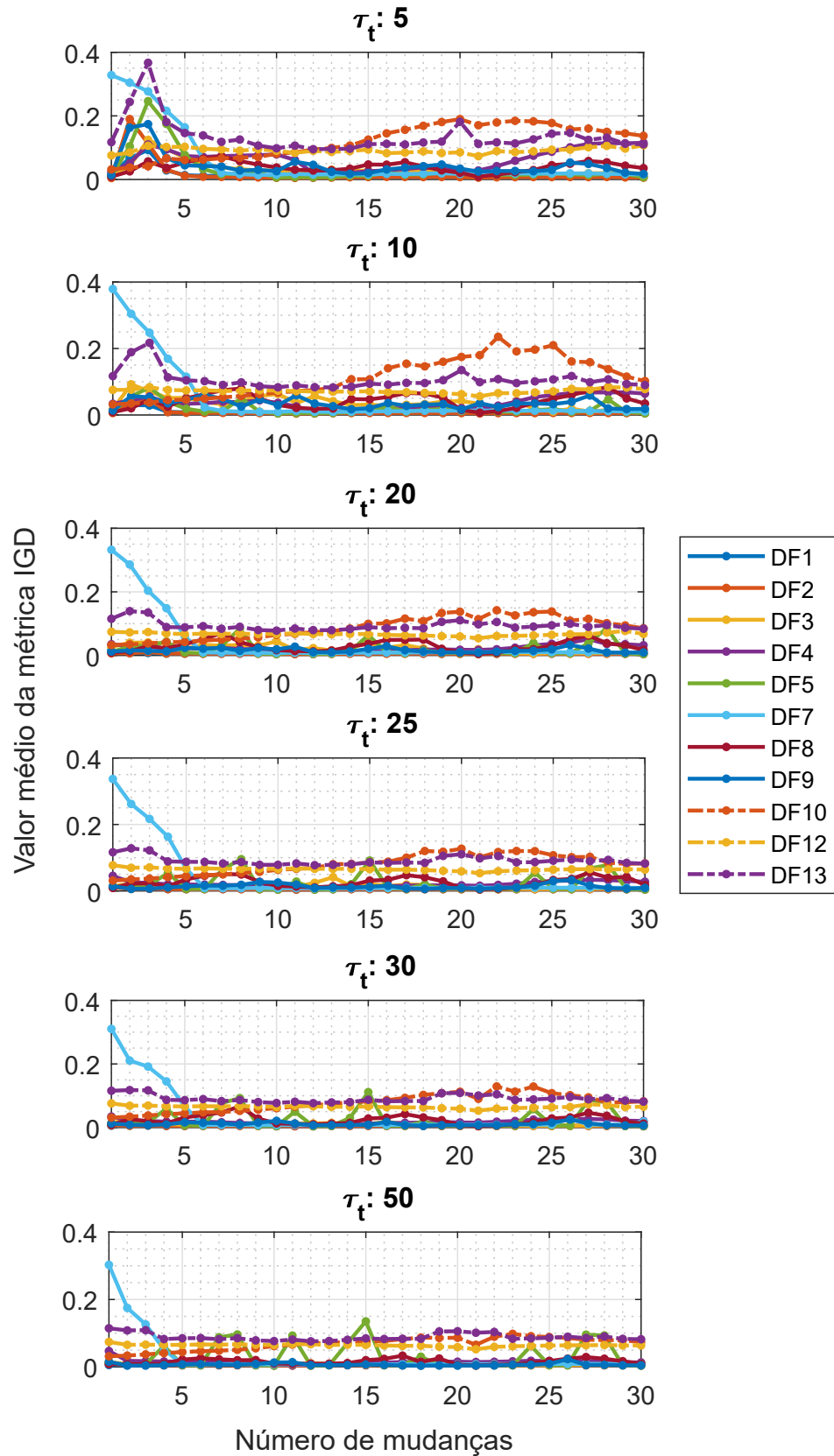
5.2 INFLUÊNCIA DA FREQUÊNCIA E DA SEVERIDADE DA MUDANÇA NO DOSP-NSDE

Para estudar o impacto da frequência de mudança na capacidade do DOSP-NSDE de responder adequadamente às mudanças nos ambientes, foi selecionado o conjunto de teste DF, no qual a severidade da mudança (n_t) foi fixada em 10 e a frequência das mudanças varia $\tau_t = [5, 10, 20, 25, 30, 50]$. O número de mudanças do ambiente foi definido como $N_c = 30$.

A Figura 14 mostra o IGD médio sobre as mudanças para cada τ_t . Quando a frequência da mudança aumenta, a proposta melhora a qualidade das soluções e pode rastrear mudanças ambientais para a maioria das funções de teste. Para um valor pequeno de τ_t , o rastreamento da POF parece ser mais difícil, como pode se ver nas figuras correspondentes a $\tau_t = 5$ e $\tau_t = 10$.

Além disso, pode se observar que para os DMOPs de três objetivos, como DF10, DF12 e DF13, o desempenho da proposta diminui, comparado a outras funções de teste,

Figura 14 – Influência da frequência da mudança.



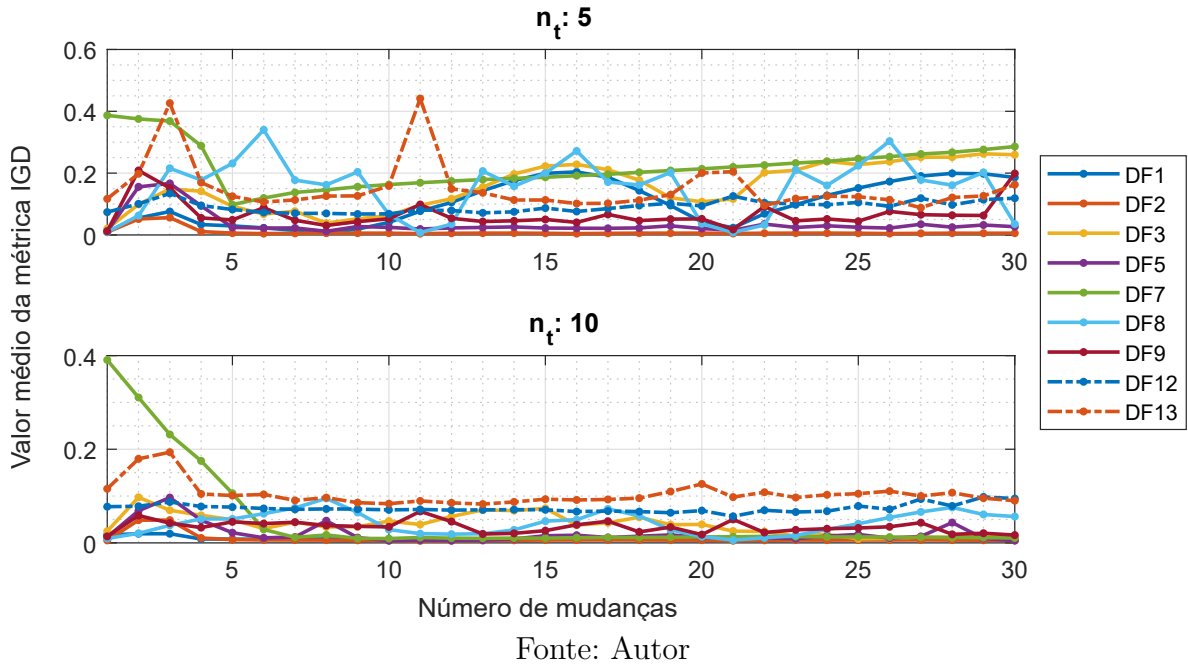
Fonte: Autor

sugerindo que uma melhoria é necessária para lidar com problemas de otimização de muitos objetivos em ambientes dinâmicos.

Por fim, é analisado o efeito dos níveis de severidade no desempenho do DOSP-NSDE, usando os problemas DF com $\tau_t = 10$ considerando mudanças ambientais severas $n_t = 5$ e moderadas $n_t = 10$. O número de alterações foi definido como $N_c = 30$.

A Figura 15 mostra os valores médios do IGD sobre as mudanças para cada n_t . Para uma mudança moderada, pode-se ver que a proposta funciona melhor, respondendo adequadamente às mudanças ambientais para todos os problemas. Para mudanças severas, os resultados parecem ser piores para a maioria das funções de teste. Estes resultados sugerem que a proposta é sensível à severidade da mudança. Analogamente ao caso com dois objetivos, o desempenho cai para problemas de três objetivos.

Figura 15 – Influência da severidade da mudança.



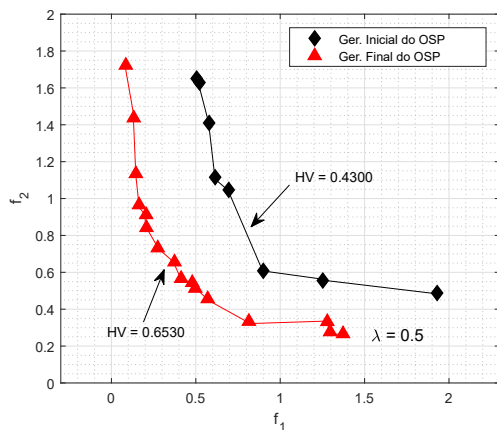
5.3 ANÁLISE DA ESTRATÉGIA OSP

Como mencionado no Capítulo 4, a estratégia OSP é acionada segundo um critério baseado em HV. A Figura 16 mostra um exemplo do critério para a execução do OSP com $\lambda = 0,5$. Os pontos com forma de diamantes e triângulos representam as soluções não dominadas das gerações inicial e final usadas no processo de regressão, respectivamente. Nestas duas gerações, o valor de HV é 0,4300 e 0,6530, respectivamente. Pode-se observar que o HV aumentou 51,86%, e as duas frentes esboçadas apresentam um movimento populacional real no espaço objetivo, demonstrando a efetividade do critério.

Quando o critério baseado em HV é atingido, o OSP é accionado. A Figura 17 ilustra o uso da previsão baseada em AR para encontrar a nova frente prevista. Os círculos e

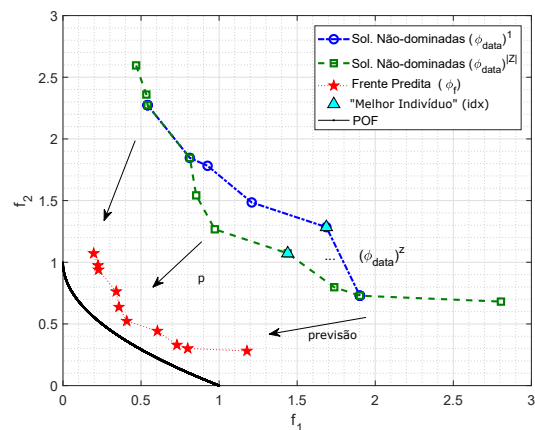
quadrados representam as soluções não dominadas das gerações inicial e final, respectivamente, do conjunto de populações usado no procedimento de previsão, ou seja, as amostras para estimar a próxima frente. Os triângulos mostram o "melhor indivíduo" nesse intervalo. Usando esse indivíduo, o modelo AR prevê cada solução em p gerações à frente. Os pontos estrelas retratam as soluções da frente prevista, isto é, a previsão resultante produzida pelo modelo AR.

Figura 16 – Exemplo do funcionamento do critério para o acionamento do OSP.



Fonte: Autor

Figura 17 – Etapa de previsão no espaço objetivo do OSP.



Fonte: Autor

A frente prevista no espaço objetivo é usada na geração das "soluções líderes", Figura 18. As estrelas mostram as soluções da frente prevista e os triângulos representam os valores objetivos das soluções líderes. Pode-se ver que nem todos os indivíduos líderes encontram-se na frente prevista, embora essa diferença não prejudique o método. Em suma, o objetivo é encontrar os indivíduos o mais próximo possível da região promissora do espaço objetivo usando o SQP.

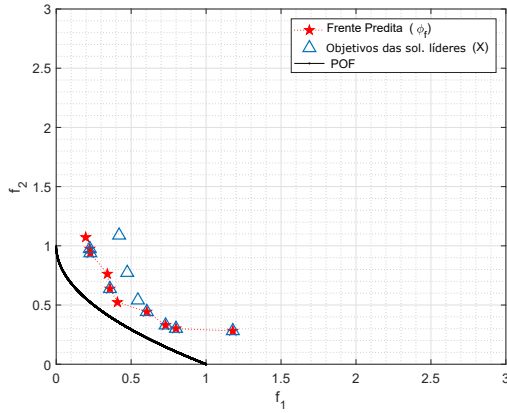
A Figura 19 mostra o resultado da aplicação do GMM-LS para completar a nova população. Pode-se observar os novos indivíduos (pontos azuis) gerados com o método GMM-LS na região prevista do espaço objetivo, delimitado pelos indivíduos da frente prevista (estrelas vermelhas). Os novos indivíduos estão próximos das soluções previstas, mas também estão esparsamente distribuídos, como esperado.

Para compreender melhor o método OSP desenvolvido, nesta seção apresenta-se a análise da estratégia quando ela é empregada como CRM, e a influência do OSP quando é usada como parte do otimizador multiobjetivo (OSP-NSDE).

5.3.1 Eficácia do Procedimento OSP como Mecanismo de Reação à Mudança

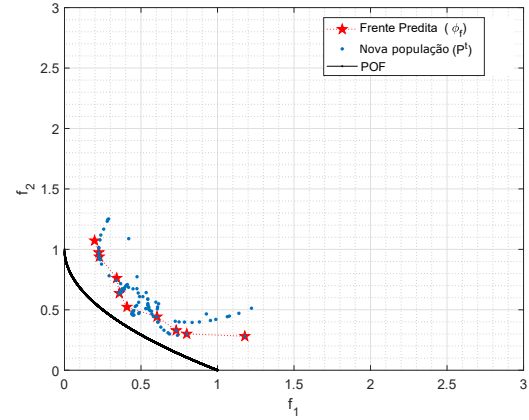
Para avaliar a influência da estratégia OSP como Mecanismo de Reação à Mudança na resolução do DMOEA, foram selecionados o NSGA-II (DEB et al., 2002) e o MOEA/D (ZHANG; LI, 2007) como AEs para otimizar. Diferentes estratégias de reação à mudança

Figura 18 – Etapa de aproximação das soluções líderes do OSP.



Fonte: Autor

Figura 19 – Etapa de geração da nova população usando o GMM-LS.



Fonte: Autor

baseados em predição foram utilizados para comparar com o OSP. As comparações incluem *Randomly Initialize Strategy* (RIS); *Feed-forward prediction strategy* (FPS) (HATZAKIS; WALLACE, 2006); *Population Prediction Strategy* (PPS) (ZHOU; JIN; ZHANG, 2014); e *Hybrid population prediction strategy based on fuzzy inference and one-step prediction* (FIOPPS) (CHEN et al., 2018), todos eles aplicados ao MOTLBO/D como o otimizador. As quatro estratégias são comparadas nos problemas FDA1-3, dMOP2 e nas funções UDF1, UDF2, UDF4 e UDF5. A configuração dos parâmetros foi a mesma adotada por (CHEN et al., 2018), para todas as estratégias de predição. Cada simulação foi executada com $N_c = 121$ mudanças ambientais. Para cada simulação, os algoritmos executam 100 gerações antes da primeira mudança.

A configuração DOSP-NSDE inclui a probabilidade de mutação de $p_m = 1/D$ e o índice de distribuição $\eta_m = 20$ para o operador de mutação. A estratégia canônica do DE (rand/1), com tamanho de passo de mutação $F = 0,5$ e taxa de cruzamento $CR = 1$ também foram usados. Para DOSP-NSDE, foi definido $\lambda = 0,6$, $\beta = 0,9$ e o horizonte de previsão inicial $p = 40$. Quando o ambiente muda, se as informações do histórico não forem suficientes para previsão, a estratégia RIS é executada.

A Tabela 5 mostra a média e o desvio padrão da métrica IGD para as estratégias comparadas com diferentes MOEAs para os problemas FDA1-3, dMOP2, UDF1,2,4,5. Pode-se observar que, para o NSGA-II, a estratégia do OSP é significativamente melhor que outras estratégias de predição para FDA1,3, dMOP2, UDF1,4. O OSP apresentou desempenho inferior ao FIOPPS para os problemas UDF2 e UDF 5 e apresenta o pior desempenho para o FDA2. No caso do MOEA/D, o OSP obtém o melhor desempenho apenas para UDF2 e UDF5. Estes resultados sugerem que a estratégia do OSP é mais apropriada para uso com MOEAs baseados em Pareto, como NSGA-II ou NSDE. A principal razão pode ser que a estratégia do OSP usa o princípio de dominância de Pareto para classificar e selecionar as soluções usadas no processo de previsão.

Tabela 5 – Média e desvio padrão da métrica IGD para as estratégias comparadas com diferentes MOEAs.

Estratégias	NSGA-II	MOEA/D	NSGA-II	MOEA/D
	FDA1		FDA2	
RIS	0,7112±0,0394 -	0,1003±0,0059 -	0,0081±0,0006 +	0,0940±0,0048 -
FPS	0,2743±0,0402 -	0,0428±0,0044 -	0,0084±0,0010 +	0,0489±0,0051 +
PPS	0,3502±0,0889 -	0,0255±0,0070 ≈	0,0076±0,0015 +	0,0250±0,0013 -
FIOPPS	0,0973±0,0061 -	0,0129±0,0004 -	0,0115±0,0010 +	0,0252±0,0006 +
OSP	0,0099± 0,0013	0,0241± 0,0022	0,0304± 0,0012	0,0715± 0,0010
Estratégias	FDA3		dMOP2	
RIS	0,4572±0,0413 -	0,0959±0,0052 +	0,8503±0,0719 -	0,1453±0,0109 -
FPS	0,3112±0,0246 -	0,0849±0,0043 +	0,2817±0,0380 -	0,0525±0,0061 -
PPS	0,5690±0,1763 -	0,0912±0,0227 +	0,4993±0,1635 -	0,0398±0,0105 -
FIOPPS	0,1297±0,0071 -	0,0668±0,0019 +	0,1276±0,0039 -	0,0148±0,0006 +
OSP	0,0303± 0,0011	0,1344± 0,0173	0,0108± 0,0012	0,0264± 0,0018
Estratégias	UDF1		UDF2	
RIS	0,1483±0,0248 -	0,1192±0,0161 +	0,0478±0,0065 +	0,0124±0,0007 +
FPS	0,1285±0,0218 ≈	0,1517±0,0639 +	0,0403±0,0056 +	0,0077±0,0003 +
PPS	0,2675±0,0476 -	0,0945±0,0223 +	0,0522±0,0289 +	0,0064±0,0003 +
FIOPPS	0,1653±0,0040 -	0,2364±0,0578 +	0,0229±0,0005 +	0,0062±0,0001 +
OSP	0,1163± 0,0274	0,1986± 0,0095	1,3177± 0,0143	1,2616± 0,0133
Estratégias	UDF4		UDF5	
RIS	0,1075±0,0117 -	0,0803±0,0064 +	0,0570±0,0075 -	0,0129±0,0007 -
FPS	0,1067±0,0066 -	0,1192±0,0421 +	0,0389±0,0041 -	0,0087±0,0005 -
PPS	0,1644±0,0394 -	0,0865±0,0049 +	0,0661±0,0283 -	0,0082±0,0016 -
FIOPPS	0,1104±0,0032 -	0,1391±0,0375 ≈	0,0239±0,0004 +	0,0071±0,0001 -
OSP	0,0600± 0,0025	0,1455± 0,0092	0,0287± 0,0022	0,0068± 0,0055

5.3.2 Eficácia do Procedimento OSP entre Mudanças (OSP-NSDE)

A eficácia do OSP entre mudanças foi validado analisando o desempenho do OSP-NSDE. Este algoritmo multi-objetivo estático foi avaliado através de simulações usando alguns *benchmarks* bem conhecidos, e foi comparado com onze algoritmos sobre MOPs selecionados. Nas tabelas que resumem os resultados estatísticos, o primeiro número representa o valor médio e o segundo o desvio padrão.

OSP-NSDE usa uma probabilidade de mutação de $p_m = 1/D$ e um índice de distribuição $\eta_m = 20$ para o operador de mutação. A estratégia canônica rand/1, com tamanho de passo de mutação $F = 0,5$ e taxa de cruzamento $CR = 1$ também foi usada. Para o OSP-NSDE, a porcentagem de variação λ foi definida para 0,5 e o horizonte de previsão inicial foi definido como $p = 100$. Os parâmetros dos algoritmos concorrentes foram definidos seguindo as suas publicações originais com o objetivo de replicar seus melhores resultados.

Primeiro, foi avaliada a influência da estratégia OSP e dos operadores de variação no processo evolucionário através da comparação do desempenho de três algoritmos:

1. NSDE: Versão original do algoritmo usando DE e distância de aglomeração;
2. NSDE modificado (Mod-NSDE): O algoritmo NSDE usando DE, mutação polinomial e mecanismo de diversidade do SPEA2;

3. OSP-NSDE: O algoritmo Mod-NSDE usando o método de previsão no espaço objetivo (OSP) e a inferência variacional para gerar indivíduos promissores.

A Tabela 6 mostra os valores médios e desvio padrão do IGD obtido pelos três algoritmos comparados para as 31 funções de teste. Observa-se que o desempenho do NSDE está abaixo dos demais. Comparando NSDE e Mod-NSDE, os resultados sugerem que há mais precisão da frente ótima de Pareto na maioria dos testes para Mod-NSDE. Comparando o Mod-NSDE e o OSP-NSDE, os resultados sugerem que o uso do procedimento OSP melhora ainda mais a qualidade das soluções na maioria das funções.

Tabela 6 – Média e desvio padrão do IGD para os três algoritmos baseados em NSDE. Os melhores resultados médios em cada caso estão em negrito. Os símbolos “ + ”, “ $\approx$ ” e “ - ” indicam que o algoritmo na coluna correspondente foi significativamente melhor que, equivalente ao, ou pior do que o OSP-NSDE.

