Electric Vehicle Routing as a Multi-Objective Evolutionary Problem

Rafael do Nascimento Moura[1], **Rodrigo Gabriel Ferreira Soares**[1]

[1]Centro de Informática – Universidade Federal de Pernambuco, Recife, Brasil

rnm4@cin.ufpe.br, rgfs@cin.ufpe.br

Abstract. *The transition to electric logistics fleets introduces complex operational challenges due to limited vehicle autonomy and recharging needs. The Electric Vehicle Routing Problem with Time Windows (EVRPTW) addresses this reality, requiring the balance of conflicting criteria: fleet size, total distance, and delivery delays. This paper contributes a systematic and statistically rigorous comparison of three Multi-Objective Evolutionary Algorithm paradigms, NSGA-II, SMS-EMOA, and MOEA/D, applied to the tri-objective EVRPTW. Experiments conducted on 50 instances statistically prove that SMS-EMOA produces Pareto fronts of superior overall quality. Additionally, results demonstrate that NSGA-II is the most effective in minimizing extreme delays, whereas MOEA/D suffers a severe structural inadequacy caused by the problem's discrete and stratified nature.*

Resumo. *A transição para frotas logísticas elétricas introduz complexos desafios operacionais devido à autonomia limitada dos veículos e necessidade de recargas. O Problema de Roteamento de Veículos Elétricos com Janelas de Tempo (EVRPTW) reflete essa realidade, exigindo o balanceamento de múltiplos critérios conflitantes: tamanho da frota, distância e atrasos. Este artigo contribui com uma comparação sistemática e estatisticamente rigorosa de três paradigmas de Algoritmos Evolutivos Multiobjetivo, NSGA-II, SMS-EMOA e MOEA/D, aplicados ao EVRPTW. Experimentos conduzidos em 50 instâncias comprovam que o SMS-EMOA produz frentes de Pareto de qualidade global superior. Adicionalmente, os resultados demonstram que o NSGA-II é o mais eficaz em atrasos extremos, enquanto o MOEA/D apresenta inadequação estrutural severa devido à natureza discreta e estratificada do problema.*

1. Introdução

Um dos desafios emergentes da eletrificação veicular é o Problema de Roteamento de Veículos Elétricos com Janelas de Tempo (EVRPTW)[Schneider et al. 2014], no qual uma frota homogênea de veículos elétricos deve atender clientes com janelas de tempo e demandas conhecidas, respeitando restrições de capacidade de carga, autonomia de bateria e necessidade de recarga. Esta otimização envolve múltiplos critérios conflitantes: 1) o tamanho da frota; 2) a distância percorrida e; 3) o atraso nas entregas. Assim, caracterizamos o EVRPTW como um problema de otimização de múltiplos objetivos.

Devido à natureza NP-difícil do problema [Lenstra and Rinnooy Kan 1981] e à complexidade combinatória das restrições elétricas, métodos exatos tornam-se proibitivos em aplicações reais, e a literatura aponta as metaheurísticas como as abordagens

mais efetivas nesses cenários [Qin et al. 2021, Ye et al. 2022, Zhou et al. 2025]. Embora diversas metaheurísticas tratem com sucesso as restrições operacionais do EVRPTW [Erdelić and Carić 2022, Mao et al. 2020], formular o roteamento como critérios conflitantes exige uma abordagem específica. Nesse ponto, os Algoritmos Evolutivos Multiobjetivo (MOEAs) destacam-se por não agregar os objetivos, mas aproximar todo o conjunto de *trade-offs* ótimos (a frente de Pareto) em uma única execução, oferecendo ao decisor um panorama completo de alternativas. Três famílias representam as principais abordagens da literatura — dominância de Pareto (NSGA-II, [Deb et al. 2002]), indicador de qualidade (SMS-EMOA, [Beume et al. 2007]) e decomposição escalar (MOEA/D, [Zhang and Li 2007]) — empregando mecanismos distintos para equilibrar convergência e diversidade, com impactos potencialmente sensíveis à estrutura do problema. Embora clássicos, a escolha específica destes algoritmos justifica-se pois eles permanecem amplamente adotados na literatura contemporânea do EVRPTW, servindo consistentemente como *baselines* de desempenho e como arquiteturas-base para modificações.

A literatura de EVRPTW recente prioriza duas direções: hibridizar MOEAs com módulos de intensificação especializados (busca de grande vizinhança, componentes meméticos, aprendizado por reforço) [Azadi et al. 2024, Zheng et al. 2025, Haddad et al. 2025] e estender a formulação com detalhes operacionais mais realistas, como frotas heterogêneas com recarga parcial [Azadi et al. 2024], coleta e entrega simultâneas [Zheng et al. 2025] e troca rápida de bateria [Haddad et al. 2025]. Ao modificar problema e algoritmo simultaneamente, porém, essas abordagens ofuscam o efeito base do mecanismo evolutivo. Falta, assim, um entendimento sistemático de como os diferentes paradigmas de seleção multiobjetivo lidam com a complexidade estrutural inerente ao EVRPTW. Este trabalho responde três Questões de Pesquisa (QP):

QP1. *Qual MOEA produz as melhores soluções para o EVRPTW com respeito ao número de veículos, à distância e ao atraso? Por quê?* Avaliamos as frentes de Pareto em 50 instâncias do problema e apontamos o melhor algoritmo para configuração do problema.