F	OSP-NSDE	NSDE	Mod-NSDE
UF Benchmark			
1	7,8451e-3 (3,22e-3)	5,8585e-2 (1,55e-2) -	1,7448e-2 (5,08e-3) -
2	1,3552e-2 (2,08e-3)	5,6050e-2 (7,16e-3) -	1,7119e-2 (2,41e-3) -
3	8,1113e-2 (9,20e-2)	2,3180e-1 (3,21e-2) -	2,9945e-2 (1,90e-2) +
4	6,2295e-2 (4,04e-3)	6,9323e-2 (7,81e-3) -	6,3520e-2 (5,58e-3) $\approx$
5	6,4435e-1 (1,63e-1)	1,0978e+0 (2,52e-1) -	4,3508e-1 (1,12e-1) +
6	9,1748e-2 (6,20e-2)	5,5476e-1 (1,09e-1) -	1,1950e-1 (5,34e-2) -
7	9,5869e-3 (2,77e-3)	3,4042e-2 (1,84e-2) -	1,1025e-2 (7,35e-4) -
8	1,3578e-1 (6,72e-2)	4,3576e-1 (1,16e-1) -	2,3312e-1 (7,39e-2) -
9	2,6274e-1 (1,19e-1)	6,9325e-1 (1,91e-1) -	4,2556e-1 (1,12e-1) -
10	1,1922e+0 (4,40e-1)	2,0298e+0 (3,32e-1) -	1,2058e+0 (1,69e-1) $\approx$
ZDT Benchmark			
1	4,0105e-3 (6,26e-5)	5,0173e+0 (4,21e+0) -	4,0011e-3 (6,10e-5) $\approx$
2	3,9806e-3 (5,90e-5)	4,3895e+1 (7,21e+0) -	3,9810e-3 (6,01e-5) $\approx$
3	4,8085e-3 (7,08e-5)	1,6408e+0 (2,23e+0) -	4,8199e-3 (9,65e-5) $\approx$
4	4,3908e-3 (1,30e-4)	2,1066e+1 (6,98e+0) -	4,3560e-3 (1,25e-4) $\approx$
6	3,1348e-3 (4,20e-5)	3,8513e-3 (1,85e-4) -	3,1276e-3 (4,43e-5) $\approx$
DTLZ Benchmark			
1	2,4862e-2 (8,89e-4)	7,4162e-1 (6,01e-1) -	2,5464e-2 (6,48e-4) -
2	6,6494e-2 (1,73e-3)	7,7810e-2 (2,36e-3) -	6,9445e-2 (2,25e-3) -
3	6,4039e-2 (1,55e-3)	1,4086e+1 (1,65e+1) -	6,8954e-2 (2,11e-3) -
4	7,5481e-2 (2,59e-3)	8,0588e-2 (3,43e-3) -	7,8406e-2 (3,23e-3) -
5	1,1093e-2 (9,88e-4)	6,6234e-3 (3,75e-4) +	1,5779e-2 (1,06e-3) -
6	4,1210e-3 (3,96e-5)	5,0451e-1 (5,71e-1) -	4,7252e-3 (3,21e-5) -
7	1,0589e-1 (9,82e-2)	1,0970e-1 (8,16e-2) -	1,3660e-1 (1,08e-1) -
WFG Benchmark			
1	1,5228e+0 (2,88e-2)	1,5129e+0 (1,77e-2) +	1,5434e+0 (3,59e-2) -
2	1,8368e-1 (8,60e-3)	2,1299e-1 (1,49e-2) -	2,0153e-1 (8,05e-3) -
3	1,9065e-1 (1,35e-2)	1,1958e-1 (1,65e-2) +	2,3223e-1 (1,58e-2) -
4	2,8318e-1 (4,76e-3)	3,4689e-1 (1,11e-2) -	3,2352e-1 (6,89e-3) -
5	2,5129e-1 (5,12e-3)	2,8287e-1 (1,23e-2) -	2,7361e-1 (6,02e-3) -
6	2,7015e-1 (1,72e-2)	3,3409e-1 (2,10e-2) -	2,7098e-1 (8,55e-3) $\approx$
7	2,6339e-1 (7,03e-3)	3,3084e-1 (1,61e-2) -	2,9532e-1 (4,76e-3) -
8	3,6006e-1 (7,99e-3)	4,3804e-1 (1,65e-2) -	3,9284e-1 (7,21e-3) -
9	2,8612e-1 (6,33e-2)	3,9148e-1 (3,95e-2) -	3,2694e-1 (6,08e-2) -
+ / - / $\approx$		3/28/0	2/21/8

Ainda, pode-se ver que para problemas com POF convexa (UF3 e WFG1), ou proble-

mas com POF desconectado (UF5) ou POF linear (WFG3), os resultados obtidos pelo OSP-NSDE foram um pouco piores quando comparados aos seus competidores.

Comparações de desempenho com MOEAs

Onze EAs do estado da arte, foram escolhidos para comparações com OSP-NSDE. Eles foram subdivididos nos seguintes subgrupos: EAs baseados em Pareto (PESA-II (CORNE et al., 2001), NSGA-II (DEB et al., 2002)); EAs baseados em decomposição, (MOEA/D (ZHANG; LI, 2007)); MOEAs da família do DE (NSDE (ANGIRA; BABU, 2005), GDE3 (KUKKONEN; LAMPINEN, 2005), MOEA/D-DE (LI; ZHANG, 2009), MOEA/PC (DENY-SIUK et al., 2015), CPS-MOEA (ZHANG; ZHOU; ZHANG, 2015)); Algoritmos da família EDA (RM-MEDA (ZHANG; ZHOU; JIN, 2008), IM-MOEA (CHENG et al., 2015), e MO-CMA-ES (IGEL; HANSEN; ROTH, 2007)). Além disso, foi considerada a versão modificada do NSDE, sem a adição da estratégia do OSP para avaliar o efeito das ferramentas estatísticas inseridas no algoritmo proposto.

Cada média e desvio padrão foram calculados para 30 execuções de cada função de teste para cada algoritmo. As Tabelas 7 e 8 mostram as performances considerando as métricas IGD e HV. Na Tabela 7, observa-se que o OSP-NSDE supera os outros concorrentes na maioria das comparações individuais em relação à métrica IGD. O número de sinais negativos é significativamente maior que o número de $\approx$ adicionado ao número de sinais + para os problemas de teste. Além disso, nota-se que o OSP-NSDE alcançou a melhor média do IGD em 9 de 31 funções de teste. Em 7 de 9 funções, tem-se problemas de três objetivos, sugerindo adequação a funções complexas. O OSP-NSDE só perdeu as competições para as funções UF (para MOEA/D-DE) e ZDTs (para MOEA/D e MOEA/D-DE). Estes resultados também reforçam que o OSP-NSDE é sensível às formas da frente de Pareto.

Na Tabela 8, observa-se que o OSP-NSDE também supera os concorrentes na maioria das comparações individuais em relação ao indicador de alta tensão. OSP-NSDE alcançou a melhor média do IGD em 7 das 31 funções de teste; 3 destes 7 são problemas de três objetivos. O OSP-NSDE só perdeu as competições para as funções ZDTs (para GDE3) e WFGs (IM-MOEA e NSGA-II). A métrica HV compara os algoritmos simultaneamente para convergência e diversidade. Por esta razão, pode se dizer que as soluções obtidas têm uma boa distribuição e qualidade para a maioria das funções.

Considerando os valores médios dos indicadores utilizados, obteve-se a classificação dos treze algoritmos concorrentes em cada instância de teste. A Tabela 9 mostra as frequências de classificação de cada algoritmo em relação a IGD, HV e métrica *Ensemble*. Os algoritmos são ordenados de acordo com o número de vezes que atingem a primeira posição. Se houver empate na primeira posição, considera-se que a segunda quebra o empate e assim por diante. Observa-se que, de acordo com o IGD e o HV, o OSP-NSDE atinge o maior número de frequências no primeiro lugar superando todos os concorrentes.

Tabela 7 – Média e desvio padrão do IGD para os algoritmos comparados. Os melhores resultados médios em cada instância estão em negrito. Os símbolos ” + ”, ” $\approx$ ” e ” - ” indicam onde o algoritmo na coluna correspondente funciona significativamente melhor que, equivalente a, ou pior do que o OSP-NSDE.

Problema	OSP-NSDE	NSDE	GDE3	MOEA/PC	CPS-MOEA	MOEA/D-DE
UF1	7,8451e-3 (3,22e-3)	5,8585e-2 (1,55e-2) -	2,6484e-2 (4,26e-3) -	7,1034e-3 (5,88e-4) $\approx$	1,2976e-2 (3,18e-3) -	5,2936e-3 (1,48e-3) +
UF2	1,3552e-2 (2,08e-3)	5,6050e-2 (7,16e-3) -	1,6841e-2 (2,46e-3) -	2,3937e-2 (2,84e-2) -	1,7422e-2 (1,71e-3) -	1,2287e-2 (2,90e-3) +
UF3	8,1113e-2 (9,20e-2)	2,3180e-1 (3,21e-2) -	3,9413e-2 (1,91e-2) $\approx$	7,9864e-2 (6,03e-2) $\approx$	2,3894e-2 (1,66e-2) +	1,6735e-2 (1,35e-2) +
UF4	6,2295e-2 (4,04e-3)	6,9323e-2 (7,81e-3) -	6,1654e-2 (6,53e-3) $\approx$	6,3656e-2 (5,26e-3) $\approx$	6,3304e-2 (6,00e-3) $\approx$	7,0434e-2 (7,10e-3) -
UF5	6,4435e-1 (1,63e-1)	1,0978e+0 (2,52e-1) -	4,7703e-1 (1,13e-1) +	4,8856e-1 (1,95e-1) +	4,7730e-1 (1,29e-1) +	3,5930e-1 (1,56e-1) +
UF6	9,1748e-2 (6,20e-2)	5,5476e-1 (1,09e-1) -	1,2449e-1 (5,11e-2) -	2,5855e-1 (1,98e-1) -	1,5172e-1 (1,34e-1) -	3,0589e-1 (1,45e-1) -
UF7	9,5869e-3 (2,77e-3)	3,4042e-2 (1,84e-2) -	5,0941e-2 (1,09e-3) -	1,6216e-2 (1,43e-2) $\approx$	1,2679e-2 (9,44e-4) -	8,9705e-3 (7,24e-3) +
UF8	1,3578e-1 (6,72e-2)	4,3576e-1 (1,16e-1) -	7,6607e-1 (2,00e-1) -	2,1937e-1 (6,36e-2) -	5,7171e-1 (1,27e-1) -	1,8235e-1 (6,71e-2) -
UF9	2,6274e-1 (1,19e-1)	6,9325e-1 (1,91e-1) -	7,0823e-1 (1,74e-1) -	1,9722e-1 (5,54e-2) $\approx$	5,5284e-1 (1,33e-1) -	2,3170e-1 (2,46e-2) $\approx$
UF10	1,1922e+0 (4,40e-1)	2,0298e+0 (3,32e-1) -	1,6596e+0 (4,16e-1) -	7,0616e-1 (1,08e-1) +	1,5506e+0 (3,71e-1) -	5,7219e-1 (5,20e-2) +
ZDT1	4,0105e-3 (6,26e-5)	5,0173e+0 (4,21e+0) -	5,9494e-3 (6,46e-5) -	4,8805e-3 (1,12e-5) -	5,0069e-3 (2,65e-4) -	3,8897e-3 (2,02e-6) +
ZDT2	3,9806e-3 (5,90e-5)	4,3895e+1 (7,21e+0) -	8,0345e-3 (5,35e-5) -	3,8160e-3 (1,64e-6) +	4,9541e-3 (2,16e-4) -	3,8087e-3 (8,33e-7) +
ZDT3	4,8085e-3 (7,08e-5)	1,6408e+0 (2,23e+0) -	6,5225e-3 (3,70e-5) -	1,5625e-2 (5,40e-4) -	5,3089e-3 (2,27e-4) -	1,0934e-2 (1,48e-5) -
ZDT4	4,3908e-3 (1,30e-4)	2,1066e+1 (6,98e+0) -	4,2844e-3 (9,07e-5) +	1,7121e+1 (1,19e+1) -	5,3997e-3 (2,75e-4) -	3,9031e-3 (1,72e-5) +
ZDT6	3,1348e-3 (4,20e-5)	3,8513e-3 (1,85e-4) -	9,1098e-3 (3,48e-5) -	3,2403e-3 (6,52e-8) -	3,8241e-3 (1,82e-4) -	3,1030e-3 (3,68e-8) +
DTLZ1	2,4862e-2 (8,89e-4)	7,4162e-1 (6,01e-1) -	3,0818e-2 (1,45e-3) -	2,8337e-2 (1,52e-4) -	3,2447e-2 (1,24e-3) -	3,0924e-2 (5,13e-5) -
DTLZ2	6,6494e-2 (1,73e-3)	7,7810e-2 (2,36e-3) -	7,8753e-2 (2,85e-3) -	6,1331e-2 (1,19e-3) +	8,0812e-2 (2,77e-3) -	7,5376e-2 (2,15e-4) -
DTLZ3	6,4039e-2 (1,55e-3)	1,4086e+1 (1,65e+1) -	7,7509e-2 (2,77e-3) -	6,1308e-2 (9,73e-4) +	7,9427e-2 (3,48e-3) -	7,5281e-2 (3,07e-4) -
DTLZ4	7,5481e-2 (2,59e-3)	8,0588e-2 (3,43e-3) -	8,0384e-2 (2,87e-3) -	6,2171e-2 (1,23e-3) +	8,1669e-2 (2,86e-3) -	9,3237e-2 (8,48e-2) -
DTLZ5	1,1093e-2 (9,88e-4)	6,6234e-3 (3,75e-4) +	5,7734e-3 (2,39e-4) +	2,9816e-2 (2,73e-3) -	7,7198e-3 (5,97e-4) +	1,4603e-2 (1,37e-5) -
DTLZ6	4,1210e-3 (3,96e-5)	5,0451e-1 (5,71e-1) -	4,5018e-3 (8,50e-5) -	2,8332e-2 (2,85e-3) -	5,3672e-3 (3,57e-4) -	1,4621e-2 (6,65e-6) -
DTLZ7	1,0589e-1 (9,82e-2)	1,0970e-1 (8,16e-2) -	1,1642e-1 (8,12e-2) -	1,4737e-1 (5,85e-3) -	1,2568e-1 (9,20e-2) -	2,1045e-1 (3,46e-2) -
WFG1	1,5228e+0 (2,88e-2)	1,5129e+0 (1,77e-2) +	1,6488e+0 (7,20e-2) -	1,8901e+0 (1,58e-1) -	1,6040e+0 (5,25e-2) -	6,3698e-1 (1,12e-1) +
WFG2	1,8368e-1 (8,60e-3)	2,1299e-1 (1,49e-2) -	2,1889e-1 (1,21e-2) -	6,8098e-1 (5,24e-3) -	2,2092e-1 (1,22e-2) -	5,6595e-1 (2,85e-2) -
WFG3	1,9065e-1 (1,35e-2)	1,1958e-1 (1,65e-2) +	1,2424e-1 (2,23e-2) +	2,1714e-1 (1,60e-2) -	1,3089e-1 (1,86e-2) +	1,0796e-1 (7,46e-4) +
WFG4	2,8318e-1 (4,76e-3)	3,4689e-1 (1,11e-2) -	3,5084e-1 (1,20e-2) -	2,9856e-1 (5,17e-3) -	3,4183e-1 (9,73e-3) -	3,6823e-1 (6,97e-3) -
WFG5	2,5129e-1 (5,12e-3)	2,8287e-1 (1,23e-2) -	2,8829e-1 (1,21e-2) -	2,8817e-1 (1,61e-3) -	2,9219e-1 (9,11e-3) -	3,6519e-1 (5,91e-3) -
WFG6	2,7015e-1 (1,72e-2)	3,3409e-1 (2,10e-2) -	3,4551e-1 (1,12e-2) -	3,0842e-1 (6,00e-2) -	3,4692e-1 (2,60e-2) -	3,8152e-1 (4,31e-2) -
WFG7	2,6339e-1 (7,03e-3)	3,3084e-1 (1,61e-2) -	3,3844e-1 (1,20e-2) -	2,7168e-1 (2,54e-3) -	3,3663e-1 (1,26e-2) -	3,5876e-1 (2,69e-3) -
WFG8	3,6006e-1 (7,99e-3)	4,3804e-1 (1,65e-2) -	4,2755e-1 (1,21e-2) -	3,1955e-1 (2,76e-3) +	4,2600e-1 (1,48e-2) -	3,9355e-1 (2,72e-3) -
WFG9	2,8612e-1 (6,33e-2)	3,9148e-1 (3,95e-2) -	3,6815e-1 (5,31e-2) -	3,3082e-1 (6,70e-2) -	3,3852e-1 (6,08e-2) -	3,5299e-1 (2,09e-2) $\approx$
+/- / $\approx$		3/28/0	4/25/2	7/19/5	4/26/1	12/17/2

Continuação

Problema	RM-MEDA	IM-MOEA	MO-CMA-ES	PESA-II	NSGA-II	MOEA/D
UF1	4,4583e-2 (1,21e-2) -	5,0602e-2 (1,33e-2) -	1,7707e-1 (1,17e-2) -	1,3178e-1 (4,30e-2) -	9,2260e-2 (2,30e-2) -	2,4618e-1 (1,01e-1) -
UF2	1,8795e-2 (3,90e-3) -	2,8355e-2 (1,39e-2) -	3,5309e-1 (1,57e-2) -	6,0008e-2 (2,74e-2) -	3,2808e-2 (1,08e-2) -	1,2507e-1 (4,93e-2) -
UF3	7,6478e-2 (5,15e-2) $\approx$	8,3367e-2 (1,57e-2) -	3,7877e-1 (8,73e-3) +	3,0256e-1 (3,00e-2) -	2,3570e-1 (6,00e-2) -	3,1602e-1 (3,40e-2) -
UF4	8,9710e-2 (5,67e-3) -	5,4832e-2 (3,31e-3) +	5,7114e-2 (3,55e-3) +	5,2333e-2 (2,53e-3) +	4,5710e-2 (5,29e-3) +	7,9803e-2 (4,28e-3) -
UF5	7,1411e-1 (1,28e-1) -	2,8977e-1 (4,69e-2) +	1,7043e+0 (8,08e-2) -	3,5691e-1 (1,15e-1) +	2,9969e-1 (1,08e-1) -	4,9197e-1 (1,18e-1) +
UF6	1,6222e-1 (5,36e-2) -	1,5940e-1 (9,73e-2) -	8,1713e-1 (3,41e-2) -	2,4663e-1 (1,19e-1) -	1,9039e-1 (1,05e-1) -	4,2468e-1 (1,22e-1) -
UF7	2,3779e-2 (3,08e-3) -	4,8200e-2 (7,51e-2) -	1,6797e-1 (2,00e-2) -	2,3374e-1 (1,57e-1) -	1,3854e-1 (1,56e-1) -	3,9549e-1 (1,91e-1) -
UF8	3,4066e-1 (9,74e-2) -	2,5639e-1 (4,61e-2) -	1,7647e+0 (4,83e-1) -	2,8505e-1 (5,24e-2) -	2,6464e-1 (9,03e-2) -	4,0388e-1 (2,71e-1) -
UF9	4,3148e-1 (1,73e-1) -	2,1697e-1 (6,51e-2) $\approx$	1,1847e+0 (2,23e-1) -	2,3165e-1 (6,59e-2) $\approx$	3,3639e-1 (9,42e-2) -	2,4997e-1 (3,11e-2) $\approx$
UF10	2,0025e+0 (2,65e-1) -	5,9721e-1 (4,19e-2) -	4,9226e+0 (1,26e+0) -	4,7460e-1 (7,65e-2) +	6,0921e-1 (2,14e-1) +	7,5101e-1 (1,48e-1) +
ZDT1	2,7991e-1 (3,22e-1) -	4,7190e-2 (2,27e-3) -	5,8206e-3 (4,45e-4) -	1,5368e-2 (4,37e-3) -	5,2129e-3 (2,63e-4) -	3,9559e-3 (2,17e-6) +
ZDT2	4,3240e-1 (3,43e-1) -	7,8476e-2 (4,24e-3) -	5,8994e-3 (4,18e-4) -	1,4877e-2 (1,87e-3) -	5,3212e-3 (1,99e-4) -	3,8084e-3 (5,93e-6) +
ZDT3	5,0225e-1 (1,67e-1) -	3,7375e-2 (1,86e-3) -	7,0478e-3 (5,99e-4) -	1,6215e-2 (2,39e-3) -	5,9215e-3 (3,01e-4) -	1,1551e-2 (1,56e-4) -
ZDT4	1,1413e+1 (2,89e+0) -	4,6522e-3 (1,05e-4) -	1,3925e+1 (3,40e+0) -	1,1641e-2 (2,71e-3) -	4,5387e-3 (1,30e-4) -	3,9886e-3 (4,21e-5) +
ZDT6	5,4075e-2 (5,62e-2) -	1,5157e+0 (8,77e-2) -	3,8195e-3 (2,53e-4) -	8,6068e-3 (1,22e-3) -	4,4325e-3 (1,96e-4) -	3,1425e-3 (2,56e-5) $\approx$
DTLZ1	1,0321e+1 (2,47e+0) -	5,7346e-1 (1,85e-1) -	2,8109e+1 (4,21e+0) -	2,4911e-2 (1,65e-3) $\approx$	2,7269e-2 (1,54e-3) -	3,0557e-2 (1,04e-6) -
DTLZ2	6,9798e-2 (2,38e-3) -	8,1267e-2 (2,35e-3) -	3,0110e-1 (1,89e-2) -	6,6165e-2 (3,32e-3) $\approx$	6,9241e-2 (2,49e-3) -	7,4464e-2 (5,75e-8) -
DTLZ3	1,5755e+2 (1,06e+1) -	1,5694e+1 (3,08e+0) -	2,1361e+2 (2,37e+1) -	7,2289e-2 (1,23e-2) -	6,9966e-2 (3,65e-3) -	7,4473e-2 (6,26e-6) -
DTLZ4	7,3553e-2 (2,57e-3) +	7,4656e-2 (1,87e-3) $\approx$	4,8805e-2 (2,69e-2) -	6,4456e-2 (2,45e-3) +	9,5841e-2 (1,61e-1) -	8,4464e-2 (2,23e-7) -
DTLZ5	5,2066e-3 (2,06e-4) +	2,2108e-2 (2,35e-3) -	2,2191e-1 (2,03e-2) -	1,0745e-2 (1,12e-3) $\approx$	5,7725e-3 (2,74e-4) +	3,3925e-2 (2,53e-6) -
DTLZ6	9,2650e-1 (7,76e-1) -	3,4439e+0 (1,42e-1) -	4,0925e-2 (8,07e-3) -	1,5491e-2 (4,04e-3) -	5,9892e-3 (3,98e-4) -	3,3929e-2 (5,24e-7) -
DTLZ7	9,9568e-2 (1,04e-2) +	3,2561e-1 (4,96e-2) -	1,0501e-1 (7,62e-3) +	1,3411e-1 (1,62e-1) -	8,5544e-2 (5,02e-2) +	1,5335e-1 (1,48e-3) -
WFG1	2,0046e+0 (8,13e-2) -	1,2384e+0 (8,96e-2) +	1,9917e+0 (3,14e-2) -	1,8313e-1 (1,87e-2) +	2,4853e-1 (2,37e-2) +	2,8687e-1 (2,98e-3) +
WFG2	2,5125e-1 (1,83e-2) -	2,8662e-1 (2,75e-2) -	2,5005e-1 (1,11e-2) -	2,0603e-1 (2,80e-2) -	1,9823e-1 (1,19e-2) -	7,9665e-1 (1,91e-1) -
WFG3	2,1719e-1 (1,91e-2) -	1,8048e-1 (1,36e-2) +	1,7482e-1 (1,48e-2) +	4,0014e-1 (2,24e-1) -	9,1748e-2 (1,81e-2) +	1,5709e-1 (4,72e-4) +
WFG4	3,3938e-1 (1,22e-2) -	3,2906e-1 (8,71e-3) -	3,2984e-1 (1,09e-2) -	2,8366e-1 (1,56e-2) $\approx$	2,6963e-1 (1,02e-2) +	2,4710e-1 (2,23e-3) +
WFG5	3,2008e-1 (1,81e-2) -	3,3610e-1 (1,17e-2) -	2,6731e-1 (8,42e-3) -	2,8317e-1 (1,76e-2) -	2,8041e-1 (9,22e-3) -	2,8645e-1 (1,57e-3) -
WFG6	3,5362e-1 (1,81e-2) -	3,3859e-1 (1,16e-2) -	3,5957e-1 (1,19e-2) -	3,0693e-1 (2,32e-2) -	3,0252e-1 (1,19e-2) -	2,9214e-1 (1,11e-2) -
WFG7	3,3677e-1 (1,52e-2) -	3,4627e-1 (1,27e-2) -	3,3611e-1 (1,43e-2) -	2,9224e-1 (1,58e-2) -	2,8066e-1 (9,07e-3) -	2,6714e-1 (8,38e-3) $\approx$
WFG8	4,3618e-1 (1,53e-2) -	4,1252e-1 (1,10e-2) -	4,2209e-1 (1,20e-2) -	3,8153e-1 (1,81e-2) -	3,7421e-1 (1,14e-2) -	2,9718e-1 (2,82e-3) +
WFG9	2,7491e-1 (1,03e-2) $\approx$	3,3716e-1 (1,24e-2) $\approx$	2,9386e-1 (3,39e-2) -	2,6828e-1 (8,26e-3) $\approx$	2,8182e-1 (2,34e-2) $\approx$	2,4335e-1 (6,08e-3) $\approx$
+/- / $\approx$	3/26/2	4/24/3	3/28/0	5/20/6	8/22/1	9/18/4

Tabela 8 – Média e desvio padrão do HV para os algoritmos comparados. Os melhores resultados médios em cada instância estão em negrito. Os símbolos ” + ”, ” $\approx$ ” e ” - ” indicam onde o algoritmo na coluna correspondente funciona significativamente melhor que, equivalente a, ou pior do que o OSP-NSDE.

Problema	OSP-NSDE	NSDE	GDE3	MOEA/PC	CPS-MOEA	MOEA/D-DE
UF1	8,6300e-1 (5,13e-3)	7,7083e-1 (3,25e-2) -	8,4867e-1 (7,28e-3) -	8,6102e-1 (9,64e-4) -	8,5445e-1 (5,23e-3) -	8,4679e-1 (2,69e-3) -
UF2	8,5536e-1 (2,69e-3)	7,9053e-1 (1,06e-2) -	8,5096e-1 (3,54e-3) -	8,4314e-1 (1,95e-2) -	8,4896e-1 (2,32e-3) -	8,3704e-1 (4,05e-3) -
UF3	7,7202e-1 (9,93e-2)	5,3115e-1 (5,27e-2) -	8,1006e-1 (2,85e-2) $\approx$	7,9459e-1 (4,72e-2) $\approx$	8,3368e-1 (2,65e-2) +	8,4644e-1 (2,18e-2) +
UF4	4,3455e-1 (6,30e-3)	4,2190e-1 (1,26e-2) -	4,3629e-1 (1,03e-2) $\approx$	4,3061e-1 (7,84e-3) -	4,3180e-1 (1,04e-2) $\approx$	4,2063e-1 (1,13e-2) -
UF5	2,9334e-2 (4,94e-2)	0,0000e+0 (0,00e+0) -	7,9928e-2 (7,22e-2) +	1,7161e-1 (9,13e-2) +	6,8275e-2 (7,11e-2) +	2,5613e-1 (1,07e-1) +
UF6	4,8860e-1 (7,50e-2)	3,6873e-2 (4,76e-2) -	4,1400e-1 (4,10e-2) -	3,4154e-1 (1,15e-1) -	4,0604e-1 (9,14e-2) -	3,1477e-1 (9,69e-2) -
UF7	6,9205e-1 (5,71e-3)	6,4827e-1 (3,12e-2) -	6,8951e-1 (1,94e-3) -	6,8221e-1 (1,85e-2) -	6,8616e-1 (1,81e-3) -	6,3285e-1 (1,24e-2) -
UF8	5,3720e-1 (1,11e-1)	1,1212e-1 (8,31e-2) -	1,4660e-2 (2,89e-2) -	4,6833e-1 (6,37e-2) -	4,0236e-2 (4,25e-2) -	4,7098e-1 (6,55e-2) -
UF9	6,1307e-1 (1,96e-1)	1,5382e-1 (1,34e-1) -	1,3449e-1 (1,12e-1) -	8,1397e-1 (5,63e-2) +	2,6651e-1 (1,11e-1) -	7,6676e-1 (2,35e-2) +
UF10	2,2871e-3 (7,62e-3)	0,0000e+0 (0,00e+0) -	0,0000e+0 (0,00e+0) -	1,2248e-1 (4,79e-2) +	9,2907e-7 (5,09e-6) $\approx$	1,3875e-1 (3,05e-2) +
ZDT1	8,7087e-1 (2,28e-4)	4,0071e-2 (1,17e-1) -	8,7100e-1 (2,28e-4) +	8,6770e-1 (4,03e-5) -	8,6947e-1 (3,92e-4) -	8,5146e-1 (3,51e-5) -
ZDT2	5,3774e-1 (1,70e-4)	0,0000e+0 (0,00e+0) -	5,3788e-1 (1,36e-4) +	5,3569e-1 (1,69e-5) -	5,3637e-1 (3,36e-4) -	5,1840e-1 (3,12e-5) -
ZDT3	1,0234e+0 (1,74e-4)	2,9252e-1 (2,13e-1) -	1,0234e+0 (1,70e-4) $\approx$	1,0144e+0 (2,60e-4) -	1,0225e+0 (1,75e-2) -	1,0205e+0 (1,34e-5) -
ZDT4	8,6970e-1 (3,01e-4)	0,0000e+0 (0,00e+0) -	8,6987e-1 (2,34e-4) +	7,2634e-2 (2,31e-1) -	8,6791e-1 (4,50e-4) -	8,5124e-1 (1,72e-4) -
ZDT6	4,3337e-1 (1,38e-4)	4,3262e-1 (1,91e-4) -	4,3342e-1 (5,85e-5) $\approx$	4,3108e-1 (4,06e-7) -	4,3261e-1 (2,22e-4) -	4,0348e-1 (1,18e-6) -
DTLZ1	1,3741e-1 (4,05e-4)	1,8840e-2 (4,00e-2) -	1,3426e-1 (8,14e-4) -	1,2384e-1 (7,93e-5) -	1,3389e-1 (7,05e-4) -	1,3357e-1 (4,12e-5) -
DTLZ2	6,9564e-1 (4,79e-3)	6,7576e-1 (5,93e-3) -	6,7231e-1 (4,85e-3) -	7,0340e-1 (1,00e-2) +	6,7005e-1 (5,74e-3) -	6,9867e-1 (4,61e-4) +
DTLZ3	7,0215e-1 (4,49e-3)	6,3544e-2 (1,86e-1) -	6,7315e-1 (7,65e-3) -	7,0951e-1 (6,98e-3) +	6,7379e-1 (5,72e-3) -	6,9779e-1 (1,14e-3) -
DTLZ4	6,7824e-1 (6,38e-3)	6,6940e-1 (7,93e-3) -	6,7030e-1 (6,03e-3) -	7,2202e-1 (6,44e-3) +	6,7035e-1 (5,80e-3) -	6,9709e-1 (4,60e-2) +
DTLZ5	1,2807e-1 (8,68e-4)	1,3156e-1 (1,79e-4) +	1,3198e-1 (1,09e-4) +	1,2330e-1 (1,26e-3) -	1,3096e-1 (2,95e-4) +	1,1055e-1 (5,00e-6) -
DTLZ6	1,3314e-1 (2,56e-5)	7,0882e-2 (6,74e-2) -	1,3326e-1 (1,90e-5) +	1,2427e-1 (1,06e-3) -	1,3296e-1 (7,74e-5) -	1,2957e-1 (1,53e-6) -
DTLZ7	1,5496e+0 (6,69e-2)	1,5524e+0 (5,37e-2) $\approx$	1,4962e+0 (4,48e-2) -	1,4497e+0 (1,38e-2) -	1,4845e+0 (5,33e-2) -	1,3763e+0 (3,60e-3) -
WFG1	1,8693e+1 (6,08e-1)	1,9925e+1 (4,04e-1) +	1,7701e+1 (6,19e-1) -	8,4453e+0 (3,57e+0) -	1,8403e+1 (7,02e-1) $\approx$	4,4847e+1 (4,08e+0) +
WFG2	5,7778e+1 (1,96e-1)	5,8028e+1 (2,48e-1) +	5,7272e+1 (2,50e-1) -	5,6157e+1 (2,00e-1) -	5,7436e+1 (2,41e-1) -	5,5931e+1 (4,25e-1) -
WFG3	5,3591e+0 (1,25e-1)	6,1925e+0 (6,16e-2) +	6,2010e+0 (7,22e-2) +	5,0534e+0 (1,49e-1) -	6,0930e+0 (6,52e-2) +	6,0757e+0 (1,49e-2) +
WFG4	3,0018e+1 (2,70e-1)	2,9290e+1 (3,03e-1) -	2,9550e+1 (4,13e-1) -	2,9541e+1 (1,79e-1) -	2,9699e+1 (2,69e-1) -	3,1651e+1 (2,28e-1) +
WFG5	3,0299e+1 (2,99e-1)	3,0332e+1 (5,13e-1) $\approx$	2,9980e+1 (5,87e-1) -	2,9438e+1 (2,06e-1) -	2,9889e+1 (3,91e-1) -	2,9098e+1 (1,27e-1) -
WFG6	3,1153e+1 (1,04e+0)	3,0209e+1 (1,17e+0) -	2,9886e+1 (4,34e-1) -	2,9205e+1 (3,53e+0) -	2,9624e+1 (1,53e+0) -	2,9258e+1 (3,51e+0) -
WFG7	3,3557e+1 (3,24e-1)	3,0587e+1 (5,14e-1) -	3,0224e+1 (3,43e-1) -	3,1518e+1 (1,90e-1) -	3,0325e+1 (3,91e-1) -	3,1676e+1 (2,35e-1) -
WFG8	2,6487e+1 (3,14e-1)	2,5290e+1 (3,86e-1) -	2,5535e+1 (4,01e-1) -	2,7873e+1 (2,92e-1) +	2,5533e+1 (3,39e-1) -	2,3642e+1 (1,50e-1) -
WFG9	2,9136e+1 (3,35e+0)	2,4848e+1 (2,16e+0) -	2,6143e+1 (3,01e+0) -	2,7629e+1 (3,80e+0) -	2,7859e+1 (3,40e+0) -	3,0798e+1 (1,56e+0) $\approx$
+/-/ $\approx$		4/25/2	7/20/4	7/23/1	4/24/3	9/21/1

Continuação

Problema	RM-MEDA	IM-MOEA	MO-CMA-ES	PESA-II	NSGA-II	MOEA/D
UF1	8,0760e-1 (1,66e-2) -	7,8823e-1 (2,16e-2) -	5,5238e-1 (2,04e-2) -	6,9096e-1 (5,82e-2) -	7,3734e-1 (3,68e-2) -	5,9062e-1 (6,51e-2) -
UF2	8,4938e-1 (3,99e-3) -	8,3855e-1 (9,56e-3) -	3,2343e-1 (1,76e-2) -	8,0850e-1 (2,01e-2) -	8,3139e-1 (1,04e-2) -	7,7570e-1 (2,75e-2) -
UF3	7,4213e-1 (8,44e-2) $\approx$	7,1471e-1 (3,00e-2) -	2,2446e-1 (6,36e-3) -	4,8490e-1 (4,54e-2) -	5,3787e-1 (7,55e-2) -	4,4823e-1 (5,91e-2) -
UF4	3,9103e-1 (9,09e-3) -	4,5143e-1 (4,53e-3) +	4,4346e-1 (6,11e-3) +	4,4805e-1 (4,89e-3) +	4,6346e-1 (9,85e-3) +	4,3074e-1 (5,41e-3) -
UF5	2,9486e-3 (1,32e-2) -	2,7694e-1 (5,07e-2) +	0,0000e+0 (0,00e+0) -	2,8071e-1 (8,09e-2) +	2,7942e-1 (8,34e-2) +	1,9782e-1 (9,01e-2) +
UF6	3,6803e-1 (6,58e-2) -	3,9344e-1 (6,15e-2) -	0,0000e+0 (0,00e+0) -	3,4693e-1 (9,40e-2) -	3,7915e-1 (7,97e-2) -	2,4416e-1 (9,10e-2) -
UF7	6,6846e-1 (4,91e-3) -	6,3961e-1 (9,08e-2) -	4,1440e-1 (3,40e-2) -	4,7307e-1 (1,30e-1) -	5,5899e-1 (1,35e-1) -	3,3904e-1 (1,52e-1) -
UF8	2,4177e-1 (1,05e-1) -	4,3865e-1 (3,35e-2) -	0,0000e+0 (0,00e+0) -	3,5293e-1 (5,43e-2) -	3,7707e-1 (1,01e-1) -	3,7726e-1 (1,42e-1) -
UF9	4,1914e-1 (2,16e-1) -	7,0681e-1 (5,49e-2) $\approx$	5,8644e-5 (2,07e-2) $\approx$	7,0273e-1 (8,42e-2) $\approx$	5,3885e-1 (1,31e-1) -	7,2624e-1 (3,89e-2) $\approx$
UF10	8,2522e-5 (4,52e-4) $\approx$	3,7271e-1 (5,95e-2) +	0,0000e+0 (0,00e+0) -	1,4793e-1 (4,59e-2) +	5,6581e-2 (5,21e-2) +	1,4015e-1 (5,43e-2) +
ZDT1	5,7060e-1 (2,88e-1) -	7,9823e-1 (3,62e-3) -	8,6837e-1 (8,71e-4) -	8,5691e-1 (4,03e-3) -	8,7016e-1 (3,26e-4) -	8,7064e-1 (2,23e-5) -
ZDT2	2,0968e-1 (2,27e-1) -	4,0625e-1 (6,44e-3) -	5,3535e-1 (7,07e-4) -	5,2265e-1 (4,86e-3) -	5,3682e-1 (3,09e-4) -	5,1847e-1 (2,21e-5) -
ZDT3	3,7341e-1 (1,75e-1) -	9,4170e-1 (3,74e-3) -	1,0225e+0 (3,50e-4) -	1,0134e+0 (3,20e-3) -	1,0231e+0 (1,82e-4) -	1,0185e+0 (1,45e-4) -
ZDT4	0,0000e+0 (0,00e+0) -	8,6947e-1 (2,34e-4) -	0,0000e+0 (0,00e+0) -	8,6139e-1 (3,23e-3) -	8,3098e-1 (1,53e-4) -	8,5029e-1 (2,55e-4) -
ZDT6	3,7706e-1 (5,72e-2) -	0,0000e+0 (0,00e+0) -	4,3265e-1 (3,36e-4) -	4,2722e-1 (1,78e-3) -	4,3198e-1 (2,05e-4) -	4,3314e-1 (1,35e-4) -
DTLZ1	0,0000e+0 (0,00e+0) -	2,5220e-3 (4,11e-3) -	0,0000e+0 (0,00e+0) -	1,3743e-1 (7,89e-4) $\approx$	1,3724e-1 (5,15e-4) $\approx$	1,3004e-1 (2,50e-6) -
DTLZ2	6,9203e-1 (5,71e-3) -	6,8785e-1 (4,81e-3) -	2,0171e-1 (3,08e-2) -	6,9529e-1 (1,09e-2) $\approx$	6,7842e-1 (4,98e-3) -	6,4485e-1 (3,30e-7) -
DTLZ3	0,0000e+0 (0,00e+0) -	0,0000e+0 (0,00e+0) -	0,0000e+0 (0,00e+0) -	6,9695e-1 (2,27e-2) $\approx$	7,1169e-1 (7,76e-3) +	7,4438e-1 (3,10e-4) +
DTLZ4	6,8965e-1 (7,82e-3) +	7,0673e-1 (3,14e-3) +	1,8810e-2 (2,85e-2) -	7,1367e-1 (6,28e-3) +	6,5324e-1 (1,08e-1) -	6,4485e-1 (8,18e-6) -
DTLZ5	1,3241e-1 (1,04e-4) +	1,2427e-1 (1,15e-3) -	1,7413e-2 (6,15e-3) -	1,2828e-1 (2,19e-3) $\approx$	1,3252e-1 (9,51e-5) +	1,2102e-1 (8,46e-7) -
DTLZ6	3,5672e-2 (5,03e-2) -	0,0000e+0 (0,00e+0) -	1,2080e-1 (2,24e-3) -	1,2740e-1 (3,65e-3) -	1,3272e-1 (8,05e-5) -	1,2102e-1 (2,52e-7) -
DTLZ7	1,4419e+0 (3,83e-2) -	9,3441e-1 (1,16e-1) -	1,5420e+0 (1,93e-2) $\approx$	1,5397e+0 (9,01e-2) $\approx$	1,2780e+0 (3,27e-2) -	1,5205e+0 (2,66e-3) -
WFG1	5,5218e+0 (1,88e+0) -	2,4900e+1 (1,54e+0) +	6,0627e+0 (9,08e-1) -	5,9119e+1 (3,42e-1) +	5,8934e+1 (2,47e-1) +	5,9044e+1 (1,86e-1) +
WFG2	5,6191e+1 (3,56e-1) -	5,7422e+1 (4,07e-1) -	5,6240e+1 (2,42e-1) -	5,9000e+1 (1,98e-1) +	5,5723e+1 (1,60e-1) -	5,3818e+1 (3,03e-1) -
WFG3	5,5280e+0 (7,63e-2) +	5,7951e+0 (9,01e-2) +	5,7940e+0 (5,93e-2) +	4,3029e+0 (9,29e-1) -	6,4213e+0 (4,74e-2) +	5,0339e+0 (6,99e-3) -
WFG4	2,9670e+1 (4,67e-1) -	3,2623e+1 (2,57e-1) +	3,0385e+1 (3,11e-1) +	3,0671e+1 (6,08e-1) +	3,3470e+1 (3,04e-1) +	3,5058e+1 (4,61e-1) +
WFG5	2,8801e+1 (7,02e-1) -	3,0650e+1 (3,55e-1) +	3,1242e+1 (2,91e-1) +	2,9717e+1 (5,17e-1) -	3,1443e+1 (3,19e-1) +	2,8777e+1 (2,12e-1) -
WFG6	2,9258e+1 (5,18e-1) -	3,1955e+1 (5,90e-1) +	2,9500e+1 (3,44e-1) -	2,9608e+1 (9,86e-1) -	3,0306e+1 (8,79e-1) -	3,0139e+1 (9,71e-1) -
WFG7	2,9809e+1 (3,59e-1) -	3,2440e+1 (2,50e-1) -	3,0441e+1 (3,22e-1) -	3,0796e+1 (6,03e-1) -	3,3248e+1 (2,93e-1) $\approx$	3,0595e+1 (1,24e-1) -
WFG8	2,4887e+1 (4,67e-1) -	2,6837e+1 (2,11e-1) +	2,6383e+1 (2,96e-1) $\approx$	2,6243e+1 (4,29e-1) -	2,4371e+1 (2,02e-1) -	3,0150e+1 (1,14e-1) +
WFG9	3,1248e+1 (3,35e-1) $\approx$	3,1952e+1 (2,54e-1) +	3,0200e+1 (1,93e+0) $\approx$	3,0803e+1 (4,59e-1) $\approx$	3,2172e+1 (1,42e+0) +	3,3209e+1 (2,62e-1) +
+/-/ $\approx$	3/25/3	11/19/1	4/24/3	7/17/7	10/19/2	7/23/1

Tabela 9 – Frequência de classificação de cada algoritmo para o IGD, HV e Métrica *Ensemble*.