QP2. *Qual o impacto das características do problema na convergência de cada algoritmo? E qual algoritmo consegue lidar melhor com estas restrições?* Avaliamos as curvas geométricas de convergência ao longo de 850.000 avaliações considerando instâncias representativas que variam a topologia espacial e a rigidez das janelas de tempo.

QP3. *Qual algoritmo lida melhor com a otimização de cada objetivo individualmente?* Realizamos experimentos extensivos para identificar o cenário em que cada algoritmo tem melhor desempenho para cada objetivo.

Nossas contribuições são: 1) a primeira comparação sistemática e estatisticamente rigorosa de NSGA-II, SMS-EMOA e MOEA/D em 50 instâncias do EVRPTW [Schneider et al. 2014] com 30 repetições, empregando a metodologia de [Demšar 2006, García and Herrera 2008]; 2) a identificação e análise da inadequação estrutural do MOEA/D para o EVRPTW; e 3) o primeiro estudo sobre as interações entre operadores genéticos de permutação e o decodificador de particionamento de rotas em EVRPTW.

2. Formulação do Problema

O EVRPTW é definido sobre um grafo completo $G = (V, A)$ onde o conjunto de nós $V = V_0 \cup \mathcal{C} \cup \mathcal{S}$ é composto pelo depósito $V_0 = \{0, n+1\}$ (com cópias para partida e

retorno), pelo conjunto de clientes $\mathcal{C} = \{1, \ldots, n\}$ e pelo conjunto de estações de recarga $\mathcal{S}$. O conjunto de arcos é $A = \{(i, j) : i, j \in V,\ i \neq j\}$, com distância d_{ij} e tempo de viagem $t_{ij} = d_{ij}/v$ associados a cada arco, onde v é a velocidade constante do veículo.

Cada cliente $i \in \mathcal{C}$ possui demanda $q_i > 0$ (representando a carga que será recolhida pelo caminhão no cliente), janela de tempo $[e_i, l_i]$ (onde e_i representa o instante mais cedo permitido para o início do atendimento e l_i o instante limite máximo) e tempo de serviço s_i. Cada veículo possui capacidade de carga Q, capacidade de bateria C, taxa de consumo energético r (que decresce o nível da bateria de forma linear e proporcional à distância percorrida) e taxa de recarga g (que determina o tempo necessário para o carregamento total dos caminhões). A política de recarga adotada é a *full recharge* (Charge-Depart-Policy — CDP), na qual o veículo recarrega completamente a bateria em cada estação visitada [Schneider et al. 2014].

2.1. Formulação Tri-Objetivo

Sejam K o conjunto de veículos disponíveis e $x_{ijk} \in \{0, 1\}$ a variável de decisão que indica se o veículo k percorre o arco (i, j). Definem-se ainda: τ_i como o instante de chegada do veículo ao nó i, u_i como a carga acumulada ao chegar ao nó i, e y_i como o nível de bateria ao chegar ao nó i. A formulação tri-objetivo minimiza:

$$f_1 = \sum_{k \in K} \sum_{j \in V \setminus \{0\}} x_{0jk}, \quad f_2 = \sum_{k \in K} \sum_{(i,j) \in A} d_{ij}\, x_{ijk}, \quad f_3 = \sum_{i \in \mathcal{C}} \max(0,\ \tau_i - l_i), \tag{1}$$

sujeito às restrições:

Atendimento e fluxo. Cada cliente é visitado exatamente uma vez, e o fluxo de cada veículo é conservado, ou seja,

$$\sum_{k \in K} \sum_{j \in V} x_{ijk} = 1, \qquad \forall\, i \in \mathcal{C} \tag{2}$$

$$\sum_{j \in V} x_{ijk} = \sum_{j \in V} x_{jik}, \qquad \forall\, i \in V,\ \forall\, k \in K \tag{3}$$

Capacidade. A carga acumulada em qualquer ponto da rota não excede a capacidade do veículo:

$$u_j \geq u_i + q_j - Q\,(1 - x_{ijk}), \qquad \forall\, (i, j) \in A,\ \forall\, k \in K \tag{4}$$

$$0 \leq u_i \leq Q, \qquad \forall\, i \in V \tag{5}$$

Tempo. O instante de chegada respeita o tempo de viagem e serviço. Chegadas antecipadas implicam espera até e_i; chegadas tardias contribuem para f_3:

$$\tau_j \geq \tau_i + s_i + t_{ij} - M\,(1 - x_{ijk}), \qquad \forall\, (i, j) \in A,\ \forall\, k \in K \tag{6}$$

Bateria. O nível de bateria ao chegar em um vértice j é limitado pelo nível do vértice anterior i subtraído do consumo no arco (i, j). As restrições ativam esse decréscimo apenas se o arco for percorrido (através da constante suficientemente grande M), garantindo a restauração total (C) caso o nó seja uma estação de recarga ou o depósito:

$$y_j \leq y_i - r \cdot d_{ij} + M\,(1 - x_{ijk}), \qquad \forall\, (i, j) \in A,\ j \notin \mathcal{S},\ \forall\, k \tag{7}$$

$$y_j = C, \qquad \forall\, j \in \mathcal{S} \cup V_0 \tag{8}$$

$$0 \leq y_i \leq C, \qquad \forall\, i \in V \tag{9}$$

Origem e destino. Cada veículo parte do depósito e retorna a ele:

$$\sum_{j \in V} x_{v0jk} \leq 1, \qquad \forall\, k \in K \tag{10}$$

$$\sum_{i \in V} x_{i,n+1,k} = \sum_{j \in V} x_{v0jk}, \qquad \forall\, k \in K \tag{11}$$

Esta formulação tri-objetivo relaxa as janelas de tempo, transformando-as em objetivo (f_3). Esta abordagem é adotada em problemas multiobjetivo com restrições (CMOPs), nos quais a transformação de restrições em objetivos é uma técnica estabelecida: 1) para ampliar o espaço de *trade-offs* [Vieira et al. 2004]; e 2) pela literatura de VRP multiobjetivo, onde o atraso é um objetivo relevante [Castro-Gutierrez et al. 2011].

3. Metodologia

Cada indivíduo é codificado como uma permutação $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ dos n clientes. As estações de recarga e o depósito não são codificados e são incorporados à solução pelo decodificador descrito abaixo. O benefício desta abordagem [Prins 2004] é a separação de responsabilidades. O algoritmo evolutivo explora a ordenação relativa dos clientes, enquanto o decodificador transforma tal ordenação em rotas viáveis, decidindo onde particionar a sequência (quantos veículos usar e quais clientes cada um atende) e onde inserir estações de recarga. Essa separação é particularmente vantajosa no EVRPTW, onde as múltiplas restrições interdependentes – capacidade, bateria, janelas de tempo, horizonte do depósito – tornam a geração direta de rotas viáveis um problema combinatório em si. O decodificador opera em três fases sequenciais: particionamento da permutação em rotas, inserção de estações de recarga e avaliação dos objetivos. A seguir, descrevemos as duas primeiras fases. A última consiste apenas no cálculo dos três objetivos.

3.1. Decodificador

A primeira fase recebe a permutação π e encontra a partição ótima em subsequências contíguas, cada uma correspondendo a uma rota, que minimiza a distância total. O problema é modelado como um caminho mais curto em um grafo auxiliar acíclico com $n+1$ nós, um para cada posição $0, 1, \ldots, n$ da permutação. Um arco (i, j) representa uma rota que serve os clientes $\pi_{i+1}, \ldots, \pi_j$ na ordem da permutação, com custo igual à distância total da rota: $d(0, \pi_{i+1}) + \sum_{k=i+1}^{j-1} d(\pi_k, \pi_{k+1}) + d(\pi_j, 0)$. Um arco é viável se duas condições de fechamento são satisfeitas: (a) a demanda acumulada $\sum_{k=i+1}^{j} q_{\pi_k}$ não excede a capacidade Q; e (b) o tempo estimado de retorno ao depósito não excede o horizonte. Crucialmente, violações de janelas de tempo dos clientes individuais não finalizam a rota, elas são permitidas nesta fase e capturadas como componentes de f_3 na avaliação final. Esta decisão permite a exploração de particionamentos que trocam pontualidade por menor distância ou menos veículos, enriquecendo o espaço de *trade-offs*. O caminho mais curto do nó 0 ao nó n é obtido por programação dinâmica em tempo $O(n^2)$, e a reconstrução via predecessores produz a partição ótima em rotas.

A segunda fase garante a viabilidade energética de cada rota, através do monitoramento do nível de bateria. Quando a bateria é insuficiente para o próximo trecho, ou quando o mecanismo de *look-ahead* preventivo detecta que ela não bastará para o próximo destino ou para a estação mais próxima, uma estação de recarga é inserida.

3.2. Algoritmos evolutivos

Os algoritmos compartilham os componentes de representação (permutação), decodificador, inicialização, crossover (OrderCrossover [Oliver et al. 1987]) e mutação (InversionMutation [Eiben and Smith 2015]). A população foi fixada em 105 indivíduos. Esse número segue o MOEA/D, que restringe o tamanho da população ao número de vetores de peso. Para três objetivos, o método de Das-Dennis gera 105 vetores. O NSGA-II e SMS-EMOA adotam o mesmo tamanho para garantir comparação justa.

A diferença fundamental entre os três mecanismos é que o NSGA-II avalia qualidade por *frente* de dominância e desempata por *crowding distance* (densidade local); o SMS-EMOA remove o indivíduo que menos contribui ao hipervolume total (Pareto-compatível); e o MOEA/D avalia cada solução apenas em relação ao subproblema escalar do vetor de peso correspondente, cooperando localmente com vizinhos.

3.3. Calibração de hiperparâmetros

Os hiperparâmetros foram ajustados via busca aleatória [Bergstra and Bengio 2012] com 100 iterações. Para garantir a capacidade de generalização dos algoritmos sem comprometer a viabilidade computacional, selecionamos uma instância de seis subgrupos. Para mitigar o viés humano nesta seleção, a instância eleita como "representativa" de cada grupo foi definida como aquela cujo *tw_ratio* (proporção de clientes com janelas de tempo restritivas) era o mais próximo da mediana de sua respectiva categoria.

Para selecioná-las, baseamo-nos nas características do *benchmark* de Schneider [Schneider et al. 2014], que divide os problemas em seis subgrupos definidos por duas dimensões: a distribuição espacial dos clientes — agrupada (C), aleatória (R) ou mista (RC) — e o regime de janelas de tempo — apertadas (série 1xx, englobando C1, R1 e RC1) ou largas (série 2xx, englobando C2, R2 e RC2).