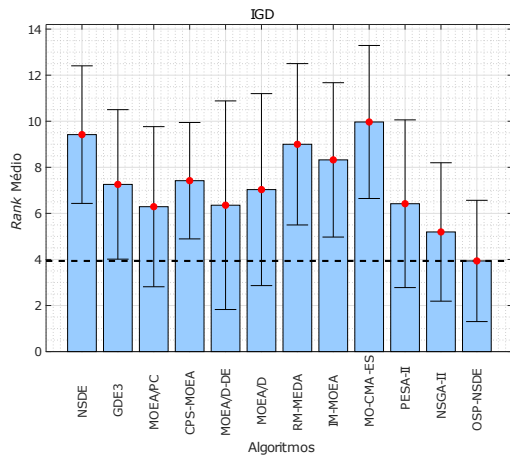
Algoritmos	<i>Rank</i>												
	1	2	3	4	5	6	7	8	9	10	11	12	13
IGD													
OSP-NSDE	9	3	5	3	3	3	3	1	1	0	0	0	0
MOEA/D-DE	7	4	0	3	0	1	3	2	1	1	4	1	4
MOEA/D	4	2	3	1	1	4	2	1	3	1	2	4	3
MOEA/PC	4	2	2	1	4	4	1	4	3	1	3	2	0
NSGA-II	3	4	1	7	4	2	3	1	3	2	0	1	0
PESA-II	2	5	2	1	3	4	2	1	2	4	3	1	1
IM-MOEA	1	2	2	1	3	0	2	5	2	3	4	3	3
RM-MEDA	1	1	2	0	1	3	2	1	3	2	4	9	2
CPS-MOEA	0	1	1	3	4	1	5	2	6	6	2	0	0
GDE3	0	0	6	2	3	4	2	1	3	3	3	4	0
MO-CMA-ES	0	0	2	1	0	3	1	4	2	1	2	3	12
NSDE	0	0	1	1	3	1	2	3	2	5	4	3	6
HV													
OSP-NSDE	7	4	5	2	2	3	4	2	0	2	0	0	0
NSGA-II	5	4	2	3	2	3	1	2	3	2	1	3	0
GDE3	5	4	2	2	2	0	4	2	2	2	1	5	0
MOEA/D	4	1	2	2	2	2	0	4	3	0	2	6	3
PESA-II	3	3	1	1	4	1	6	2	3	3	3	0	1
MOEA/PC	3	2	2	0	3	4	0	2	5	6	2	1	1
IM-MOEA	2	1	7	3	2	2	2	3	1	0	4	1	3
MOEA/D-DE	1	3	0	6	1	4	3	1	1	4	4	2	1
NSDE	1	1	1	2	1	2	1	0	3	6	6	2	5
CPS-MOEA	0	1	1	4	6	2	1	7	4	2	3	0	0
MO-CMA-ES	0	1	1	1	2	2	4	1	0	2	1	2	14
RM-MEDA	0	1	0	2	1	2	3	3	2	2	3	9	3
Métrica <i>Ensemble</i>													
OSP-NSDE	7	7	3	3	3	4	0	3	1	0	0	0	0
MOEA/D	7	5	2	3	1	2	0	1	0	2	2	4	2
MOEA/D-DE	4	2	3	2	0	0	4	2	5	3	0	4	2
NSGA-II	3	4	4	4	5	0	2	0	3	3	2	1	0
MOEA/PC	2	2	5	2	3	2	2	3	1	2	5	1	1
MO-CMA-ES	2	0	0	1	1	3	2	0	1	3	1	2	15
GDE3	1	5	6	3	1	3	4	1	3	1	2	1	0
CPS-MOEA	1	0	0	3	2	5	3	3	2	5	3	3	1
IM-MOEA	1	0	0	0	1	1	2	4	4	3	4	8	3
PESA-II	0	2	4	2	3	2	4	5	1	2	3	3	0
RM-MEDA	0	1	1	3	2	1	2	2	8	1	4	4	2
NSDE	0	0	1	1	5	3	3	1	1	6	5	0	5

No caso da métrica *Ensemble*, OSP-NSDE une-se a MOEA/D na primeira posição com a mesma frequência. No entanto, olhando mais de perto para as outras frequências das diferentes categorias, conclui-se que OSP-NSDE tem posições mais frequentes até a

classificação 9 e não tem presença da 10 à 13. Isto sugere que OSP-NSDE apresenta o comportamento mais robusto.

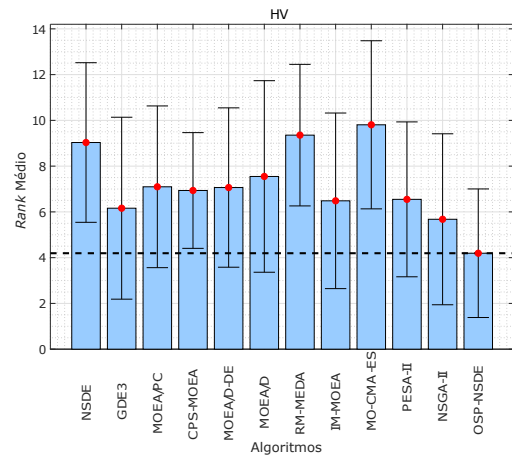
Outra maneira de avaliar o desempenho é comparando a média e desvio padrão das métricas de IGD, HV e *Ensemble* para todos os algoritmos; ver as Figuras 20, 21, e 22. A linha preta pontilhada representa o valor médio do OSP-NSDE. Nas Figuras 20, 21, e 22, observa-se que o OSP-NSDE atinge as melhores médias do *rank* em comparação com os outros algoritmos. As diferenças em favor do OSP-NSDE são acentuadas no *rank* da métrica de *Ensemble*.

Figura 20 – Classificação média baseada no IGD para todos os algoritmos testados. A linha tracejada mostra os valores médios do OSP-NSDE.



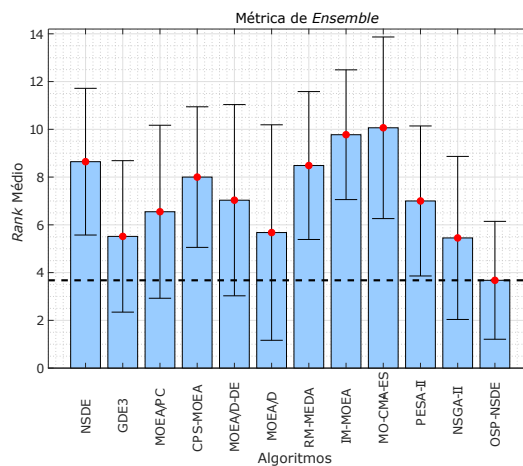
Fonte: Autor

Figura 21 – Classificação média baseada no HV para todos os algoritmos testados. A linha tracejada mostra os valores médios do OSP-NSDE.



Fonte: Autor

Figura 22 – Classificação média baseada na Métrica *Ensemble* para todos os algoritmos testados. A linha tracejada mostra os valores médios do OSP-NSDE.

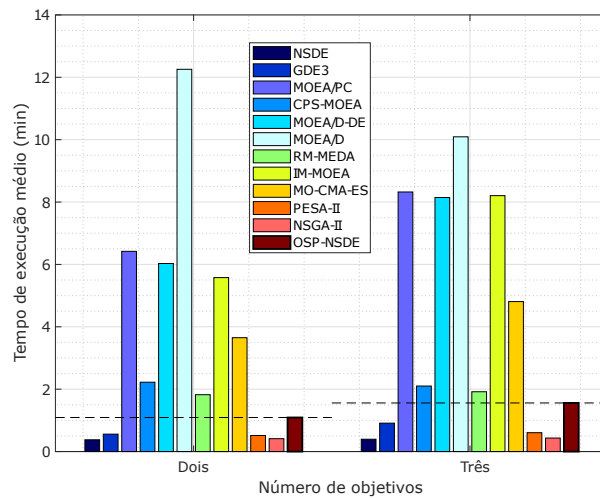


Fonte: Autor

Os tempos médios de execução dos algoritmos para dois grupos de funções de teste,

as funções de dois objetivos e de três objetivos, são mostrados na Figura 23. O tempo de execução do OSP-NSDE aumenta com o aumento do número de objetivos, embora o OSP-NSDE tenha um dos menores tempos de execução. MOEA/D-DE, MOEA/D, IM-MOEA e MOEA/PC são os algoritmos mais lentos. Além disso, o OSP-NSDE sempre foi mais rápido que dois outros métodos (CPS-MOEA e MO-CMA-ES), tão rápido quanto o RM-MEDA, e mais lento que o NSGA-II, NSDE, GDE3 e PESA-II. Apesar do fato de que o OSP-NSDE tem um tempo de execução maior que os quatro últimos algoritmos, ele apresenta melhor desempenho que todos os métodos comparados na maioria das funções. Vale ressaltar que o crescimento do tempo não é muito alto porque a estratégia do OSP é acionada algumas vezes; quando ela é mais eficaz para o processo evolucionário.

Figura 23 – Tempo médio de execução dos algoritmos para os diferentes grupos de funções de teste.



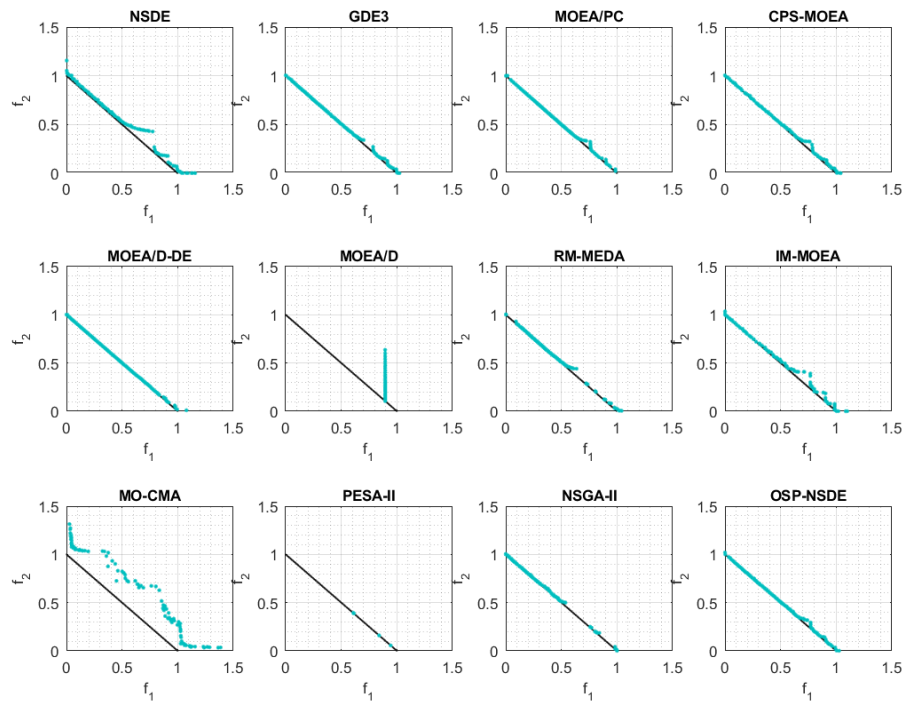
Fonte: Autor

As Figuras 24, 25, 26 e 27 mostram as melhores frentes de Pareto obtidas por cada algoritmo (pontos azuis) para quatro instâncias de teste: UF7, ZDT3, DTLZ7 e WFG6. Nas Figuras, os pontos pretos representam a verdadeira Frente Ótima de Pareto. Os gráficos corroboram os resultados e conclusões obtidos com a análise feita nas Tabelas 7 e 8.

O OSP-NSDE também foi comparado com algoritmos baseados em PSO. Como competidores foram selecionados o MOPSO (COELLO; LECHUGA, 2002), dMOPSO (MARTÍNEZ; COELLO, 2011), MMPSO (LIN et al., 2015), e CMOPSO (ZHANG et al., 2018). Todos os algoritmos foram executados no PlatEMO, com as mesmas condições experimentais anteriores. Os resultados foram determinados para 30 execuções de cada função de teste para cada algoritmo.

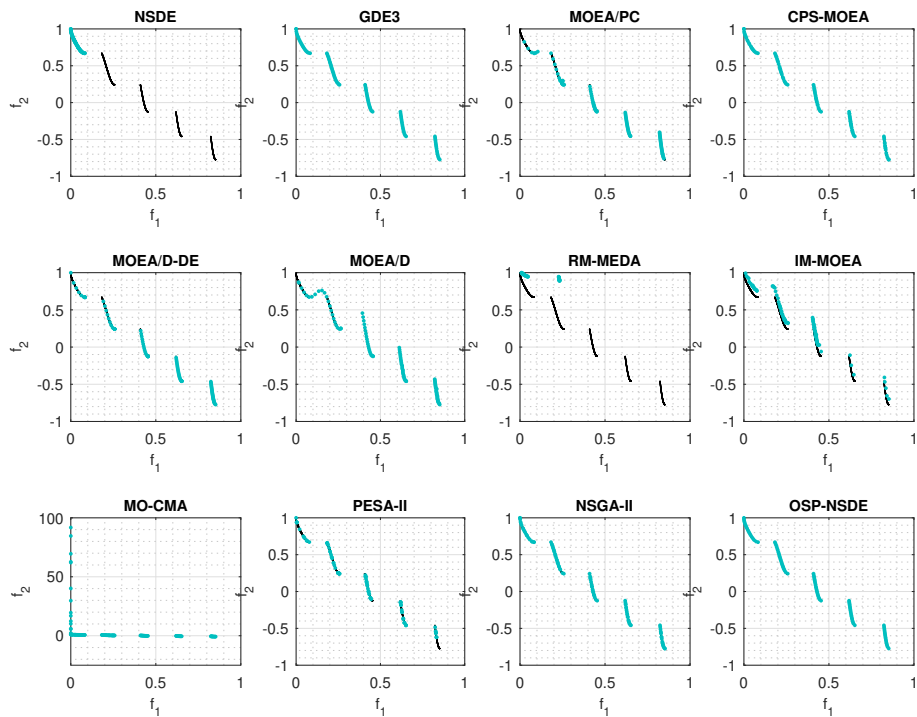
A Tabela 10 mostra o desempenho dos algoritmos considerando a métrica IGD. Pode-se observar que o OSP-NSDE supera os concorrentes na maioria das comparações individuais em relação ao indicador. O OSP-NSDE obteve o melhor desempenho para o conjunto de

Figura 24 – Frente de Pareto aproximada encontrada pelos algoritmos para UF7



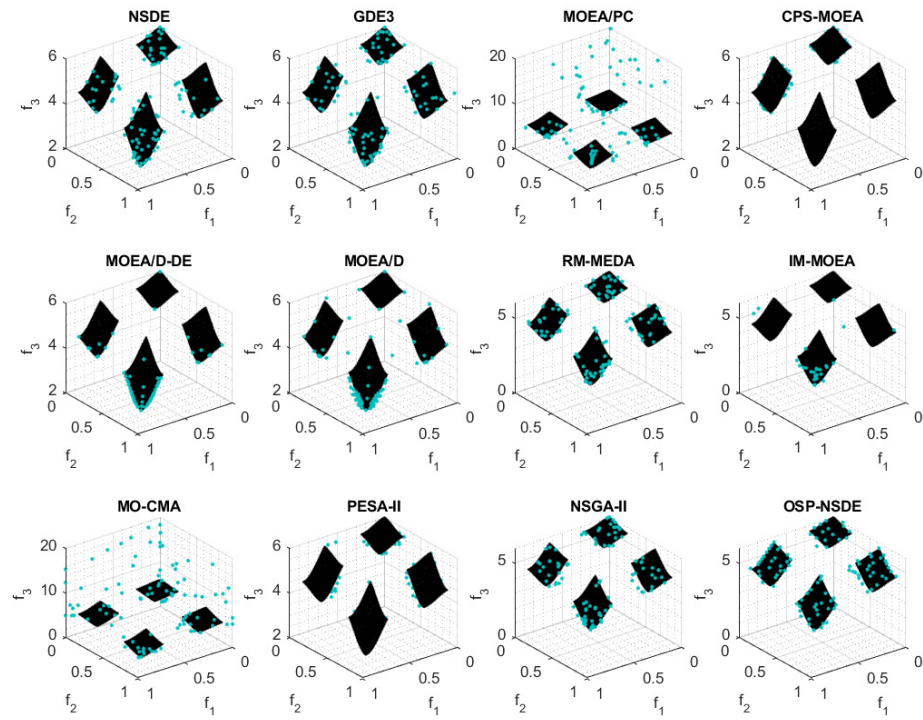
Fonte: Autor

Figura 25 – Frente de Pareto aproximada encontrada pelos algoritmos para ZDT3



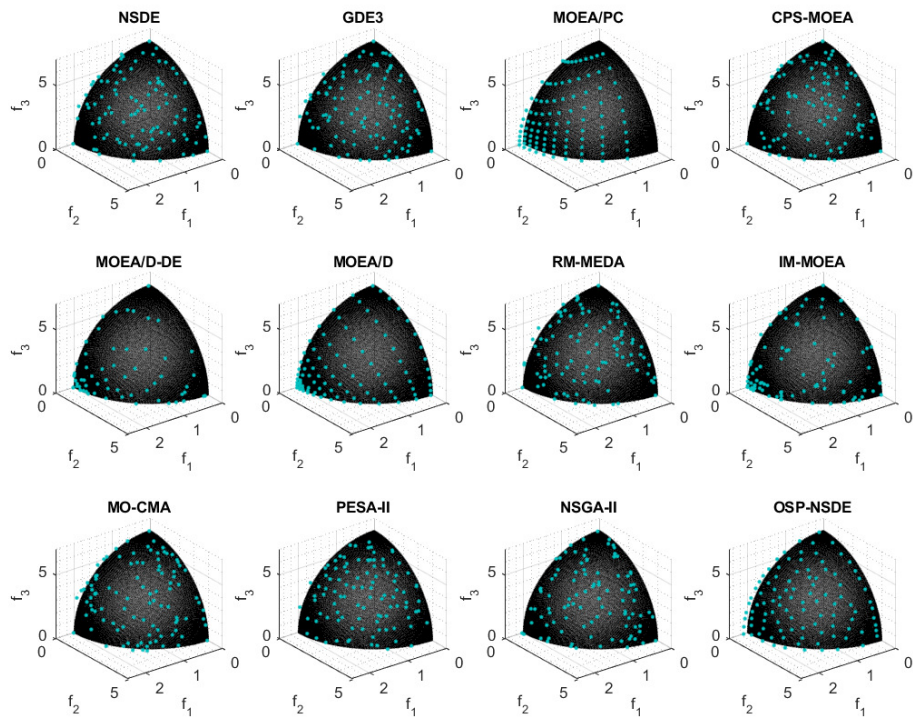
Fonte: Autor

Figura 26 – Frente de Pareto aproximada encontrada pelos algoritmos para DTLZ7



Fonte: Autor

Figura 27 – Frente de Pareto aproximada encontrada pelos algoritmos para WFG6



Fonte: Autor

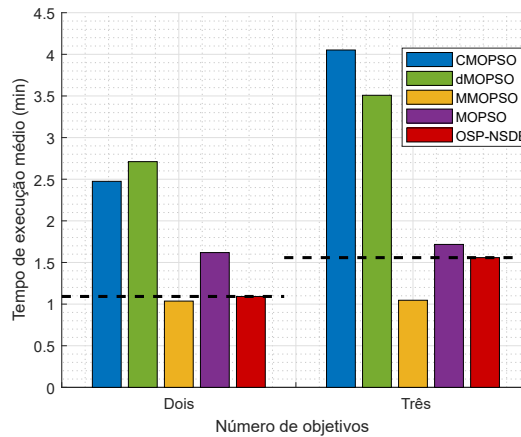
testes de UF e o DTLZ. O algoritmo proposto perdeu para o CMOPSO nas funções do WFG e empatou com o competidor nos problemas do ZDT.

Tabela 10 – Média e desvio padrão do IGD para os algoritmos baseados em PSO comparados. Os melhores resultados médios em cada instância estão em negrito. Os símbolos “+”, “ $\approx$ ” e “-” indicam onde o algoritmo na coluna correspondente funciona significativamente melhor que, equivalente a, ou pior do que o OSP-NSDE.