As distribuições de probabilidade da busca aleatória foram: $p_c \sim$ Beta$(3;\ 1{,}5)$ em $[0, 1]$ e $p_m \sim$ Beta$(3;\ 1)$ em $[0, 1]$. Para o MOEA/D, usamos: $n_neighbors \sim U[3, 25]$, $prob_nm \sim$ Beta$(3;\ 1)$, decomposição $\sim$ Categórica uniforme $\{$Tchebycheff, Weighted Sum, PBI$\}$, e vetores de referência $\sim$ Categórica uniforme $\{$Das-Dennis, *energy*$\}$.

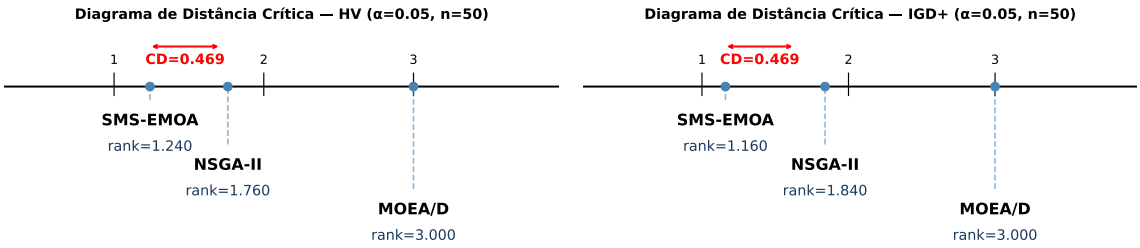
4. Experimentos

Para a avaliação final dos algoritmos, utilizamos as 50 instâncias restantes de grande porte [Schneider et al. 2014] (100 clientes, 21 estações), excluindo as 6 já utilizadas na calibração. Cada algoritmo foi avaliado em 30 execuções independentes em cada instância. O critério de parada de 850.000 avaliações foi calibrado considerando um ponto de estabilização com melhoria $< 2\%$, ao qual foi aplicada margem de segurança de 20%.

Foram utilizadas métricas que capturaram aspectos distintos e frequentemente discordantes. O Hipervolume (HV) [Zitzler and Thiele 1998] foi adotado como métrica principal por ser a única métrica de qualidade Pareto-compatível. O HV é sensível tanto à proximidade à frente verdadeira quanto à diversidade (espalhamento das soluções). Para cada instância, o ponto de referência é o máximo de cada objetivo entre todas as 90 execuções (3 algoritmos $\times$ 30 execuções), multiplicado por 1,1. A adoção desta margem garante que as soluções localizadas nos extremos da frente de Pareto possuam uma contribuição de volume estritamente positiva, evitando penalizações geométricas arbitrárias no cálculo da

Tabela 1. Ranks médios de Friedman ($n = 50$, $CD = 0{,}469$, $\alpha = 0{,}05$).

Métrica	SMS-EMOA	NSGA-II	MOEA/D	χ^2
HV (↑)	**1,24**	1,76	3,00	81,76
IGD+ (↓)	**1,16**	1,84	3,00	86,56



(a) Métrica HV **(b) Métrica IGD+**

Figura 1. Diagramas de Diferença Crítica ($n = 50$, $CD = 0{,}469$, $\alpha = 0{,}05$).

métrica [Ishibuchi et al. 2017]. O IGD+ [Ishibuchi et al. 2015] foi adotado como métrica complementar para avaliar especificamente a cobertura uniforme da frente. A frente de referência para cada instância é a união não-dominada de todas as soluções das 90 execuções.

Esta comparação tem duas etapas [Demšar 2006, García and Herrera 2008]. Na primeira, o teste de Friedman verifica se existe diferença global entre os algoritmos. Na segunda, o Wilcoxon *signed-rank* pareado com correção de Holm-Bonferroni identifica quais pares diferem significativamente, procedimento que [García and Herrera 2008] demonstram ser mais poderoso que o anterior por controlar o erro tipo I apenas sobre os pares testados em vez de sobre todas as comparações possíveis. Usamos também o teste de Friedman por subgrupo, correlação de Spearman entre *tw_ratio* e diferença de desempenho entre pares, curvas de convergência de HV ao longo das avaliações (mediana $\pm$ IQR), e contagem de vitórias por objetivo individual.

5. Qualidade Global e Efeito do Tipo de Instância

Nesta seção, analisamos empiricamente todas as instâncias, bem como as especificidades de cada subgrupo. A Tabela 1 apresenta os ranks médios de Friedman calculados sobre as 50 instâncias. O teste rejeitou a equivalência entre os algoritmos com p-value $< 0{,}001$ para ambas as métricas ($\chi^2 = 81{,}76$ para HV, $\chi^2 = 86{,}56$ para IGD+). Todas as diferenças pareadas excedem a distância crítica de Nemenyi ($CD = 0{,}469$), estabelecendo o ranking SMS-EMOA $>$ NSGA-II $>$ MOEA/D com significância estatística para todos os pares, conforme ilustrado nos diagramas de diferença crítica da Figura 1.

O SMS-EMOA obteve HV superior em 38 de 50 instâncias (76%). Sua vantagem é mais acentuada em IGD+ (rank 1,16) do que em HV (rank 1,24), indicando que a superioridade reside principalmente na cobertura uniforme da frente, considerando que o mecanismo de seleção por contribuição ao hipervolume elimina soluções redundantes e preserva aquelas que cobrem regiões da frente que ficariam descobertas. O NSGA-II, com rank 1,76 em HV, é o segundo melhor, significativamente inferior ao SMS-EMOA com diferença de 0,52 em rank médio, a menor entre os três pares. O MOEA/D obteve rank de 3,0 em ambas as métricas. Seu HV mediano correspondeu tipicamente de 60 a

Tabela 2. Proporção de instâncias com diferença significativa (Wilcoxon *signed-rank*, $\alpha = 0{,}05$, correção de Holm-Bonferroni.)

Par	HV	IGD+
SMS-EMOA vs. NSGA-II	4/50 (8%)	17/50 (34%)
NSGA-II vs. MOEA/D	48/50 (96%)	36/50 (72%)
SMS-EMOA vs. MOEA/D	50/50 (100%)	50/50 (100%)

Tabela 3. Ranks médios de Friedman por subgrupo para a métrica HV.

Subgrupo	Distribuição	n	SMS-EMOA	NSGA-II	MOEA/D
C1	Agrupada, Janela apertada	8	**1,12**	1,88	3,00
C2	Agrupada, Janela larga	7	**1,43**	1,57	3,00
R1	Aleatória, Janela apertada	11	**1,36**	1,64	3,00
R2	Aleatória, Janela larga	10	**1,10**	1,90	3,00
RC1	Mista, Janela apertada	7	**1,43**	1,57	3,00
RC2	Mista, Janela larga	7	**1,00**	2,00	3,00

85% do melhor entre os outros dois algoritmos. Esse resultado é notável e não pode ser atribuído a configuração inadequada, dado que a calibração explorou 100 configurações incluindo 3 funções de decomposição e 2 métodos de vetores de referência. A Tabela 2 detalha a análise por instância via *Wilcoxon signed-rank* pareado.

Para avaliarmos o ranking global em contextos específicos, o teste de Friedman foi repetido para cada um dos seis subgrupos que consideram a distribuição espacial e as janelas de tempo. A Tabela 3 apresenta os ranks médios por subgrupo. A vantagem do SMS-EMOA varia com a geometria do problema. Nas instâncias agrupadas com janelas apertadas, a diferença é a maior: rank 1,12 vs. 1,88 (diferença de 0,76). Nestas instâncias, os clientes formam agrupamentos geográficos com janelas de tempo restritivas, criando rotas compactas onde o conflito entre distância e atraso é mais severo e pequenas reduções na distância percorrida costumam resultar em grandes penalizações de tempo. A seleção por contribuição ao hipervolume do SMS-EMOA distribui soluções de forma mais eficaz nesse cenário. Nas instâncias com distribuição mista ou aleatória (RC1, R1), as diferenças são intermediárias (0,14 até 0,28), refletindo um espaço de busca com transições mais graduais e suaves entre os objetivos, que ambos os algoritmos baseados em dominância/indicador conseguem explorar de forma similar. As instâncias RC2 apresentam a maior diferença entre SMS-EMOA e NSGA-II (1,00 vs. 2,00). Janelas largas produzem soluções mais diversas em mais regiões viáveis da frente, cenário onde a seleção baseada em hipervolume é particularmente eficaz por atribuir valor explícito à cobertura de regiões vazias.

As correlações negativas e significativas entre *tw_ratio* e a vantagem sobre o MOEA/D ($\rho = -0{,}585$ para HV, $\rho = -0{,}764$ para IGD+) confirmam que janelas de tempo mais restritivas amplificam as limitações do MOEA/D. Este fato pode ser devido à estrutura da frente de Pareto. No EVRPTW, f_1 (número de veículos) é inteiro com baixa cardinalidade (tipicamente 3 a 20 valores distintos para instâncias de 100 clientes) produzindo uma frente estratificada composta por camadas discretas por valor de f_1, com *trade-off* aproximadamente linear entre f_2 e f_3 dentro de cada camada. O MOEA/D distribui 105 vetores de peso uniformemente no simplex normalizado, a região geométrica triangular contínua onde a soma dos pesos dos objetivos é igual a um, assumindo correspondência contínua com a frente. Na frente estratificada, múltiplos vetores convergem para os mesmos extremos de cada camada, desperdiçando subproblemas. Em experimen-

tos preliminares, apenas 9 das 105 soluções eram distintas (utilização de 9% dos vetores). Instâncias com janelas apertadas produzem menos soluções viáveis e menos camadas efetivas de f_1, exacerbando esse desperdício. Esse diagnóstico empírico é consistente com os resultados da literatura [Qi et al. 2015], que demonstraram como a natureza discreta da minimização de frotas faz com que múltiplos vetores de decomposição de Tchebycheff interceptem a frente de Pareto nas mesmas soluções duplicadas, esvaziando a capacidade de busca do MOEA/D. O MOEA/D também apresenta colapso parcial na manutenção de diversidade com mediana de 99 soluções na frente de Pareto (105 para os outros).