Problem	OSP-NSDE	MOPSO (2002)	dMOPSO (2011)	MMOPSO (2015)	CMOPSO (2018)
UF1	7,85e-3 (3,22e-3)	5,00e-1 (1,03e-1) -	2,85e-1 (7,55e-2) -	5,935e-2 (6,34e-2) -	3,15e-2 (3,14e-3) -
UF2	1,36e-2 (2,08e-3)	1,13e-1 (1,04e-2) -	8,63e-2 (1,26e-2) -	1,77e-2 (1,22e-3) -	2,41e-2 (4,32e-3) -
UF3	8,11e-2 (9,20e-2)	5,60e-1 (2,04e-2) -	3,20e-1 (6,47e-3) -	9,74e-2 (4,03e-2) -	1,20e-1 (2,44e-2) -
UF4	6,23e-2 (4,04e-3)	1,07e-1 (1,57e-2) -	1,14e-1 (8,35e-3) -	4,53e-2 (2,65e-3) +	7,01e-2 (3,53e-3) -
UF5	6,44e-1 (1,63e-1)	3,23e+0 (2,35e-1) -	2,31e+0 (2,95e-1) -	3,68e-1 (1,44e-1) +	2,67e-1 (1,25e-1) +
UF6	9,17e-2 (6,20e-2)	2,54e+0 (6,19e-1) -	1,36e+0 (2,68e-1) -	3,17e-1 (1,03e-1) -	9,34e-2 (1,61e-2) -
UF7	9,59e-3 (2,77e-3)	6,39e-1 (1,50e-1) -	2,94e-1 (1,11e-1) -	1,77e-1 (2,22e-1) -	1,86e-2 (2,08e-3) -
UF8	1,36e-1 (6,72e-2)	5,03e-1 (8,12e-2) -	3,21e-1 (2,99e-2) -	2,39e-1 (8,26e-2) -	7,52e-1 (1,31e-1) -
UF9	2,63e-1 (1,19e-1)	6,27e-1 (5,99e-2) -	6,02e-1 (3,22e-2) -	3,33e-1 (6,76e-2) -	9,85e-1 (2,02e-1) -
UF10	1,19e+0 (4,40e-1)	2,63e+0 (3,86e-1) -	9,46e-1 (0,00e+0) +	8,17e-1 (2,16e-1) +	3,51e+0 (7,49e-1) -
ZDT1	4,01e-3 (6,26e-5)	2,60e+1 (1,16e+1) -	7,05e-3 (1,07e-3) -	5,01e-3 (2,60e-4) -	3,88e-3 (3,10e-5) +
ZDT2	3,98e-3 (5,90e-5)	4,50e+1 (1,14e+1) -	4,59e-1 (2,79e-1) -	5,22e-3 (3,09e-4) -	3,99e-3 (2,71e-5) $\approx$
ZDT3	4,81e-3 (7,08e-5)	3,32e+1 (1,43e+1) -	7,55e-2 (1,81e-1) -	5,85e-3 (2,98e-4) -	4,99e-3 (6,96e-5) -
ZDT4	4,39e-3 (1,30e-4)	1,65e+1 (7,13e+0) -	4,18e-3 (1,99e-4) +	4,82e-3 (2,98e-4) -	3,85e-3 (2,13e-5) +
ZDT6	3,13e-3 (4,20e-5)	6,55e-3 (7,40e-4) -	3,10e-3 (3,77e-8) $\approx$	4,21e-3 (1,63e-4) -	3,30e-3 (2,41e-5) -
DTLZ1	2,49e-2 (8,89e-4)	3,32e+0 (1,34e+0) -	4,18e-1 (4,66e-1) -	2,81e-2 (1,29e-3) -	2,56e-2 (1,81e-4) -
DTLZ2	6,65e-2 (1,73e-3)	4,02e-1 (1,77e-2) -	1,29e-1 (7,57e-3) -	7,01e-2 (2,96e-3) -	5,78e-2 (1,04e-3) +
DTLZ3	6,40e-2 (1,55e-3)	4,19e+1 (2,73e+1) -	3,65e+0 (6,26e+0) -	7,08e-2 (2,23e-3) -	5,19e+0 (4,54e+0) -
DTLZ4	7,55e-2 (2,59e-3)	2,57e-1 (4,80e-2) -	2,29e-1 (4,39e-2) -	7,06e-2 (1,37e-3) +	6,09e-2 (1,65e-3) +
DTLZ5	1,11e-2 (9,88e-4)	1,02e-2 (1,43e-3) $\approx$	4,28e-2 (7,19e-3) -	6,50e-3 (9,36e-4) +	5,06e-3 (2,15e-4) +
DTLZ6	4,12e-3 (3,96e-5)	1,76e+0 (1,37e+0) -	3,39e-2 (7,53e-6) -	7,07e-3 (1,36e-3) -	4,20e-3 (5,93e-5) -
DTLZ7	1,06e-1 (9,82e-2)	1,56e+0 (1,04e+0) -	1,37e-1 (6,07e-3) -	8,67e-2 (7,81e-3) +	1,36e-1 (1,32e-1) -
WFG1	1,52e+0 (2,88e-2)	1,87e+0 (3,16e-1) -	1,53e+0 (8,33e-3) $\approx$	2,67e-1 (2,42e-2) +	1,19e+0 (2,19e-1) +
WFG2	1,84e-1 (8,60e-3)	2,63e-1 (2,13e-2) -	9,57e-1 (1,74e-1) -	2,08e-1 (1,02e-2) -	1,60e-1 (4,69e-3) +
WFG3	1,91e-1 (1,35e-2)	5,42e-1 (1,23e-1) -	4,18e-1 (1,17e-1) -	1,11e-1 (1,99e-2) +	1,48e-1 (1,28e-2) +
WFG4	2,83e-1 (4,76e-3)	8,08e-1 (6,88e-2) -	3,23e-1 (2,58e-2) -	2,84e-1 (1,20e-2) $\approx$	2,60e-1 (4,94e-3) +
WFG5	2,51e-1 (5,12e-3)	5,20e-1 (1,20e-1) -	2,64e-1 (1,46e-2) -	2,92e-1 (9,08e-3) -	2,53e-1 (6,16e-3) $\approx$
WFG6	2,70e-1 (1,72e-2)	6,82e-1 (1,20e-1) -	3,31e-1 (2,49e-2) -	3,29e-1 (6,90e-2) -	2,75e-1 (3,82e-2) -
WFG7	2,63e-1 (7,03e-3)	8,84e-1 (1,64e-1) -	3,83e-1 (1,49e-2) -	2,86e-1 (9,26e-3) -	2,30e-1 (3,95e-3) +
WFG8	3,60e-1 (7,99e-3)	1,02e+0 (4,58e-2) -	5,75e-1 (3,20e-2) -	3,70e-1 (1,25e-2) $\approx$	3,69e-1 (4,05e-3) $\approx$
WFG9	2,86e-1 (6,33e-2)	4,43e-1 (5,27e-2) -	2,63e-1 (5,49e-3) +	2,88e-1 (1,25e-2) $\approx$	2,18e-1 (3,25e-3) +
+ / - / $\approx$		0/30/1	3/26/2	8/20/3	12/16/3

Além disso, os tempos médios de execução dos algoritmos são mostrados na Figura 28. Pode-se observar que o OSP-NSDE sempre foi mais rápido que o MOPSO, dMOPSO e CMOPSO, e mais lento que o MMOPSO.

Figura 28 – Tempo médio de execução dos algoritmos baseados em PSO e o OSP-NSDE para os diferentes grupos de funções de teste.



Fonte: Autor

5.4 COMPARAÇÃO DE DESEMPENHO DO DOSP-NSDE COM DMOEAS BASEADOS EM PREVISÃO

Neste experimento, o DOSP-NSDE foi comparado com alguns algoritmos relevantes baseados em previsão. As configurações dos parâmetros são as mesmas da Subsecção 5.3.1.

A média e desvio padrão dos valores médios do IGD, SP e HV são mostrados nas Tabelas 11,12,13, respectivamente, em que os melhores valores estão em negrito. Os símbolos “+”, “ $\approx$ ” e “-”, indicam onde o concorrente trabalha significativamente melhor que, equivalente a, ou pior que o DOSP-NSDE.

Na Tabela 11, observa-se que o DOSP-NSDE supera os demais concorrentes na maioria das comparações individuais, ou seja, o número de sinais negativos é significativamente maior que o número de $\approx$ adicionados ao número de sinais + para os problemas de teste. Percebe-se que o DOSP-NSDE é melhor que as outras quatro estratégias em 6 das 8 funções. Para o problema do FDA2, todas as soluções são estatisticamente equivalentes. Para a função UDF2, a proposta apresentou desempenho inferior ao PPS. Esta função é um DMOP Tipo II complexo com POS polinomial associado a um deslocamento vertical.

Tabela 11 – Média e desvio padrão do IGD para os DMOEAs comparados.

Estratégias	FDA1	FDA2	FDA3	dMOP2
RIS	0,0307 $\pm$ 0,0013 -	0,0216 $\pm$ 0,0008 $\approx$	0,0827 $\pm$ 0,0034 -	0,0580 $\pm$ 0,0253 -
FPS	0,0174 $\pm$ 0,0013 -	0,0212$\pm$0,0003 $\approx$	0,0602 $\pm$ 0,0045 -	0,0614 $\pm$ 0,1297 -
PPS	0,0097 $\pm$ 0,0031 -	0,0217 $\pm$ 0,0005 $\approx$	0,0551 $\pm$ 0,0023 -	0,0114 $\pm$ 0,0027 -
OSP	0,0055 $\pm$ 0,0004	0,0214 $\pm$ 0,0004	0,0221$\pm$ 0,0010	0,0067$\pm$ 0,0026
Estratégias	UDF1	UDF2	UDF4	UDF5
RIS	0,1564 $\pm$ 0,0151 -	0,0320 $\pm$ 0,0021 -	0,1204 $\pm$ 0,0055 -	0,0289 $\pm$ 0,0029 -
FPS	0,1880 $\pm$ 0,0756 -	0,0168 $\pm$ 0,0014 -	0,1279 $\pm$ 0,0115 -	0,0231 $\pm$ 0,0024 -
PPS	0,2928 $\pm$ 0,1295 -	0,0085$\pm$0,0007 +	0,1530 $\pm$ 0,0658 -	0,0126 $\pm$ 0,0041 -
OSP	0,0491$\pm$ 0,0030	0,0138 $\pm$ 0,0035	0,0477$\pm$ 0,0036	0,0097$\pm$ 0,0005

Considerando o indicador SP (Tabela 12), observa-se que o DOSP-NSDE alcançou o melhor valor médio em 5 de 8 funções de teste. O DOSP-NSDE só perdeu as competições para a função UDF1. Corroborando os resultados na tabela do IGD, o algoritmo proposto não apresenta diferença estatística com seus concorrentes para o problema FDA2.

Tabela 12 – Média e desvio padrão do SP para os DMOEAs comparados.

Estratégias	FDA1	FDA2	FDA3	dMOP2
RIS	0,0073±0,0001 -	0,0288±0,0001 $\approx$	0,0134±0,0008 +	0,0064±0,0003 -
FPS	0,0067±0,0001 -	0,0270 ±4 × 10 ⁻⁵ $\approx$	0,0134±0,0009 +	0,0052±0,0002 -
PPS	0,0058±9 × 10 ⁻⁵ -	0,0274 ±2 × 10 ⁻⁵ $\approx$	0,0119±0,0008 +	0,0042 ±8 × 10 ⁻⁵ -
OSP	0,0035 ±4 × 10 ⁻⁵	0,0277 ±6 × 10 ⁻⁵	0,0158±0,0009	0,0035 ±0,0002
Estratégias	UDF1	UDF2	UDF4	UDF5
RIS	0,0252±0,0059 -	0,0199±0,0023 -	0,0179±0,0026 $\approx$	0,0178±0,0028 -
FPS	0,0194±0,0043 $\approx$	0,0107±0,0022 -	0,0180±0,0019 -	0,0129±0,0014 -
PPS	0,0169±0,0045 +	0,0054±0,0013 -	0,0166±0,0019 $\approx$	0,0084±0,0010 -
OSP	0,0191±0,0037	0,0045±0,0008	0,0162±0,0092	0,0071±0,0003

A Tabela 13, referente aos resultados obtidos usando a métrica HV, mostra que o DOSP-NSDE supera os concorrentes em 5 de 8 funções de teste. Para os demais problemas, a proposta não tem diferença estatística com seus adversários.

Tabela 13 – Média e desvio padrão do HV para os DMOEAs comparados.

Estratégias	FDA1	FDA2	FDA3	dMOP2
RIS	0,8673±0,0001 $\approx$	1,1663 ±2 × 10 ⁻⁵ $\approx$	1,1261±0,0008 -	0,6706±0,0019 -
FPS	0,8661±0,0010 $\approx$	1,1663 ±2 × 10 ⁻⁵ $\approx$	1,1199±0,0052 -	0,6555±0,0443 -
PPS	0,8596±0,0006 $\approx$	1,1662 ±4 × 10 ⁻⁵ $\approx$	1,1033±0,0018 -	0,6628±0,0005 -
OSP	0,8576±0,0003	1,1599 ±3 × 10 ⁻⁵	4,5123±0,0024	0,7596±0,0019
Estratégias	UDF1	UDF2	UDF4	UDF5
RIS	0,6041±0,0110 $\approx$	0,8407±0,0005 $\approx$	0,5682±0,0032 $\approx$	0,7277±0,0007 -
FPS	0,6031±0,0745 $\approx$	0,8379±0,0011 $\approx$	0,5638±0,0097 $\approx$	0,7247±0,0014 -
PPS	0,5649±0,0811 $\approx$	0,8306±0,0018 $\approx$	0,5360±0,0329 -	0,7168±0,0017 -
OSP	0,5913±0,0021	0,8512±0,0005	0,5845±0,0005	0,7692±0,0034

Os resultados nas Tabelas 11-13 mostram que o DOSP-NSDE alcança bons valores dos indicadores de desempenho para a maioria das funções de teste. O principal motivo pode estar associado ao fato de o DOSP-NSDE extrair informações tanto do espaço objetivo quanto do espaço de decisão. No primeiro, a informação estatística captura melhor o comportamento probabilístico do movimento da frente. No último, soluções mais dispersas são criadas na vizinhança dos melhores indivíduos, levando a um *trade-off* entre exploração e exploração.

Vale salientar que o custo computacional do algoritmo proposto parece muito caro, mas não é, porque a execução do OSP na fase estacionária não é aplicada em todas as gerações do processo evolutivo.

Outro experimento teve como objetivo comparar o desempenho do DOSP-NSDE com alguns algoritmos atuais baseados em previsão, como DSS (WU; JIN; LIU, 2015), MOEA/D-KF (MURUGANANTHAM; TAN; VADAKKEPAT, 2016), DMS (RUAN et al., 2017), CKPS (ZOU et al., 2017), e FIOPPS (CHEN et al., 2018), considerando os problemas de teste FDA1-FDA3 e dMOP2. A configuração dos parâmetros desses algoritmos foi a mesma usada em (CHEN et al., 2018).

A média e o desvio padrão dos valores do IGD é mostrada na Tabela 14. Percebe-se que o DOSP-NSDE apresenta desempenho significativamente melhor que seus concorrentes nos problemas FDA1 e dMOP2, e não há diferença estatística com o vencedor no problema do FDA3. O pior desempenho foi obtido para as funções do FDA2 (DMOP Tipo II), perdendo para todos os algoritmos. Os resultados sugerem que a proposta é competitiva em relação aos atuais DMOEAs baseados em predição.

Tabela 14 – Média e desvio padrão do IGD para os DMOEAs baseados em predição atuais.

Estratégias	FDA1	FDA2	FDA3	dMOP2
DSS	0,0079±0,0058 -	0,0113±0,0087+	0,0247±0,0139 -	0,0097±0,0052 -
MOEA/D-KF	0,0076±0,0061 -	0,0092±0,0092+	0,0247±0,0139 -	0,0080±0,0052 -
DMS	0,0078±0,0094 -	0,0164±0,0008+	0,0219±0,0082 ≈	0,0081±0,0052 -
CKPS	0,0096±0,0053 -	0,0104±0,0087+	0,0243±0,0128 -	0,0299±0,0253 -
FIOPPS	0,0069±0,0001 -	0,0216±4 × 10 ⁻⁵ ≈	0,0528±0,0029 -	0,0077±0,0004 -
OSP	0,0055±0,0004	0,0214± 0,0004	0,0221± 0,0010	0,0067± 0,0026

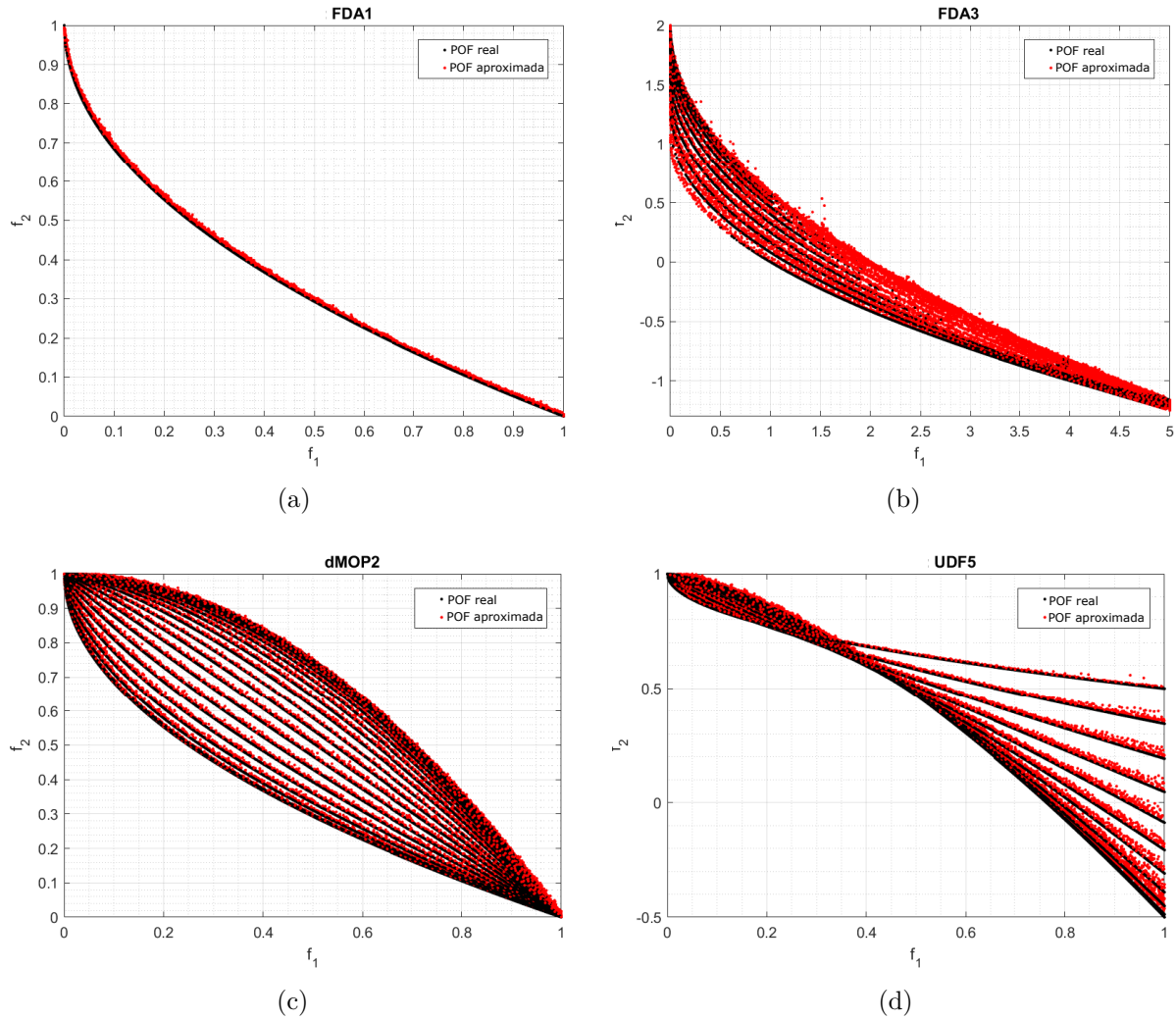
A Figura 29 mostra as melhores frentes de Pareto obtidas pelo DOSP-NSDE (pontos vermelhos) em cada mudança, para quatro instâncias de teste: dMOP2, FDA1, FDA3 e UDF5. Na figura, os pontos pretos representam a verdadeira Frente Ótima de Pareto. Pode ser visto que o DOSP-NSDE pode responder a mudanças ambientais com precisão, encontrando soluções com boas aproximações.

5.5 RESOLUÇÃO DE UM PROBLEMA DO MUNDO REAL: OTIMIZAÇÃO DINÂMICA DOS PARÂMETROS DO CONTROLADOR ADAPTATIVO LIVRE DE MODELO PARA REATOR DE TANQUE AGITADO CONTINUAMENTE (CSTR)

Nos últimos anos, os EAs foram empregados para resolver problemas do mundo real. O grau de dificuldade de um problema do mundo real pode ser desconhecido de antemão. Para tais problemas, o otimizador deve lidar com a não-linearidade e descontinuidade, restrições complexas, objetivos conflitantes, ruído e incerteza, entre outros desafios.

Na indústria, um sistema pode mudar com o tempo devido ao seu envelhecimento ou à aleatoriedade intrínseca de algumas plantas. Nesses casos, um controlador adaptativo é necessário para obter um bom desempenho do sistema de malha fechado projetado (FARINA; DEB; AMATO, 2004). Como exemplo, considera-se o controle do reator tanque continuamente agitado (CSTR). O CSTR é um tipo de reator químico no qual sucessivas

Figura 29 – Frente de Pareto aproximada obtida pelo DOSP-NSDE em todas as mudanças, nas funções: (a) FDA1; (b) FDA3; (c) DMOP2; (d) UDF5.



Fonte: Autor

reações químicas multicomponentes, isotérmicas e de fase líquida podem ser realizadas com sucesso. O conteúdo do reator é bem agitado e uniformemente distribuído ao longo do processo. Os reagentes de entrada (A,B) são misturados em um tanque com volume constante através de um agitador e o produto C é retornado. Nesta planta, o controle da concentração do produto C é feito ajustando a taxa de alimentação molar do reagente B. O comportamento dinâmico deste sistema CSTR não-linear de terceira ordem pode ser expresso pelas seguintes equações diferenciais (UÇAK; GÜNEL, 2019):

$$\begin{aligned}
 \dot{x}_1 &= 1 - 4x_1 + 0,5x_2^2 \\
 \dot{x}_2 &= -x_2 + 3x_1 - (0,5 + a(t))x_2^2 + u(t) \\
 \dot{x}_3 &= -x_3 + a(t)x_2^2 + d(t)
 \end{aligned} \tag{5.1}$$

onde $x_1(t)$, $x_2(t)$, e $x_3(t)$ denotam os estados obtidos das concentrações dos reagentes A e B e do produto C; $u(t)$ é o sinal de controle, e $x_3(t)$ é a saída do sistema. $a(t) =$

$1 + 0,1 \sin(0,032 * \pi * t)$ é um parâmetro dependente do tempo que pode ser variado para simular o envelhecimento acima mencionado ou a aleatoriedade inerente no sistema.

Neste estudo, foi usado o controle adaptativo livre de modelo (*Model-Free Adaptive Control: MFAC*) (HOU; JIN, 2013) para controlar o sistema não-linear com incertezas. MFAC é um método típico de controle direcionado por dados (*Data-Driven Control: DDC*). Neste tipo de controle, apenas os dados de entrada e saída da planta controlada são diretamente usados para projetar um controlador e para a análise de sistema em malha fechada, sem qualquer dinâmica de modelo envolvida.

Considere o seguinte sistema não-linear de uma entrada e uma saída (SISO) de tempo discreto (HOU; JIN, 2013):

$$y(k+1) = f(y(k), \dots, y(k-n_y), u(k), \dots, u(k-n_u)) \quad (5.2)$$

onde n_y, n_u são as ordens desconhecidas da saída $y(k)$ e da entrada $u(k)$ do sistema no instante k , respectivamente, e $f(\dots)$ é uma função não-linear desconhecida.

As seguintes hipóteses são consideradas para a planta controlada (HOU; JIN, 2013): H1: a derivada parcial de $f(\dots)$ em relação à entrada do controle $u(k)$ é contínua. H2: o sistema 5.2 é generalizado Lipschitz, isto é, $|\Delta y(k+1)| \leq b|\Delta u(k)|$ para qualquer k e $\Delta u(k) \neq 0$, com $\Delta y(k+1) = y(k+1) - y(k)$, $\Delta u(k) = u(k) - u(k-1)$. A constante b é positiva.