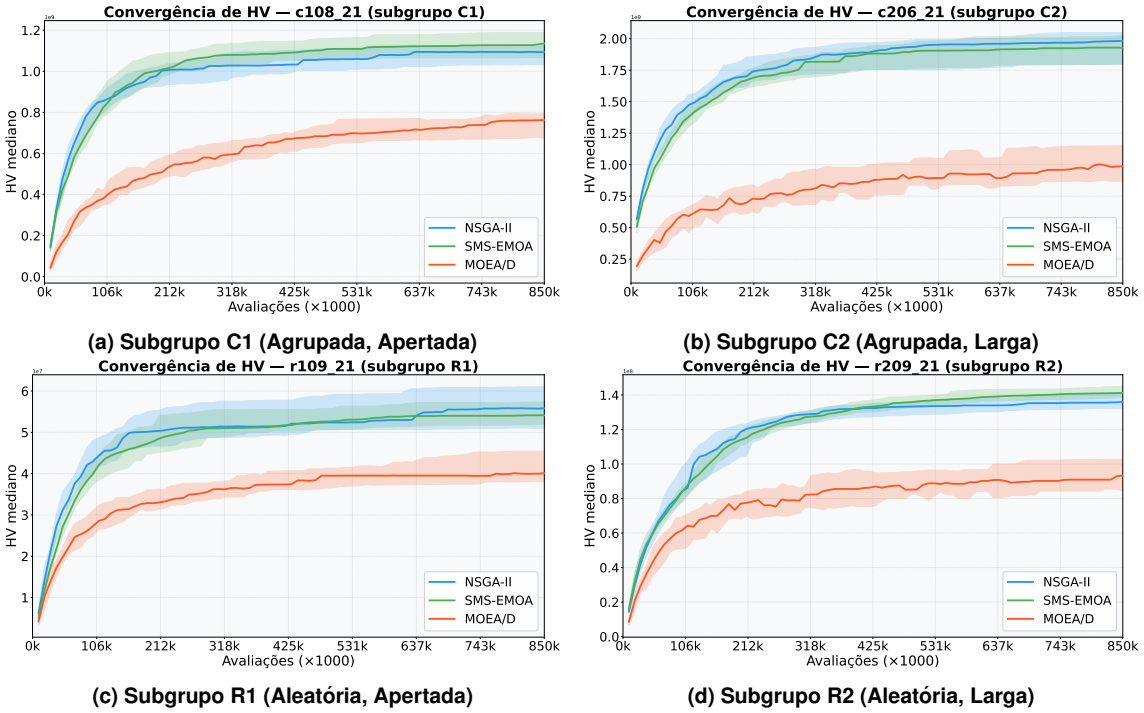
Entre SMS-EMOA e NSGA-II, a correlação com *tw_ratio* não é significativa para HV ($\rho = -0{,}065$, $p = 0{,}654$) mas significativa para IGD+ ($\rho = -0{,}321$, $p = 0{,}023$). Isto sugere que a vantagem do SMS-EMOA em *cobertura* se intensifica em instâncias com janelas apertadas, onde o espaço viável é menor e a distribuição uniforme de soluções é mais desafiadora. Neste caso, a seleção por contribuição ao hipervolume se mostra mais eficaz que o *crowding distance*.

Respondemos a QP1 demonstrando que SMS-EMOA produz frentes de melhor qualidade global, significativamente superior ao NSGA-II e ao MOEA/D. O ranking é consistente em todos os 6 subgrupos de instâncias. A inferioridade do MOEA/D é decorrente da incompatibilidade entre vetores de peso contínuos e frente estratificada por f_1 inteiro.

6. Dinâmica de Convergência

Esta seção investiga o comportamento dinâmico e a velocidade de convergência dos algoritmos. A Figura 2 apresenta as curvas de convergência de HV para 4 instâncias, pois RC1 e RC2 apresentaram comportamentos semelhantes a R1 e R2. A seleção destas instâncias mitigou o viés de análise, adotando os problemas com *tw_ratio* mais próximo à mediana de cada categoria. As curvas mostram a mediana e o intervalo interquartil (IQR) das 30 execuções independentes de cada algoritmo. Em instâncias com janelas apertadas, o NSGA-II converge mais rapidamente nas primeiras 100 a 200 mil avaliações, atingindo HV mediano superior ao SMS-EMOA nessa fase. A explicação está no mecanismo de seleção considerando que o *non-dominated sorting* com *crowding distance* opera sobre a população inteira a cada geração, produzindo pressão seletiva imediata e forte desde as primeiras iterações. O SMS-EMOA, com estratégia $(\mu+1)$, substitui apenas um indivíduo por geração, resultando em convergência inicial mais lenta.

A evolução posterior nessas instâncias, contudo, é separada pela topologia geográfica. Nas topologias aleatória e mista, o espaço de busca restrito mantém as curvas do SMS-EMOA e NSGA-II sobrepostas até o final, o que é consistente com a baixa taxa de significância por instância observada no teste de Wilcoxon. Por outro lado, na topologia agrupada, o isolamento entre *clusters* cria *trade-offs* mais acentuados. Nesse cenário, o SMS-EMOA é superior ao NSGA-II entre 300 e 500 mil avaliações e mantém esta vantagem até o final, explorando e refinando as regiões extremas de *trade-off* de maneira contínua. Nas instâncias com janelas largas, o SMS-EMOA apresenta desempenho superior desde as avaliações iniciais ou supera o NSGA-II precocemente (em torno de 50.000 avaliações). A diferença para o NSGA-II é mais visível do que nas instâncias com janelas apertadas, com pouca ou nenhuma sobreposição de bandas de IQR. Isto se deve às janelas largas combinadas com clientes espalhados expandirem o espaço de soluções viáveis de forma combinatória, criando frentes de Pareto com vastas regiões a serem cobertas.



(a) Subgrupo C1 (Agrupada, Apertada)

(b) Subgrupo C2 (Agrupada, Larga)

(c) Subgrupo R1 (Aleatória, Apertada)

(d) Subgrupo R2 (Aleatória, Larga)

Figura 2. Gráficos de convergência de HV (mediana $\pm$ IQR).