Um sistema não-linear de tempo discreto que satisfaz as hipóteses H1-H2 pode ser transformado em um modelo de linearização dinâmica de forma compacta (*Compact Form Dynamic Linearization - CFDL*). O CFDL constrói uma série de modelos lineares dinâmicos e equivalentes ao longo dos pontos dinâmicos de operação da planta controlada, em vez de identificar um único modelo global não-linear do processo. Assim, para o sistema não-linear 5.2 satisfazer às hipóteses H1 e H2, então deve existir uma derivada pseudo-parcial (*Pseudo-Partial-Derivative - PPD*) $\phi(k)$, tal que se $\Delta u(k) \neq 0$, o sistema 5.2 pode ser descrito como o seguinte modelo de CFDL (HOU; JIN, 2013):

$$y(k+1) = y(k) + \phi(k)\Delta u(k) \text{ com } |\phi(k)| \leq b \quad (5.3)$$

A Eq. 5.3 é um sistema linear dinâmico com um parâmetro de variação temporal lento, se $\Delta u(k) \neq 0$ e se taxa de mudança do sinal de entrada de controle $\Delta u(k)$ não for muito alta. Isto deve ser garantido para o projeto do sistema de controle.

Uma função objetivo é usada para a estimação dos parâmetros da PPD $\phi(k)$:

$$J(\phi(k)) = |y(k) - y(k-1) - \phi(k)\Delta u(k-1)|^2 + \mu|\phi(k) - \hat{\phi}(k-1)|^2 \quad (5.4)$$

onde μ é um fator de peso positivo. Assim, para qualquer instante k , pode-se estimar a PPD (usando apenas os dados de I/O da planta controlada) como:

$$\phi(k) = \phi(k-1) + \frac{\eta\Delta u(k-1)}{\mu + |\Delta u(k-1)|^2}(\Delta y(k) - \hat{\phi}(k-1)\Delta u(k-1)) \quad (5.5)$$

onde $\eta \in (0, 1)$ é o tamanho do passo.

Para que a condição $\Delta u(k) \neq 0$ seja satisfeita e, conseqüentemente fazer com que o algoritmo de estimação de parâmetros consiga rastrear o parâmetro variante no tempo, a seguinte condição de reinicialização deve ser considerada ((HOU; JIN, 2013)):

$$\begin{aligned} \hat{\phi}(k) = \hat{\phi}(1), \text{ se } |\hat{\phi}(k)| \leq \epsilon, \text{ ou } |\Delta u(k-1)| \leq \epsilon, \\ \text{ou } \text{sin}(\hat{\phi}(k)) \neq \text{sin}(\hat{\phi}(1)) \end{aligned} \quad (5.6)$$

onde $\hat{\phi}(1)$ é o valor inicial para o PPD estimado e ϵ é uma constante pequena positiva.

Para o esquema de controle, uma função ponderada de custo de entrada de um passo é adotada e dada por:

$$J(u(k)) = |y^*(k+1) - y(k+1)|^2 + \lambda |u(k) - u(k-1)|^2 \quad (5.7)$$

onde $y^*(k+1)$ é o sinal de referência ou sinal esperado da saída do sistema, e λ é um fator de peso positivo.

A substituição de 5.3 em 5.7 e a resolução da equação $\frac{\partial J(u(k))}{\partial u(k)} = 0$, fornece o sinal de controle seguinte (HOU; JIN, 2013):

$$u(k) = u(k-1) + \frac{\rho \phi(k)}{\lambda + |\phi(k)|^2} (y^*(k+1) - y(k+1)) \quad (5.8)$$

onde $\rho \in (0, 1)$ é o tamanho do passo.

Para encontrar os parâmetros ótimos do MFAC, foram usadas duas funções objetivos propostas em (SOUSA; FRANCA; ARAÚJO, 2015). A primeira (Eq. 5.9), definida como a média do sinal de erro acumulado ao quadrado, é usada para minimizar o erro de estado estacionário. A segunda função (Eq. 5.10) busca reduzir o máximo erro durante o estado de transição. Esta função é calculada usando a diferença quadrática acumulada entre as derivadas dos sinais de referência e de saída.

$$f_1(y^*(k), y(k)) = \frac{1}{N} \sum_{k=1}^N [e(k)^2], \text{ com } e(k) = y^*(k) - y(k) \quad (5.9)$$

$$f_2(y^{*'}(k), y'(k)) = \frac{1}{N-1} \sum_{k=1}^{N-1} [(y^{*'}(k) - y'(k))^2] \quad (5.10)$$

Um DMOEA produz uma frente de Pareto em cada ambiente, portanto, uma única solução deve ser selecionada, em cada mudança, entre as várias opções do conjunto de Pareto. Frequentemente, o ponto de joelho (*knee-point*) na POF é a solução escolhida se o tomador de decisão não tiver uma preferência especial. Portanto, emprega-se um método de *knee-point* (*Distance to the extreme line method*), que seleciona a solução (x^*) mais próxima de um dado ponto de referência, em geral, a origem de coordenadas (LIANG et al., 2018):

$$x^* = \min_{1 \leq i \leq NP} (d(x_i)) \mid d_i = \sqrt{f_{1N}(x_i)^2 + f_{2N}(x_i)^2} \quad (5.11)$$

onde f_{1_N} e f_{2_N} são as funções objetivo normalizadas.

O DOSP-NSDE foi executado para responder ao máximo de 10 de mudanças na dinâmica da planta, com uma população de 20 indivíduos. A gravidade da mudança (n_t) foi de 10 e a frequência de mudança $\tau_t = 10$. A configuração DOSP-NSDE inclui a probabilidade de mutação de $p_m = 1/D$ e o índice de distribuição $\eta_m = 20$ para o operador de mutação. A estratégia canônica do DE (rand/1), com tamanho de passo de mutação $F = 0,5$ e taxa de cruzamento $CR = 1$ também foram usados. Para o DOSP-NSDE, foi definido $\lambda = 0,1$, $\beta = 0,9$ e o horizonte de previsão inicial $p = 40$. Quando o ambiente muda, se o histórico de mudanças não for longo o suficiente para a previsão, a estratégia RIS é executada. As condições iniciais do sistema CSTR controlado são: $x_1 = 0,8796$; $x_2 = 0,8796$; $x_3 = 0,3467$; $u_1 = 0$; $\phi(\mathbf{1}) = 1$; e $\epsilon = 10^{-5}$. O método ode23tb^[2] do MatLab foi empregado nas simulações com o período de amostragem de 0,1s.

O controlador foi projetado para o rastreamento dos seguintes sinais de referência:

1. Sinal de escada

$$y^*(k+1) = \begin{cases} 0,4, & k \leq 100 \text{ e } 400 < k \leq 500 \\ 0,5, & 100 < k \leq 200 \text{ e } 300 < k \leq 400 \\ 0,6, & 200 < k \leq 300 \end{cases}$$

2. Sinal senoidal

$$y^*(k+1) = 0,15 \cdot \text{sen}(2 * \pi / f(k)) + 0,5$$

$$f(k) = \begin{cases} 310, & 0 < k \leq 310 \\ 190, & 310 < k \leq 500 \\ 100, & 500 < k \leq 600 \\ 80, & 600 < k \leq 680 \\ 30, & 680 < k \leq 710 \\ 40, & 710 < k \leq 750 \\ 30, & 750 < k \leq 780 \\ 20, & 780 < k \leq 800 \end{cases}$$

3. Sinal Misto

$$y^*(k+1) = \begin{cases} 0,2(-1)^{\text{round}(k/100)} + 0,5, \\ \text{se } 0 < k \leq 300 \text{ e } k > 700 \\ 0,2\text{sen}(k\pi/100) + 0,1\cos(k\pi/50) + 0,5, \\ \text{se } 300 < k \leq 700 \end{cases}$$

^[2] Resolve equações diferenciais rígidas usando regra trapezoidal e a fórmula de diferenciação inversa

O desempenho do rastreamento do controlador MFAC para os três sinais de referência e seus sinais de controle são ilustrados na Figura 30. Pode ser visto que o controlador tem um desempenho de rastreamento bem-sucedido para o sinal de referência da escada em cada mudança da dinâmica do sistema. O pior desempenho é obtido no rastreamento do sinal de referência senoidal. Para isso, o DOSP-NSDE não tem tempo suficiente, entre as mudanças, para encontrar o melhor conjunto de parâmetros do MFAC. Isso sugere que, para mudanças rápidas, o DOSP-NSDE pode não ser a melhor escolha para o ajuste paramétrico do MFAC no rastreamento de sinais do tipo sinusoidal. O controlador MFAC com ajuste de parâmetros usando DOSP-NSDE teve desempenho aceitável rastreando o sinal de referência misto.

Além disso, o desempenho do rastreamento do controlador MFAC é comparado com a configuração de parâmetros encontrada por diferentes otimizadores multiobjetivo. Para isso, foram selecionados o NSGA-II (DEB et al., 2002), NSGA-III (DEB; JAIN, 2014), e MOEA/D (ZHANG; LI, 2007) como otimizadores para comparar com o DOSP-NSDE. Para fazer uma comparação justa, usou-se nos otimizadores selecionados, o mesmo método de detecção de mudanças e a mesma estratégia de reação às mudanças (estratégia OSP) empregada pelo DOSP-NSDE. O conjunto de parâmetros do controlador MFAC é selecionado a partir do PF encontrado por cada algoritmo usando um tomador de decisão baseado em ponto de joelho. As avaliações foram realizadas em relação aos desempenhos de rastreamento das soluções encontradas por cada DMOEA.

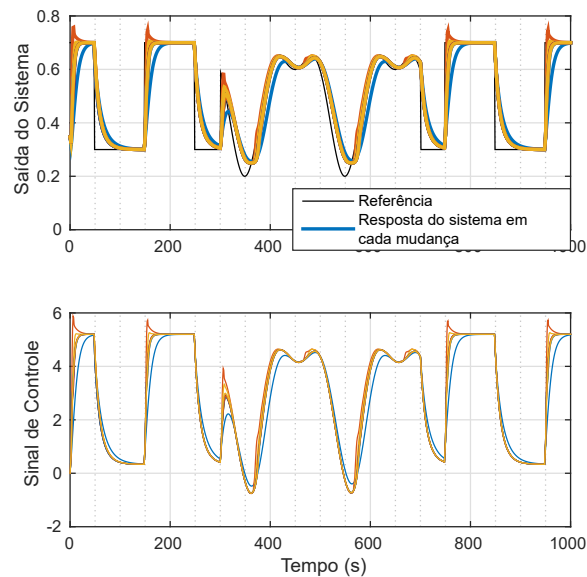
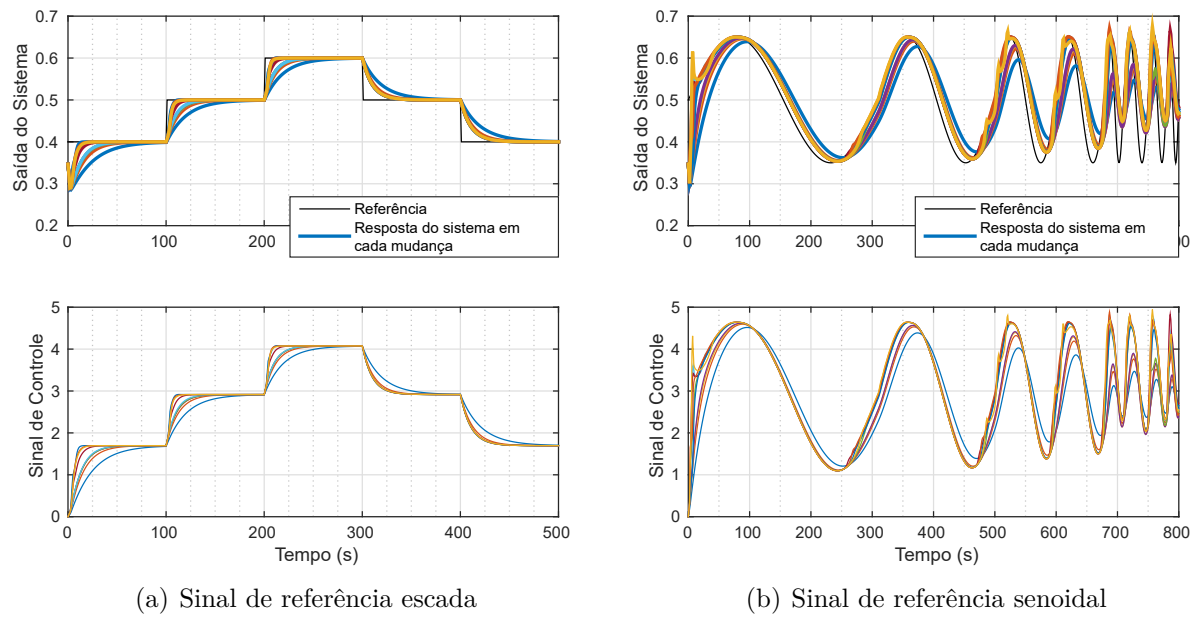
A média e a variância da Soma dos Quadrados dos Erros (*Error Sum of Squares: SSE*) e a média e a variância do sinal de controle para o controlador MFAC com parâmetros encontrados por todos os otimizadores são mostradas na Tabela 15. Nota-se que todos os controladores têm desempenho de controle aceitável e adequado. Pode-se observar que a solução encontrada pelo DOSP-NSDE apresentou o melhor desempenho em comparação com a solução determinada pelo NSGA-II e MOEA/D, e o DOSP-NSDE apresentou desempenho similar ao NSGA-III.

Tabela 15 – Média e variância da Soma dos Quadrados dos Erros (SSE), e média e variância do sinal de controle para o controlador MFAC ajustado com os parâmetros encontrados pelos otimizadores comparados.

	Escada		Senoidal		Misto	
	μ	σ^2	μ	σ^2	μ	σ^2
Soma dos Quadrados dos Erros (SSE)						
NSGA-II	0,2382	0,0024	6,5719	0,6435	15,8026	16,5613
NSGA-III	0,3141	0,0021	2,4786	0,7421	6,9851	2,5379
MOEA/D	0,7909	0,7117	5,2524	8,7833	14,3218	88,5280
DOSP-NSDE	0,2405	0,0013	2,2126	0,7771	6,2571	1,8860
Sinal de Controle						
NSGA-II	2,6496	0,8543	2,7809	0,7600	2,7456	2,7393
NSGA-III	2,6196	0,8744	2,9505	1,2996	2,8863	4,0678
MOEA/D	2,5527	0,9144	2,8030	0,9214	2,8177	2,8947
DOSP-NSDE	2,6301	0,8676	2,9921	1,3421	2,9341	4,1276

Os resultados mostrados na tabela sugerem que os parâmetros encontrados no DOSP-NSDE permitem que o controlador MFAC supere seus concorrentes para a maioria dos sinais de referência. Em relação ao sinal de controle, observamos valores semelhantes na média em todas as soluções.

Figura 30 – Rastreamento do controlador MFAC para os sinais de referência e seus sinais de controle, usando DOSP-NSDE para determinar os parâmetros ótimos do MFAC.



Fonte: Autor

6 CONCLUSÃO

Nesta tese foi elaborada uma estratégia (OSP) usando os conceitos da modelagem probabilística, que foi inserida em EAs para resolver problemas de otimização multiobjetivos dinâmicos. A estratégia permite rastrear e prever as mudanças de paisagem, bem como acelerar a convergência de um algoritmo antes de uma nova modificação do ambiente, melhorando também a qualidade das soluções obtidas em relação aos algoritmos da literatura. O algoritmo proposto, Evolução Diferencial com Ordenação pela Não Dominância com Previsão no Espaço Objetivo (DOSP-NSDE), usa o OSP tanto na fase estática do processo evolucionário (quando não há mudanças no ambiente) como no mecanismo de reação à mudança para prever a nova localização da frente ótima. O algoritmo extrai informações do espaço de decisão e do espaço objetivo. A informação estatística levantada a partir do espaço objetivo (processo de previsão empregando um modelo de séries temporais AR) reflete melhor o comportamento probabilístico do movimento das soluções não dominadas e, assim, aumenta a velocidade de convergência do algoritmo em relação às gerações. As informações geradas a partir do espaço de decisão (o processo de completar a nova população usando o GMM e a VI nos dá uma população com soluções mais dispersas na vizinhança da frente prevista. O algoritmo emprega uma condição de previsão com base na taxa de crescimento do HV aproximado para acionar o OSP. O critério envolve um limiar adaptativo e um intervalo de gerações variável para a previsão.

Da revisão do estado da arte dos Algoritmos Evolucionários Multiobjetivos em Ambientes Dinâmicos, foi proposta uma nova taxonomia baseada na forma em que os MOEAs dinâmicos inserem diversidade. Baseados nessa nova classificação, observa-se que os métodos exógenos, especialmente os métodos baseados em previsão, foram promissores para aumentar a diversidade e assim lidar melhor com a nova paisagem, aumentando a velocidade da convergência, isto é, diminuir o número de gerações que o algoritmo leva para determinar soluções de qualidade.

Da revisão do estado da arte da Modelagem Probabilística, foram escolhidos e aplicados os métodos probabilísticos que dão suporte ao algoritmo dinâmico proposto nesta tese. O estudo realizado mostrou que o GMM oferece uma base probabilística sólida para a análise das populações de um EA, pois já foi demonstrado que o uso de um conjunto de gaussianas pode aproximar qualquer distribuição de um conjunto de dados. Além disso, o uso do modelo de série temporal AR, permite realizar a previsão nos EAs.

Os resultados experimentais sugerem que o DOSP-NSDE é competitivo na maioria dos problemas de teste considerando as três métricas utilizadas. O desempenho do DOSP-NSDE tem várias explicações. Em primeiro lugar, o uso da mutação polinomial junto aos operadores DE, melhorou significativamente a POF obtida pelo algoritmo nos estados estacionários do processo de otimização. No entanto, o elemento principal é o método OSP,

cuja eficácia foi demonstrada nos resultados experimentais. Esta estratégia, aplicada no momento adequado, aumentou a velocidade da convergência do algoritmo e melhorou a qualidade das soluções.

Os experimentos realizados para analisar a influência da estratégia OSP na parte estacionária do processo de otimização, sugerem que o MOEA OSP-NSDE superou significativamente os concorrentes baseados em EA e os concorrentes baseados em PSO na maioria dos problemas de teste. Para as três métricas de desempenho, o OSP-NSDE apresentou a melhor média de todos os algoritmos comparados. O OSP-NSDE teve um tempo de execução menor em comparação com nove algoritmos e um tempo de execução mais longo que os outros quatro. O tempo de execução aumenta devido à execução da estratégia do OSP. No entanto, esse aumento não foi alto porque o OSP só é acionado quando é mais útil para o processo evolucionário.

Os resultados também sugerem que a estratégia do OSP como mecanismo de reação à mudança, é mais apropriada para uso com MOEAs baseados em Pareto, como o NSGA-II ou o NSDE. Além disso, a análise de sensibilidade à frequência da mudança sugere capacidade de rastrear mudanças ambientais quando a frequência da mudança aumenta. Finalmente, a proposta parece ser sensível à severidade da mudança.

A complexidade computacional total do DOSP-NSDE é alta. No entanto, é importante observar que o pior caso ocorre apenas algumas vezes devido ao critério baseado em HV. Este demonstra que mesmo procedimentos complexos podem ser inseridos em tempo hábil e trazer poucos custos para os modelos.

A eficácia do DOSP-NSDE também foi verificada em um problema do mundo real. Com base nos resultados experimentais, pode-se dizer que o ajuste dos parâmetros do MFAC com o algoritmo proposto permite que o MFAC tenha um desempenho de controle aceitável e adequado.

6.1 LIMITAÇÕES E TRABALHOS FUTUROS.

Experimentos preliminares sugerem que o desempenho do DOSP-NSDE é degradado com um aumento no número de funções objetivo. O problema principal é que a estratégia do OSP (no processo estacionário) se torna pouco executada devido ao critério de acionamento baseado no HV. Contudo, quando o número de objetivos é grande, a taxa de soluções não dominadas aumenta rápida e significativamente. Por esse motivo, o teste que estabelece que a percentagem de soluções não dominadas deve estar abaixo de 90% é facilmente alcançável, impedindo a execução da estratégia.

Outra limitação do algoritmo proposto é o uso da métrica HV como parte do mecanismo de acionamento da estratégia OSP durante a resolução de problemas de otimização com muitos objetivos (MaOPs), problemas que têm mais de 4 funções objetivo. Neste caso, o HV pode tornar o algoritmo computacionalmente caro para trabalhar com os MaOPs.

Ainda relacionado à resolução de MaOPs, o desempenho dos MOEAs baseados em Pareto deteriora-se rapidamente à medida que o número de objetivos aumenta, principalmente devido ao fenômeno conhecido como resistência à dominância (PURSHOUSE; FLEMING, 2007). Em outras palavras, o domínio de Pareto perde a pressão de seleção quando aborda os MaOPs.

Para abordar estas limitações e algumas outras referentes a como o algoritmo lida com as mudanças do ambiente, várias adaptações ao DOSP-NSDE podem ser consideradas tais como:

1. Investigar outros indicadores para usar como acionador da condição OSP. Indicadores de qualidade intimamente relacionados a HV podem ser usados, como o R2, *Inverted Generational Distance Plus* (IGD+) ou o indicador ϵ (LIEFOOGHE; DERBEL, 2016), ou outra aproximação para HV como as propostas em (ISHIBUCHI et al., 2009) e em (BADER; ZITZLER, 2011).
2. Usar outro método para classificar as soluções, ao invés da dominância de Pareto. Assim, pode se usar métodos como *Grid Dominance* (YANG et al., 2013), ou α -*Dominance Expansion* (LIU et al., 2017), entre outros.
3. Fixar o número de soluções não dominadas que serão usadas na previsão, em vez de usar todas as soluções não dominadas encontradas. Isso tornaria o algoritmo mais rápido em relação ao tempo.
4. Investigar outros métodos para realizar a otimização intermediária na estratégia do OSP, o que tornaria o algoritmo menos complexo computacionalmente.
5. Melhorar o DOSP-NSDE para lidar com mudanças indetectáveis, ocorrência em algumas aplicações, como o controle de sistemas incertos ou sistemas ruidosos.