Neste cenário, a seleção por contribuição ao hipervolume do SMS-EMOA é particularmente eficaz por direcionar a busca para regiões vazias da frente desde o início, enquanto o *crowding distance* do NSGA-II reage apenas à densidade local já existente, levando a uma estagnação mais precoce.

Uma exceção ao padrão 2xx foi observada em c206_21, onde o NSGA-II ultrapassa o SMS-EMOA por volta de 600 mil avaliações e termina com HV mediano superior. Instâncias agrupadas com janelas largas produzem frentes com estrutura mais regular (camadas de f_1 uniformes), cenário onde o *crowding distance* pode ser tão eficaz quanto a contribuição ao hipervolume para manutenção de diversidade. Em todas as 6 instâncias representativas, o MOEA/D converge para um patamar de HV claramente inferior aos outros dois algoritmos, e a distância não diminui ao longo da execução. Isso confirma que a inferioridade do MOEA/D é estrutural e não resultado de convergência lenta que poderia ser resolvida com mais avaliações. Notavelmente, a curva do MOEA/D apresenta um padrão de "escada" com saltos discretos de HV, especialmente visível nas instâncias 2xx. Esses saltos correspondem à descoberta de novas camadas de f_1 na frente estratificada.

Respondemos a QP2 demonstrando que o NSGA-II converge mais rápido nas primeiras avaliações sob janelas apertadas, mas é similar ou inferior ao SMS-EMOA em diferentes topologias geográficas (C1). Em instâncias com janelas largas e dispersas (R2, RC2), o SMS-EMOA é superior desde o início da otimização. O NSGA-II se sobressai apenas no cenário agrupado de janelas largas (C2). O MOEA/D é inferior em todos os cenários, apresentando saltos discretos de convergência.

7. Desempenho em objetivos individuais

A Figura 3 apresenta a contagem de vitórias por objetivo individual. Para cada instância e cada objetivo, identifica-se qual algoritmo encontrou o melhor valor mediano (menor f_1,

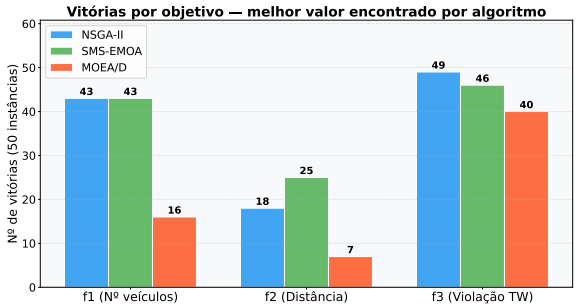


Figura 3. Vitórias por objetivo individual (melhor mediana por instância, $n = 50$).

menor f_2, menor f_3) sobre as 30 execuções.

Em f_1, NSGA-II e SMS-EMOA tiveram 43 vitórias cada, indicando que ambos são igualmente eficazes em minimizar o tamanho da frota. O MOEA/D obtém 16 vitórias. Em 30% das instâncias, ele encontra soluções com o mesmo número mínimo de veículos que os outros algoritmos. f_1 não é, portanto, o fator que diferencia os três algoritmos.

Em f_2, o SMS-EMOA teve 25 vitórias, contra 18 do NSGA-II e 7 do MOEA/D. A superioridade do SMS-EMOA em encontrar rotas mais curtas é consistente com seu mecanismo de seleção. A contribuição ao hipervolume favorece soluções que expandem a frente em regiões de alto volume, e no espaço tri-objetivo do EVRPTW, essas regiões frequentemente correspondem a combinações de distância baixa com atraso moderado.

Em f_3, o NSGA-II teve 49 vitórias em 50 instâncias. O SMS-EMOA obtém 46 e o MOEA/D 40. Em uma aplicação, um operador logístico que priorize a pontualidade nas entregas encontraria no NSGA-II as melhores soluções extremas para esse critério.

Com 40 vitórias (80% das instâncias) em f_3, o MOEA/D encontrou soluções com atraso tão baixo quanto os outros algoritmos. Combinado com suas 16 vitórias em f_1, isto demonstra que a inferioridade do MOEA/D demonstrada anteriormente não decorre de incapacidade de convergir para boas soluções pontuais. O MOEA/D é competitivo nos extremos da frente, sua limitação é a cobertura e distribuição de soluções entre esses extremos, resultando em frentes com lacunas e concentrações que penalizam severamente o HV e o IGD+. Isto é consequência do desperdício de vetores de peso na frente estratificada. Os subproblemas do MOEA/D convergem para extremos das camadas de f_1 (que são soluções boas individualmente), mas não conseguem manter soluções intermediárias.

Esta distinção entre qualidade pontual e qualidade global é relevante para diferentes contextos de decisão. Se o decisor precisa de uma única solução extrema (menor frota, menor distância ou menor atraso), as diferenças entre os algoritmos são modestas e todos são competitivos. Se o decisor precisa de um bom compromisso para uma análise comparativa, o SMS-EMOA é o mais adequado, seguido pelo NSGA-II.

Respondemos a QP3 demonstrando que os algoritmos possuem vantagens complementares em objetivos individuais: NSGA-II é superior para minimizar atraso (f_3), SMS-EMOA para distância (f_2), e ambos são equivalentes em número de veículos (f_1). O MOEA/D é competitivo em extremos individuais, mas limitado em cobertura.

8. Conclusão

Este trabalho avaliou os algoritmos NSGA-II, SMS-EMOA e MOEA/D no EVRPTW tri-objetivo. Os resultados revelam que a frente de Pareto estratificada — causada pela natureza discreta da frota (f_1) — penaliza severamente o MOEA/D devido ao desperdício de vetores de peso contínuos. Em contrapartida, o SMS-EMOA obteve a melhor qualidade global e distribuição de soluções, sendo ideal para mapear exaustivamente as alternativas operacionais e *trade-offs*. Já o NSGA-II desponta como uma opção pragmática por sua rápida convergência inicial e superioridade absoluta na minimização de atrasos (f_3). Conclui-se, portanto, que a escolha do método deve alinhar-se à geometria do problema e às prioridades logísticas. Como trabalhos futuros, propõe-se o desenvolvimento de vetores de peso adaptativos para o MOEA/D e a incorporação de modelos não lineares de recarga e degradação de bateria.

Referências

Azadi, A. H. S., Khalilzadeh, M., Antucheviciene, J., Heidari, A., and Soon, A. (2024). A sustainable multi-objective model for capacitated-electric-vehicle-routing-problem considering hard and soft time windows as well as partial recharging. *Biomimetics*, 9(4):242.

Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *J. Mach. Learn. Res.*, 13(10):281–305.

Beume, N., Naujoks, B., and Emmerich, M. (2007). Sms-emoa: Multiobjective selection based on dominated hypervolume. *Eur. J. Oper. Res.*

Castro-Gutierrez, J., Landa-Silva, D., and Moreno Pérez, J. (2011). Nature of real-world multi-objective vehicle routing with evolutionary algorithms. In *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, pages 257–264. IEEE.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Trans. Evol. Comput.*, 6(2):182–197.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.*, 7:1–30.

Eiben, A. E. and Smith, J. E. (2015). *Introduction to Evolutionary Computing*. Springer, 2 edition.

Erdelić, T. and Carić, T. (2022). Goods delivery with electric vehicles: Electric vehicle routing optimization with time windows and partial or full recharge. *Energies*, 15(1):285.

García, S. and Herrera, F. (2008). An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *J. Mach. Learn. Res.*, 9:2677–2694.

Haddad, A., Tlili, T., Dahmani, N., and Krichen, S. (2025). Integrating nsga-ii and q-learning for solving the multi-objective electric vehicle routing problem with battery swapping stations. *Int. J. Intell. Transp. Syst. Res.* Online first, doi:10.1007/s13177-025-00486-9.

Ishibuchi, H., Imada, R., Setoguchi, Y., and Nojima, Y. (2017). Reference point specification in hypervolume calculation for fair comparison and efficient search. In *Proc. Genet. Evol. Comput. Conf. (GECCO)*, pages 585–592. ACM.

Ishibuchi, H., Masuda, H., Tanigaki, Y., and Nojima, Y. (2015). Modified distance calculation in generational distance and inverted generational distance. In *Proc. Int. Conf. Evol. Multi-Criterion Optim. (EMO)*. Springer.

Lenstra, J. K. and Rinnooy Kan, A. H. G. (1981). Complexity of vehicle routing and scheduling problems. *Networks*, 11(2):221–227.

Mao, H., Shi, J., Zhou, Y., and Zhang, G. (2020). The electric vehicle routing problem with time windows and multiple recharging options. *IEEE Access*, 8:114864–114875.

Oliver, I. M., Smith, D. J., and Holland, J. R. C. (1987). A study of permutation crossover operators on the travelling salesman problem. In *Proc. 2nd Int. Conf. Genet. Algorithms*, pages 224–230. Lawrence Erlbaum.

Prins, C. (2004). A simple and effective evolutionary algorithm for the vehicle routing problem. *Comput. Oper. Res.*, 31(12):1985–2002.

Qi, Y., Hou, Z., Li, H., Huang, J., and Li, X. (2015). A decomposition based memetic algorithm for multi-objective vehicle routing problem with time windows. *Comput. Oper. Res.*, 62:61–77.

Qin, H., Su, X., Ren, T., and Luo, Z. (2021). A review on the electric vehicle routing problems: Variants and algorithms. *Front. Eng. Manag.*, 8(3):370–389.

Schneider, M., Stenger, A., and Goeke, D. (2014). The electric vehicle-routing problem with time windows and recharging stations. *Transp. Sci.*, 48(4):500–520.

Vieira, D. A. G., Adriano, R., de Vasconcelos, J. A., and Krähenbühl, L. (2004). Treating constraints as objectives in multiobjective optimization problems using niched pareto genetic algorithm. *IEEE Trans. Magn.*, 40(2):1188–1191.

Ye, C., He, W., and Chen, H. (2022). Electric vehicle routing models and solution algorithms in logistics distribution: A systematic review. *Environ. Sci. Pollut. Res.*, 29:57067–57090.

Zhang, Q. and Li, H. (2007). Moea/d: A multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.*, 11(6):712–731.

Zheng, Z., Liu, S., and Ong, Y.-S. (2025). Hybrid memetic search for electric vehicle routing with time windows, simultaneous pickup-delivery, and partial recharges. *arXiv preprint*.

Zhou, Y., Lei, Q., Li, L., and Wu, Z. (2025). An overview of energy replenishment strategies for the electric vehicle routing problem: Models and solution algorithms. *Energies*, 18(23):6196.

Zitzler, E. and Thiele, L. (1998). Multiobjective optimization using evolutionary algorithms: A comparative case study. In *Proc. Int. Conf. Parallel Probl. Solving from Nat. (PPSN V)*, volume 1498, pages 292–301. Springer.