6.2 TRABALHOS PUBLICADOS

Como parte da pesquisa desenvolvida foram publicados os seguintes artigos:

- Em Periódicos:
 1. **"Multi-Objective Evolutionary Algorithm with Prediction in the Objective Space"** dos autores **Elaine Guerrero-Peña e Aluizio F.R. Araújo** no **Information Sciences**, 2019.

- Em conferência internacionais de Qualis A1:

1. *"Searching a Probabilistic Model for Differential Evolution Population"* dos autores *Elaine Guerrero-Peña, Wagner J. F. Silva e Aluizio F.R. Araújo* no *IEEE Congress on Evolutionary Computation (IEEE CEC)* realizado em Vancouver, Canadá, 24-29 Julho 2016.
2. *"A Gaussian Mixture Model Based Local Search for Differential Evolution Algorithm"* dos autores *Elaine Guerrero-Peña e Aluizio F.R. Araújo* no *IEEE Congress on Evolutionary Computation (IEEE CEC)* realizado em San Sebastian, Espanha, 05-08 Junho 2017.
3. *"Multi-Objective Evolutionary Algorithm with Gaussian Process Regression"* dos autores *Elaine Guerrero-Peña, e Aluizio F.R. Araújo* no *IEEE Congress on Evolutionary Computation (IEEE CEC)* realizado no Rio de Janeiro, Brasil, 08-13 Julho 2018.

- Em conferência internacionais de Qualis A2:

1. *"A New Dynamic Multi-objective Evolutionary Algorithm without Chan-ge Detector"* dos autores *Elaine Guerrero-Peña e Aluizio F.R. Araújo*, no *IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2019)* realizado no Bari, Itália, 06-09 Outubro 2019.

- Em conferências nacionais:

1. *"Adaptação Paramétrica Evolucionária no Controle Adaptativo Livre de Modelo"* dos autores *Elaine Guerrero-Peña e Aluizio F.R. Araújo*, no *14 Simpósio Brasileiro de Automação Inteligente (SBAI 2019)* realizado em Ouro Preto, Minas Gerais, Brasil, 27-30 Outubro 2019.

Outros trabalhos relacionados à pesquisa foram desenvolvidos e estão submetidos para periódicos e conferências. Estes são:

1. *"Taxonomy of Dynamic Multi-Objective Evolutionary Algorithms: A Survey"* dos autores *Elaine Guerrero-Peña, Aluizio F.R. Araújo, e Cícero Garrozi*, submetido para o periódico *ACM Survey*.
2. *"Dynamic Multi-Objective Evolutionary Algorithm with Objective Space Prediction Strategy"* dos autores *Elaine Guerrero-Peña e Aluizio F.R. Araújo*, submetido para o periódico *Applied Soft Computing*.

3. *"Evolutionary Trajectory Planning with Obstacles for a Mobile Manipulator"* dos autores *Antônio R. C. Gonçalves, Elaine Guerrero-Peña, Aluizio F.R. Araújo, e Adrien Durand-Petiteville*, submetido para *International Federation of Automatic Control World Congress (IFAC 2020)*.

REFERÊNCIAS

- AHMED, A.; ALY, M.; GONZALEZ, J.; NARAYANAMURTHY, S.; SMOLA, A. J. Scalable inference in latent variable models. In: ACM. *Proceedings of the fifth ACM international conference on Web search and data mining*. [S.l.], 2012. p. 123–132.
- AKAIKE, H. Fitting autoregressive models for prediction. *Annals of the institute of Statistical Mathematics*, Springer, v. 21, n. 1, p. 243–247, 1969.
- ANGIRA, R.; BABU, B. Non-dominated Sorting Differential Evolution (NSDE): An Extension of Differential Evolution for Multi-objective Optimization. In: *IICA*. [S.l.: s.n.], 2005. p. 1428–1443.
- ANTONIAK, C. E. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *The annals of statistics*, JSTOR, p. 1152–1174, 1974.
- AUGER, F.; HILAIRET, M.; GUERRERO, J. M.; MONMASSON, E.; ORLOWSKA-KOWALSKA, T.; KATSURA, S. Industrial applications of the Kalman filter: A review. *IEEE Transactions on Industrial Electronics*, IEEE, v. 60, n. 12, p. 5458–5471, 2013.
- AVDAGIĆ, Z.; KONJICIJA, S.; OMANOVIĆ, S. Evolutionary approach to solving non-stationary dynamic multi-objective problems. In: *Foundations of Computational Intelligence Volume 3*. [S.l.]: Springer, 2009. p. 267–289.
- AZEVEDO, C. R.; ARAÚJO, A. F. Correlation between diversity and hypervolume in evolutionary multiobjective optimization. In: IEEE. *2011 IEEE Congress of Evolutionary Computation (CEC)*. [S.l.], 2011. p. 2743–2750.
- AZEVEDO, C. R.; ARAÚJO, A. F. Generalized immigration schemes for dynamic evolutionary multiobjective optimization. In: IEEE. *Evolutionary Computation (CEC), 2011 IEEE Congress on*. [S.l.], 2011. p. 2033–2040.
- AZZOUZ, R.; BECHIKH, S.; Ben Said, L. A Multiple Reference Point-based evolutionary algorithm for dynamic multi-objective optimization with undetectable changes. *Proceedings of the 2014 IEEE Congress on Evolutionary Computation, CEC 2014*, p. 3168–3175, 2014.
- AZZOUZ, R.; BECHIKH, S.; SAID, L. B. A dynamic multi-objective evolutionary algorithm using a change severity-based adaptive population management strategy. *Soft Computing*, Springer, p. 1–22, 2015.
- AZZOUZ, R.; BECHIKH, S.; SAID, L. B. Multi-objective optimization with dynamic constraints and objectives: new challenges for evolutionary algorithms. In: ACM. *Proceedings of the 2015 on Genetic and Evolutionary Computation Conference*. [S.l.], 2015. p. 615–622.
- AZZOUZ, R.; BECHIKH, S.; SAID, L. B. Dynamic Multi-objective Optimization Using Evolutionary Algorithms: A Survey. In: *Recent Advances in Evolutionary Multi-objective Optimization*. [S.l.]: Springer, 2017. p. 31–70.

- BADER, J.; ZITZLER, E. HypE: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary computation*, MIT Press, v. 19, n. 1, p. 45–76, 2011.
- BAR-SHALOM, Y.; LI, X. R.; KIRUBARAJAN, T. *Estimation with applications to tracking and navigation: theory algorithms and software*. [S.l.]: John Wiley & Sons, 2004.
- BINGUL, Z. Adaptive genetic algorithms applied to dynamic multiobjective problems. *Applied Soft Computing*, Elsevier, v. 7, n. 3, p. 791–799, 2007.
- BISHOP, C. Pattern Recognition and Machine Learning. 2nd printing edn. *Springer, New York*, 2007.
- BISWAS, S.; DAS, S.; SUGANTHAN, P. N.; COELLO, C. A. C. Evolutionary multiobjective optimization in dynamic environments: A set of novel benchmark functions. In: IEEE. *2014 IEEE Congress on Evolutionary Computation (CEC)*. [S.l.], 2014. p. 3192–3199.
- BLEI, D. M. Build, compute, critique, repeat: Data analysis with latent variable models. *Annual Review of Statistics and Its Application*, Annual Reviews, v. 1, p. 203–232, 2014.
- BRANKE, J.; BRANKE, J. U. Memory Enhanced Evolutionary Algorithms for Changing Optimization Problems. In: *Congress on Evolutionary Computation {CEC}99*. [S.l.]: IEEE, 1999.
- BUI, L. T.; ABBASS, H. A.; BRANKE, J. Multiobjective optimization for dynamic environments. In: IEEE. *2005 IEEE Congress on Evolutionary Computation*. [S.l.], 2005. v. 3, p. 2349–2356.
- BUI, L. T.; NGUYEN, M.-H.; BRANKE, J.; ABBASS, H. A. Tackling dynamic problems with multiobjective evolutionary algorithms. In: *Multiobjective problem solving from nature*. [S.l.]: Springer, 2008. p. 77–91.
- CÁMARA, M. *Parallel Processing for Dynamic Multi-objective Optimization*. Phd Thesis (PhD Thesis) — Ph. D Dissertation, Department of Computer Architecture and Computer, 2010.
- CÁMARA, M.; ORTEGA, J.; TORO, F. J. Parallel processing for multi-objective optimization in dynamic environments. In: IEEE. *Parallel and Distributed Processing Symposium, 2007. IPDPS 2007. IEEE International*. [S.l.], 2007. p. 1–8.
- CHEN, D.; ZOU, F.; LU, R.; WANG, X. A hybrid fuzzy inference prediction strategy for dynamic multi-objective optimization. *Swarm and Evolutionary Computation*, Elsevier, 2018.
- CHEN, H.; LI, M.; CHEN, X. Using diversity as an additional-objective in dynamic multi-objective optimization algorithms. *2nd International Symposium on Electronic Commerce and Security, ISECS 2009*, v. 1, p. 484–487, 2009.
- CHEN, M.; GUO, Y.; LIU, H.; WANG, C. The Evolutionary Algorithm to Find Robust Pareto-Optimal Solutions over Time. *Mathematical Problems in Engineering*, Hindawi, v. 2015, 2015.

- CHEN, R.; LI, K.; YAO, X. Dynamic Multiobjectives Optimization With a Changing Number of Objectives. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 22, n. 1, p. 157–171, 2018.
- CHEN, S. Kalman filter for robot vision: a survey. *IEEE Transactions on Industrial Electronics*, IEEE, v. 59, n. 11, p. 4409–4420, 2011.
- CHEN, X.; ZHANG, D.; ZENG, X. A Stable Matching-Based Selection and Memory Enhanced MOEA/D for Evolutionary Dynamic Multiobjective Optimization. *2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI)*, p. 478–485, 2015. Available at: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7372173>>.
- CHENG, R.; JIN, Y.; NARUKAWA, K. Adaptive reference vector generation for inverse model based evolutionary multiobjective optimization with degenerate and disconnected pareto fronts. In: SPRINGER. *International Conference on Evolutionary Multi-Criterion Optimization*. [S.l.], 2015. p. 127–140.
- CHENG, R.; JIN, Y.; NARUKAWA, K.; SENDHOFF, B. A multiobjective evolutionary algorithm using Gaussian process-based inverse modeling. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 19, n. 6, p. 838–856, 2015.
- COBB, H. G. *An investigation into the use of hypermutation as an adaptative operator in genetic algorithms having continuous*. Naval Res. Lab: Washington, 1990.
- COELLO, C. A. C. *Recent trends in evolutionary multiobjective optimization*. [S.l.]: Evolutionary multiobjective optimization: theoretical advances and applications, 2005. 85233–85787 p.
- COELLO, C. C.; LECHUGA, M. S. MOPSO: A proposal for multiple objective particle swarm optimization. In: IEEE. *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*. [S.l.], 2002. v. 2, p. 1051–1056.
- COHON, J. L.; CHURCH, R. L.; SHEER, D. P. Generating multiobjective trade-offs: An algorithm for bicriterion problems. *Water Resources Research*, Wiley Online Library, v. 15, n. 5, p. 1001–1010, 1979.
- COLLINS, M.; DASGUPTA, S.; SCHAPIRE, R. E. A generalization of principal components analysis to the exponential family. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2001. p. 617–624.
- CORNE, D. W.; JERRAM, N. R.; KNOWLES, J. D.; OATES, M. J. PESA-II: Region-based selection in evolutionary multiobjective optimization. In: MORGAN KAUFMANN PUBLISHERS INC. *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*. [S.l.], 2001. p. 283–290.
- DANTZIG, G. B.; ORDEN, A.; WOLFE, P.; OTHERS. The generalized simplex method for minimizing a linear form under linear inequality restraints. *Pacific Journal of Mathematics*, v. 5, n. 2, p. 183–195, 1955.
- DEB, K. *Multi-objective optimization using evolutionary algorithms*. [S.l.]: John Wiley & Sons, 2001.

- DEB, K. A Robust Evolutionary Framework for Multi-Objective Optimization. In: ACM. *Proceedings of the 10th annual conference on Genetic and evolutionary computation*. [S.l.], 2008. p. 633–640.
- DEB, K.; AGRAWAL, S. A niched-penalty approach for constraint handling in genetic algorithms. In: *Proceedings of the international conference on artificial neural networks and genetic algorithms (ICANNGA-99)*. [S.l.: s.n.], 1999. p. 235–243.
- DEB, K.; JAIN, H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints. *IEEE Trans. Evolutionary Computation*, v. 18, n. 4, p. 577–601, 2014.
- DEB, K.; KARTHIK, S. et al. Dynamic multi-objective optimization and decision-making using modified NSGA-II: a case study on hydro-thermal power scheduling. In: SPRINGER. *Evolutionary Multi-Criterion Optimization*. [S.l.], 2007. p. 803–817.
- DEB, K.; PRATAP, A.; AGARWAL, S.; MEYARIVAN, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, IEEE, v. 6, n. 2, p. 182–197, 2002.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, JSTOR, p. 1–38, 1977.
- DENYSIUK, R.; COSTA, L.; SANTO, I. E.; MATOS, J. C. MOEA/PC: Multiobjective evolutionary algorithm based on polar coordinates. In: SPRINGER. *International Conference on Evolutionary Multi-Criterion Optimization*. [S.l.], 2015. p. 141–155.
- DURBIN, R.; EDDY, S. R.; KROGH, A.; MITCHISON, G. *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. [S.l.]: Cambridge university press, 1998.
- EIBEN, A.; SMITH, J. *Introduction to evolutionary computing*. [S.l.]: Springer, 2003.
- FABIUS, J. et al. Asymptotic behavior of Bayes' estimates. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 35, n. 2, p. 846–856, 1964.
- FARIA, S. M. F. de S. *Modelos de mistura: aplicações em análise de regressão*. Phd Thesis (PhD Thesis) — Universidade do Porto, 2006.
- FARINA, M.; DEB, K.; AMATO, P. Dynamic Multiobjective Optimization Problems: Test Cases, Approximations, and Applications. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, p. 425–442, 2004.
- FERGUSON, T. S. A Bayesian analysis of some nonparametric problems. *The annals of statistics*, JSTOR, p. 209–230, 1973.
- FERGUSON, T. S. Prior distributions on spaces of probability measures. *The annals of statistics*, JSTOR, p. 615–629, 1974.
- FONSECA, C. M.; KNOWLES, J. D.; THIELE, L.; ZITZLER, E. A tutorial on the performance assessment of stochastic multiobjective optimizers. In: *Third International Conference on Evolutionary Multi-Criterion Optimization (EMO 2005)*. [S.l.: s.n.], 2005. v. 216, p. 240.

- FREEDMAN, D. A. On the asymptotic behavior of Bayes' estimates in the discrete case. *The Annals of Mathematical Statistics*, JSTOR, p. 1386–1403, 1963.
- FRIEDMAN, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association*, Taylor & Francis, v. 32, n. 200, p. 675–701, 1937.
- FU, X.; SUN, J. A new learning based dynamic multi-objective optimisation evolutionary algorithm. In: IEEE. *Evolutionary Computation (CEC), 2017 IEEE Congress on*. [S.l.], 2017. p. 341–348.
- GAMERMAN, D.; LOPES, H. F. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. [S.l.]: CRC Press, 2006.
- GARIS, H. de. A Dynamic Multi-Objective Evolutionary Algorithm Based on an Orthogonal Design. *2006 IEEE International Conference on Evolutionary Computation*, p. 573–580, 2006. Available at: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1688361>>.
- GARROZI, C.; ARAUJO, A. F. A memetic algorithm with one-step local search to guide diversity increase in Dynamic Multiobjective problems. In: IEEE. *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*. [S.l.], 2011. p. 649–654.
- GEE, S. B.; TAN, K. C.; ABBASS, H. A. A Benchmark Test Suite for Dynamic Evolutionary Multiobjective Optimization. *IEEE transactions on cybernetics*, IEEE, v. 47, n. 2, p. 461–472, 2017.
- GEE, S. B.; TAN, K. C.; ALIPPI, C. Solving multiobjective optimization problems in unknown dynamic environments: An inverse modeling approach. *IEEE transactions on cybernetics*, IEEE, v. 47, n. 12, p. 4223–4234, 2017.
- GELFAND, A. E.; SMITH, A. F. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, Taylor & Francis Group, v. 85, n. 410, p. 398–409, 1990.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *Journal of Applied Statistics*, Taylor & Francis, v. 20, n. 5-6, p. 25–62, 1984.
- GOH, C. K.; TAN, K. C. A competitive-cooperative coevolutionary paradigm for dynamic multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, v. 13, n. 1, p. 103–127, 2009. ISSN 1089778X.
- GREFENSTETTE, J. J. et al. Genetic algorithms for changing environments. In: *Parallel Problem Solving from Nature-PPSN*. [S.l.: s.n.], 1992. v. 2, p. 137–144.
- GREWAL, M. S.; ANDREWS, A. P. Applications of Kalman filtering in aerospace 1960 to the present [historical perspectives]. *IEEE Control Systems Magazine*, IEEE, v. 30, n. 3, p. 69–78, 2010.
- GUAN, S.-U.; CHEN, Q.; MO, W. Evolving dynamic multi-objective optimization problems with objective replacement. *Artificial Intelligence Review*, Springer, v. 23, n. 3, p. 267–293, 2005.

- GUERRERO-PÉÑA, E.; ARAÚJO, A. F. A Gaussian Mixture Model based local search for Differential Evolution Algorithm. In: IEEE. *Evolutionary Computation (CEC), 2017 IEEE Congress on*. [S.l.], 2017. p. 1885–1892.
- HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, Biometrika Trust, v. 57, n. 1, p. 97–109, 1970.
- HATZAKIS, I.; WALLACE, D. Dynamic multi-objective optimization with evolutionary algorithms: a forward-looking approach. In: ACM. *Proceedings of the 8th annual conference on Genetic and evolutionary computation*. [S.l.], 2006. p. 1201–1208.
- HE, Z.; YEN, G. G. Ranking many-objective evolutionary algorithms using performance metrics ensemble. In: IEEE. *Evolutionary Computation (CEC), 2013 IEEE Congress on*. [S.l.], 2013. p. 2480–2487.
- HE, Z.; YEN, G. G. Many-objective evolutionary algorithm: Objective space reduction and diversity improvement. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 20, n. 1, p. 145–160, 2016.
- HELBIG, M.; DEB, K.; ENGELBRECHT, A. Key challenges and future directions of dynamic multi-objective optimisation. In: IEEE. *Evolutionary Computation (CEC), 2016 IEEE Congress on*. [S.l.], 2016. p. 1256–1261.
- HELBIG, M.; ENGELBRECHT, A. *Benchmark functions for cec 2015 special session and competition on dynamic multi-objective optimization*. [S.l.], 2015.
- HELBIG, M.; ENGELBRECHT, A. P. Issues with performance measures for dynamic multi-objective optimisation. In: IEEE. *Computational Intelligence in Dynamic and Uncertain Environments (CIDUE), 2013 IEEE Symposium on*. [S.l.], 2013. p. 17–24.
- HELBIG, M.; ENGELBRECHT, A. P. Performance measures for dynamic multi-objective optimisation algorithms. *Information Sciences*, Elsevier, v. 250, p. 61–81, 2013.
- HELBIG, M.; ENGELBRECHT, A. P. Population-based metaheuristics for continuous boundary-constrained dynamic multi-objective optimisation problems. *Swarm and Evolutionary computation*, Elsevier, v. 14, p. 31–47, 2014.
- HOFFMAN, M. D.; BLEI, D. M.; WANG, C.; PAISLEY, J. Stochastic Variational Inference. *The Journal of Machine Learning Research*, JMLR. org, v. 14, n. 1, p. 1303–1347, 2013.
- HOU, Z.; JIN, S. *Model free adaptive control: theory and applications*. [S.l.]: CRC press, 2013.
- IGEL, C.; HANSEN, N.; ROTH, S. Covariance matrix adaptation for multi-objective optimization. *Evolutionary computation*, MIT Press, v. 15, n. 1, p. 1–28, 2007.
- ISHIBUCHI, H.; MASUDA, H.; NOJIMA, Y. A study on performance evaluation ability of a modified inverted generational distance indicator. In: ACM. *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*. [S.l.], 2015. p. 695–702.

- ISHIBUCHI, H.; TSUKAMOTO, N.; SAKANE, Y.; NOJIMA, Y. Hypervolume approximation using achievement scalarizing functions for evolutionary many-objective optimization. In: IEEE. *Evolutionary Computation, 2009. CEC'09. IEEE Congress on.* [S.l.], 2009. p. 530–537.
- JIANG, M.; HUANG, Z.; LIMING, Q.; HUANG, W. et al. Transfer Learning based Dynamic Multiobjective Optimization Algorithms. *IEEE Transactions on Evolutionary Computation*, IEEE, 2017.
- JIANG, M.; QIU, L.; HUANG, Z.; YEN, G. G. Dynamic Multi-objective Estimation of Distribution Algorithm Based on Domain Adaptation and Nonparametric Estimation. *Information Sciences*, Elsevier, 2018.
- JIANG, S.; YANG, S. A Framework of Scalable Dynamic Test Problems for Dynamic Multi-objective Optimization. In: IEEE. *2014 IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments (CIDUE)*. [S.l.], 2014. p. 32–39.
- JIANG, S.; YANG, S. A Steady-state and Generational Evolutionary Algorithm for Dynamic Multiobjective Optimization. *IEEE Transactions on Evolutionary Computation*, PP, n. 99, p. 1–1, 2016. ISSN 1089-778X.
- JIANG, S.; YANG, S. Evolutionary dynamic multiobjective optimization: Benchmarks and algorithm comparisons. *IEEE transactions on cybernetics*, IEEE, v. 47, n. 1, p. 198–211, 2017.
- JIANG, S.; YANG, S.; YAO, X.; TAN, K. C.; KAISER, M.; KRASNOGOR, N. Benchmark Problems for CEC2018 Competition on Dynamic Multiobjective Optimisation. *2018 IEEE Congress on Evolutionary Computation (CEC)*, IEEE, 2018.
- JIN, Y.; BRANKE, J. Evolutionary Optimization in Uncertain Environments – A Survey. *IEEE Transactions on Evolutionary Computation*, p. 303–317, 2005.
- JIN, Y.; SENDHOFF, B. Constructing dynamic optimization test problems using the multi-objective optimization concept. In: *Applications of Evolutionary Computing*. [S.l.]: Springer, 2004. p. 525–536.
- JORDAN, M. I. *Learning in graphical models*. [S.l.]: Springer Science & Business Media, 1999.
- JORDAN, M. I.; GHAHRAMANI, Z.; JAAKKOLA, T. S.; SAUL, L. K. An introduction to variational methods for graphical models. *Machine learning*, Springer, v. 37, n. 2, p. 183–233, 1999.
- KALMAN, R. E. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, American Society of Mechanical Engineers, v. 82, n. 1, p. 35–45, 1960.
- KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: IEEE. *Neural Networks, 1995. Proceedings., IEEE International Conference on.* [S.l.], 1995. v. 4, p. 1942–1948.

- KHALED, A.; TALUKDER, A.; KIRLEY, M. A pareto following variation operator for evolutionary dynamic multi-objective optimization. In: IEEE. *Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence)*. IEEE Congress on. [S.l.], 2008. p. 2270–2277.
- KINAS, P. G.; ANDRADE, H. A. *Introdução à análise bayesiana (com R)*. [S.l.]: Buqui Livros Digitais, 2014.
- KOLLER, D.; FRIEDMAN, N. *Probabilistic graphical models: principles and techniques*. [S.l.]: MIT press, 2009.
- KOO, W. T.; GOH, C. K.; TAN, K. C. A predictive gradient strategy for multiobjective evolutionary algorithms in a fast changing environment. *Memetic Computing*, Springer, v. 2, n. 2, p. 87–110, 2010.
- KUKKONEN, S.; LAMPINEN, J. GDE3: The third evolution step of generalized differential evolution. In: IEEE. *Evolutionary Computation, 2005. The 2005 IEEE Congress on*. [S.l.], 2005. v. 1, p. 443–450.
- KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. *The annals of mathematical statistics*, JSTOR, p. 79–86, 1951.
- LAVINE, M. Some aspects of Polya tree distributions for statistical modelling. *The annals of statistics*, JSTOR, p. 1222–1235, 1992.
- LAVINE, M. More aspects of Polya tree distributions for statistical modelling. *The Annals of Statistics*, JSTOR, p. 1161–1176, 1994.
- LI, H.; ZHANG, Q. Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II. *IEEE Transactions on evolutionary computation*, IEEE, v. 13, n. 2, p. 284–302, 2009.
- LI, Y.-L.; ZHAN, Z.-H.; GONG, Y.-J.; CHEN, W.-N.; ZHANG, J.; LI, Y. Differential evolution with an evolution path: A deep evolutionary algorithm. *IEEE T. Cybernetics*, v. 45, n. 9, p. 1798–1810, 2015. Available at: <<http://dblp.uni-trier.de/db/journals/tcyb/tcyb45.html#LiZGC0L15>>.
- LI, Z.; CHEN, H.; XIE, Z.; CHEN, C.; SALLAM, A. Dynamic multiobjective optimization algorithm based on average distance linear prediction model. *The Scientific World Journal*, Hindawi, v. 2014, 2014.
- LIANG, J.; ZHU, X.; YUE, C.; LI, Z.; QU, B. Performance Analysis on Knee Point Selection Methods for Multi-Objective Sparse Optimization Problems. In: IEEE. *2018 IEEE Congress on Evolutionary Computation (CEC)*. [S.l.], 2018. p. 1–8.
- LIEFOOGHE, A.; DERBEL, B. A correlation analysis of set quality indicator values in multiobjective optimization. In: ACM. *Proceedings of the Genetic and Evolutionary Computation Conference 2016*. [S.l.], 2016. p. 581–588.
- LIN, Q.; LI, J.; DU, Z.; CHEN, J.; MING, Z. A novel multi-objective particle swarm optimization with multiple search strategies. *European Journal of Operational Research*, Elsevier, v. 247, n. 3, p. 732–744, 2015.

- LIU, C.-a. New Dynamic Multiobjective Evolutionary Algorithm with Core Estimation of Distribution. *2010 International Conference on Electrical and Control Engineering*, v. 4, p. 1345–1348, 2010. Available at: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5630466>>.
- LIU, C. A.; WANG, Y. Dynamic multi-objective optimization evolutionary algorithm. *Proceedings - Third International Conference on Natural Computation, ICNC 2007*, v. 4, n. Icnc, p. 456–459, 2007.
- LIU, J.; WANG, Y.; WANG, X.; SUI, X.; GUO, S.; LIU, L. An alpha-Dominance Expansion Based Algorithm for Many-Objective Optimization. In: IEEE. *Computational Intelligence and Security (CIS), 2017 13th International Conference on*. [S.l.], 2017. p. 6–10.
- LIU, M.; LIU, Y. A dynamic evolutionary multi-objective optimization algorithm based on decomposition and adaptive diversity introduction. In: IEEE. *Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on*. [S.l.], 2016. p. 235–240.
- LIU, M.; ZENG, W. A fast evolutionary algorithm for dynamic bi-objective optimization problems. In: IEEE. *Computer Science & Education (ICCSE), 2012 7th International Conference on*. [S.l.], 2012. p. 130–134.
- LIU, M.; ZHENG, J.; WANG, J.; LIU, Y.; JIANG, L. An adaptive diversity introduction method for dynamic evolutionary multiobjective optimization. In: IEEE. *Evolutionary Computation (CEC), 2014 IEEE Congress on*. [S.l.], 2014. p. 3160–3167.
- LIU, R.; NIU, X.; FAN, J.; MU, C.; JIAO, L. An orthogonal predictive model-based dynamic multi-objective optimization algorithm. *Soft Computing*, Springer, v. 19, n. 11, p. 3083–3107, 2015.
- MARTÍNEZ, S. Z.; COELLO, C. A. C. A multi-objective particle swarm optimizer based on decomposition. In: ACM. *Proceedings of the 13th annual conference on Genetic and evolutionary computation*. [S.l.], 2011. p. 69–76.
- MAULDIN, R. D.; SUDDERTH, W. D.; WILLIAMS, S. Polya trees and random distributions. *The Annals of Statistics*, JSTOR, p. 1203–1221, 1992.
- MCKAY, M. D.; BECKMAN, R. J.; CONOVER, W. J. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, Taylor & Francis, v. 21, n. 2, p. 239–245, 1979.
- MCKIGHT, P. E.; NAJAB, J. Kruskal-Wallis Test. *Corsini Encyclopedia of Psychology*, Wiley Online Library, 2010.
- MCKNIGHT, P. E.; NAJAB, J. Mann-Whitney U Test. *Corsini Encyclopedia of Psychology*, Wiley Online Library, 2010.
- MEHNEN, J.; WAGNER, T.; RUDOLPH, G. *Evolutionary optimization of dynamic multiobjective functions*. [S.l.]: Citeseer, 2006.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. *The journal of chemical physics*, AIP Publishing, v. 21, n. 6, p. 1087–1092, 1953.

-
- MIETTINEN, K. *Nonlinear multiobjective optimization*. [S.l.]: Springer Science & Business Media, 2012.
- MINKA, T. P. *A family of algorithms for approximate Bayesian inference*. Phd Thesis (PhD Thesis) — Massachusetts Institute of Technology, 2001.
- MITCHELL, M. *An Introduction to Genetic Algorithms*. [S.l.]: The MIT Press, 1998.
- MOHAMED, S.; GHAHRAMANI, Z.; HELLER, K. A. Bayesian exponential family PCA. In: *Advances in Neural Information Processing Systems*. [S.l.: s.n.], 2009. p. 1089–1096.
- MURPHY, K. P. *Machine learning: a probabilistic perspective*. [S.l.]: MIT press, 2012.
- MURUGANANTHAM, A.; TAN, K. C.; VADAKKEPAT, P. Evolutionary dynamic multiobjective optimization via Kalman filter prediction. *IEEE transactions on cybernetics*, IEEE, v. 46, n. 12, p. 2862–2873, 2016.
- MURUGANANTHAM, A.; ZHAO, Y.; GEE, S. B.; QIU, X.; TAN, K. C. Dynamic multiobjective optimization using evolutionary algorithm with Kalman filter. *Procedia Computer Science*, Elsevier, v. 24, p. 66–75, 2013.
- NAYLOR, J. C.; SMITH, A. F. Applications of a method for the efficient computation of posterior distributions. *Applied Statistics*, JSTOR, p. 214–225, 1982.
- NEMENYI, P. Distribution-free multiple comparisons. In: INTERNATIONAL BIOMETRIC SOC 1441 I ST, NW, SUITE 700, WASHINGTON, DC 20005-2210. *Biometrics*. [S.l.], 1962. p. 263.
- NEWMAN, D.; ASUNCION, A.; SMYTH, P.; WELLING, M. Distributed algorithms for topic models. *The Journal of Machine Learning Research*, JMLR. org, v. 10, p. 1801–1828, 2009.
- NOCEDAL, J.; WRIGHT, S. J. *Sequential quadratic programming*. [S.l.]: Springer, 2006.
- ORBANZ, P.; TEH, Y. W. Bayesian nonparametric models. In: *Encyclopedia of Machine Learning*. [S.l.]: Springer, 2010. p. 81–89.
- PEARL, J. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. [S.l.]: Morgan Kaufmann, 1988.
- PENG, Z.; ZHENG, J.; ZOU, J. A Population Diversity Maintaining Strategy Based on Dynamic Environment Evolutionary Model for Dynamic Multiobjective Optimization. *2014 IEEE Congress on Evolutionary Computation (CEC'2014)*, p. 274–281, 2014.
- PENG, Z.; ZHENG, J.; ZOU, J.; LIU, M. Novel prediction and memory strategies for dynamic multiobjective optimization. *Soft Computing*, Springer, v. 19, n. 9, p. 2633–2653, 2015.
- POLETO, F. Z. *Análise de dados categorizados com omissao em variáveis explicativas e respostas*. Phd Thesis (PhD Thesis) — Universidade de Sao Paulo, 2011.
- PURSHOUSE, R. C.; FLEMING, P. J. On the evolutionary optimization of many conflicting objectives. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 11, n. 6, p. 770–784, 2007.

- RABINER, L. R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, IEEE, v. 77, n. 2, p. 257–286, 1989.
- RESENDE, C. M. C. *Inferência Bayesiana Aproximada em Modelos de Espaço de Estados*. Phd Thesis (PhD Thesis) — Dissertação de mestrado (Mestrado em Estatística). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.
- RINCÓN, L. Introducción a los procesos estocásticos. *Ciudad de México: Universidad Nacional Autónoma de México*, 2012.
- ROBBINS, H.; MONRO, S. A stochastic approximation method. *The annals of mathematical statistics*, JSTOR, p. 400–407, 1951.
- ROBERT, C.; CASELLA, G. Monte Carlo Statistical Methods Springer-Verlag. *New York*, 2004.
- ROWEIS, S. EM algorithms for PCA and SPCA. *Advances in neural information processing systems*, MORGAN KAUFMANN PUBLISHERS, p. 626–632, 1998.
- ROY, R.; MEHNEN, J. Dynamic multi-objective optimisation for machining gradient materials. *CIRP Annals-Manufacturing Technology*, Elsevier, v. 57, n. 1, p. 429–432, 2008.
- RUAN, G.; YU, G.; ZHENG, J.; ZOU, J.; YANG, S. The effect of diversity maintenance on prediction in dynamic multi-objective optimization. *Applied Soft Computing*, Elsevier, v. 58, p. 631–647, 2017.
- RUE, H.; MARTINO, S.; CHOPIN, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, Wiley Online Library, v. 71, n. 2, p. 319–392, 2009.
- SAHMOUD, S.; TOPCUOGLU, H. R. A Memory-Based NSGA-II Algorithm for Dynamic Multi-objective Optimization Problems. In: *Applications of Evolutionary Computation*. [S.l.]: Springer, 2016. p. 296–310.
- SAHMOUD, S.; TOPCUOGLU, H. R. Sensor-based change detection schemes for dynamic multi-objective optimization problems. In: IEEE. *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on*. [S.l.], 2016. p. 1–8.
- SCHOTT, J. R. *Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization*. [S.l.], 1995.
- SCHÜTZE, O.; ESQUIVEL, X.; LARA, A.; COELLO, C. A. C. Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization. *Evolutionary Computation, IEEE Transactions on*, IEEE, v. 16, n. 4, p. 504–522, 2012.
- SMITH, A.; SKENE, A.; SHAW, J.; NAYLOR, J. Progress with numerical and graphical methods for practical Bayesian statistics. *The Statistician*, JSTOR, p. 75–82, 1987.
- SMOLA, A.; NARAYANAMURTHY, S. An architecture for parallel topic models. *Proceedings of the VLDB Endowment*, VLDB Endowment, v. 3, n. 1-2, p. 703–710, 2010.

- SOUSA, J. T. G. de; FRANCA, J. E. M. da; ARAÚJO, A. F. R. Differential Evolution-Based Parameter Tuning in Model-Free Adaptive Control. In: IEEE. *2015 IEEE International Conference on Systems, Man, and Cybernetics*. [S.l.], 2015. p. 1726–1731.
- STATISTICS, E. S. E.-U. Eurostat Statistics Explained. *Eurostat Statistics Explained-Unemployment Statistics*, 2010.
- STORN, R.; PRICE, K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, Springer, v. 11, n. 4, p. 341–359, 1997.
- TANTAR, E.; TANTAR, A.-A.; BOUVRY, P. On dynamic multi-objective optimization, classification and performance measures. In: IEEE. *Evolutionary Computation (CEC), 2011 IEEE Congress on*. [S.l.], 2011. p. 2759–2766.
- TEH, Y. W.; JORDAN, M. I. Hierarchical Bayesian nonparametric models with applications. *Bayesian nonparametrics*, Camb. Ser. Stat. Probab. Math, v. 1, 2010.
- TEH, Y. W.; JORDAN, M. I.; BEAL, M. J.; BLEI, D. M. Hierarchical dirichlet processes. *Journal of the american statistical association*, v. 101, n. 476, 2006.
- TIAN, Y.; CHENG, R.; ZHANG, X.; JIN, Y. PlatEMO: a matlab platform for evolutionary multi-objective optimization. *arXiv preprint arXiv:1701.00879*, 2017.
- TIERNEY, L.; KADANE, J. B. Accurate approximations for posterior moments and marginal densities. *Journal of the american statistical association*, Taylor & Francis Group, v. 81, n. 393, p. 82–86, 1986.
- TIERNEY, L.; KASS, R. E.; KADANE, J. B. Approximate marginal densities of nonlinear functions. *Biometrika*, Biometrika Trust, v. 76, n. 3, p. 425–433, 1989.
- TIPPING, M. E.; BISHOP, C. M. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 61, n. 3, p. 611–622, 1999.
- TRIVEDI, A.; SRINIVASAN, D.; SANYAL, K.; GHOSH, A. A survey of multiobjective evolutionary algorithms based on decomposition. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 21, n. 3, p. 440–462, 2017.
- UÇAK, K.; GÜNEL, G. Ö. Model free adaptive support vector regressor controller for nonlinear systems. *Engineering Applications of Artificial Intelligence*, Elsevier, v. 81, p. 47–67, 2019.
- VACHHANI, V. L.; DABHI, V. K.; PRAJAPATI, H. B. Survey of multi objective evolutionary algorithms. In: IEEE. *Circuit, Power and Computing Technologies (ICCPCT), 2015 International Conference on*. [S.l.], 2015. p. 1–9.
- VELDHUIZEN, D. A. V. *Multiobjective evolutionary algorithms: classifications, analyses, and new innovations*. [S.l.], 1999.
- VELDHUIZEN, V.; A., D.; LAMONT, G. B. *Multiobjective Evolutionary Algorithm Research: A History and Analysis*. [S.l.]: Graduate School of Engineering, Air Force Institute of Technology, Wright-Patterson {AFB}, 1998.

WAINWRIGHT, M. J.; JORDAN, M. I. Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning*, Now Publishers Inc., v. 1, n. 1-2, p. 1–305, 2008.

WAN, S.; WANG, D. A Novel Differential Evolution for Dynamic Multiobjective Optimization with Adaptive Immigration Scheme. In: IEEE. *Proceedings of 2013 3rd International Conference on Computer Science and Network Technology*. [S.l.], 2013. p. 502–507.

WANG, Y.; DANG, C. An evolutionary algorithm for dynamic multi-objective optimization. *Applied Mathematics and Computation*, Elsevier Inc., v. 205, n. 1, p. 6–18, 2008. ISSN 00963003. Available at: <<http://dx.doi.org/10.1016/j.amc.2008.05.151>>.

WANG, Y.; LI, B. Investigation of memory-based multi-objective optimization evolutionary algorithm in dynamic environment. In: IEEE. *Evolutionary Computation, 2009. CEC'09. IEEE Congress on*. [S.l.], 2009. p. 630–637.

WANG, Y.; LI, B. Multi-strategy ensemble evolutionary algorithm for dynamic multi-objective optimization. *Memetic Computing*, v. 2, n. 1, p. 3–24, 2010. ISSN 18659284.

WELLING, M.; TEH, Y. W. Bayesian learning via stochastic gradient Langevin dynamics. In: *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*. [S.l.: s.n.], 2011. p. 681–688.

WHILE, L.; HINGSTON, P.; BARONE, L.; HUBAND, S. A faster algorithm for calculating hypervolume. *IEEE transactions on evolutionary computation*, IEEE, v. 10, n. 1, p. 29–38, 2006.

WOOLDRIDGE, J. M. *Introductory econometrics: A modern approach*. [S.l.]: Nelson Education, 2015.

WU, Y.; JIN, Y.; LIU, X. A directed search strategy for evolutionary dynamic multiobjective optimization. *Soft Computing*, Springer, v. 19, n. 11, p. 3221–3235, 2015.

YANG, S.; LI, M.; LIU, X.; ZHENG, J. A grid-based evolutionary algorithm for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 17, n. 5, p. 721–736, 2013.

YOUNES, A. *Adapting Evolutionary Approaches for Optimization in Dynamic Environments*. [S.l.]: University of Waterloo: Waterloo, 2006. 288 p.

YUPING, W.; HAILIN, L.; LEUNG, Y.-W. An effective uniform genetic algorithm for hard optimization problems. In: IEEE. *Intelligent Control and Automation, 2000. Proceedings of the 3rd World Congress on*. [S.l.], 2000. v. 1, p. 656–660.

ZHANG, J.; ZHOU, A.; ZHANG, G. A classification and pareto domination based multiobjective evolutionary algorithm. In: IEEE. *Evolutionary Computation (CEC), 2015 IEEE Congress on*. [S.l.], 2015. p. 2883–2890.

ZHANG, Q.; LI, H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on evolutionary computation*, IEEE, v. 11, n. 6, p. 712–731, 2007.

- ZHANG, Q.; ZHOU, A.; JIN, Y. Modelling the regularity in an estimation of distribution algorithm for continuous multiobjective optimisation with variable linkages. *Dept. Comput. Sci., Univ. Essex, Colchester, UK, Tech. Rep. CSM-459*, 2006.
- ZHANG, Q.; ZHOU, A.; JIN, Y. RM-MEDA: A regularity model-based multiobjective estimation of distribution algorithm. *IEEE Transactions on Evolutionary Computation*, IEEE, v. 12, n. 1, p. 41–63, 2008.
- ZHANG, X.; ZHENG, X.; CHENG, R.; QIU, J.; JIN, Y. A competitive mechanism based multi-objective particle swarm optimizer with fast convergence. *Information Sciences*, Elsevier, v. 427, p. 63–76, 2018.
- ZHANG, Z. Multiobjective optimization immune algorithm in dynamic environments and its application to greenhouse control. *Applied Soft Computing Journal*, v. 8, n. 2, p. 959–971, 2008. ISSN 15684946.
- ZHANG, Z.; QIAN, S. Artificial immune system in dynamic environments solving time-varying non-linear constrained multi-objective problems. *Soft Computing*, v. 15, n. 7, p. 1333–1349, 2011. ISSN 14327643.
- ZHAO, W.; WANG, L.; YIN, Y.; WANG, B.; TANG, Y. Sequential quadratic programming enhanced backtracking search algorithm. *Frontiers of Computer Science*, Springer, v. 12, n. 2, p. 316–330, 2018.
- ZHENG, B. A New Dynamic Multi-objective Optimization Evolutionary Algorithm. In: IEEE. *Third international conference on natural computation (ICNC 2007)*. [S.l.], 2007. v. 5, p. 565–570.
- ZHOU, A.; JIN, Y.; ZHANG, Q. A Population prediction strategy for evolutionary dynamic multiobjective optimization. *IEEE Transactions on Cybernetics*, v. 44, n. 1, p. 40–53, 2014. ISSN 21682267.
- ZHOU, A.; JIN, Y.; ZHANG, Q.; SENDHOFF, B.; TSANG, E. Prediction-based Population Re-initialization for Evolutionary Dynamic Multi-objective Optimization. *Emo*, p. 1–14, 2007. ISSN 03029743.
- ZITZLER, E. *Evolutionary algorithms for multiobjective optimization: Methods and applications*. [S.l.]: Citeseer, 1999.
- ZITZLER, E.; BROCKHOFF, D.; THIELE, L. The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration. In: SPRINGER. *Evolutionary multi-criterion optimization*. [S.l.], 2007. p. 862–876.
- ZITZLER, E.; LAUMANN, M.; THIELE, L. SPEA2: Improving the strength Pareto evolutionary algorithm. *TIK-report*, Eidgenössische Technische Hochschule Zürich (ETH), Institut für Technische Informatik und Kommunikationsnetze (TIK), v. 103, 2001.
- ZITZLER, E.; THIELE, L. Multiobjective optimization using evolutionary algorithms—a comparative case study. In: SPRINGER. *Parallel problem solving from nature—PPSN V*. [S.l.], 1998. p. 292–301.

ZOU, J.; LI, Q.; YANG, S.; BAI, H.; ZHENG, J. A prediction strategy based on center points and knee points for evolutionary dynamic multi-objective optimization. *Applied Soft Computing*, Elsevier, v. 61, p. 806–818, 2017.