

Transactions on
Computational Modeling and Intelligent Systems

Volume 5, October 2026, Article 10115
https://doi.org/10.65112/tcmis.10115

eISSN: 3108-5202
www.tcmis.org

RESEARCH ARTICLE

Exact and evolutionary three-objective configuration of battery electric buses: Cost, energy reserve and equipment value

Ali Yiğit Sabaner*, Ece Nur Polat and Ramazan Yaman

Department of Industrial Engineering, Balıkesir University, Balıkesir, Türkiye.

**Corresponding Author. Email: aliyigit.sabaner@balikesir.edu.tr (A. Y. Sabaner)*

Article Information

Received: 30 September 2026
Accepted: 29 October 2026
Published: 30 October 2026

AMS 2020 Classification:
90C29, 90C11, 90C59, 90B50

Abstract

Configuring a battery electric bus from a supplier catalogue requires the joint choice of mandatory modules and optional equipment under mass, energy, charging and performance requirements. This study formulates the task as a three-objective combinatorial problem that minimises a present-value cost index, maximises the deterministic end-of-duty energy reserve and maximises the value of the selected equipment. The vehicle and duty model is calibrated to measurements of a bus manufacturer (SORT-2 consumption 0.77 kWh/km; field consumption 0.85–1.35 kWh/km; registered gross mass 19,000–19,500 kg). Two cases are solved exactly over 1,050 scenarios of base mass, HVAC load, gross-mass limit and daily duty: a catalogue benchmark with an ordinal equipment score, and the manufacturer's real catalogue of 20 option groups whose value is measured by sales shares in ten countries. Exact fronts are obtained by certified enumeration and by a new group-wise Pareto merge, and an ε-constraint mixed-integer programme reproduces the enumeration in all 1,050 scenarios. The benchmark yields 2–63 and the real catalogue 192–748 distinct Pareto objective vectors per scenario. Within the measured consumption range the gross-mass limit, not energy, is binding except on long duties. NSGA-II and NSGA-III, run under one protocol, found the complete exact front in 97.0 % and 97.3 % of 7,190 runs each; when fronts exceed the population size, and to a lesser extent even when they do not, the final population retains only part of the front that the search has found, so the population must be sized to the front or complemented by an archive. Decisions on battery and motor are unchanged under all 47 tested monotone transformations of the ordinal score, whereas equipment packages are not.

Keywords: Electric bus, product configuration, multi-objective optimization, ε-constraint method, NSGA-III, Pareto front, equipment value

1. Introduction

Battery electric buses are now procured by many transit operators, and the resulting planning problems, from charging infrastructure and fleet composition to vehicle and crew scheduling, have become an active research area [1]. Before any of these operational decisions, however, each vehicle has to be specified: a battery, a traction motor, a gearbox, a charging interface, a thermal system, a body layout and a set of optional equipment must be selected from the catalogues of manufacturers and suppliers. This choice fixes the investment, the energy consumption and the energy reserve of the vehicle for many years, and life-cycle studies show that it strongly affects total cost of ownership [2] and the battery capacity needed to complete a duty cycle [3], [4].

From an operations research perspective the specification task is a product configuration problem: one alternative must be chosen from each mandatory component group and optional equipment may be added, subject to compatibility, capacity and performance rules [5], [6]. In the electric-bus setting the rules are strongly coupled through physics. Every added kilogram increases traction energy, every kilowatt of continuous auxiliary load consumes energy over the whole service day, and a larger battery increases both usable energy and vehicle mass, which a registered gross-mass limit caps. Consequently, the cheapest feasible package rarely offers the largest energy margin, and equipment that adds safety, comfort or information consumes both money and energy.

Optional equipment is bought because it has value for operators and passengers, but this value is hard to quantify. If it is ignored, a cost–reserve model never selects more equipment than required: in the case studied here the two-objective special case has at most 3 distinct Pareto objective vectors per scenario and a single one in 152 of 719 feasible scenarios (Section 4.10). If equipment is valued by catalogue scores, these are typically ordinal, and an additive aggregate of ordinal scores supports only conclusions that are invariant under every order-preserving transformation of the scale [7]; measured, additive value functions require stronger information [8]. A configuration study that adds an equipment-value objective therefore has to show which of its conclusions depend on how value is measured.

This paper formulates catalogue-based electric-bus configuration as a three-objective problem (minimise a present-value cost index, maximise the deterministic end-of-duty energy reserve and maximise equipment value) and studies it on two cases. The vehicle and duty model of both cases is calibrated to measurements supplied by a battery-electric bus manufacturer. Case 1 is a catalogue benchmark whose equipment carries ordinal scores; it is used to compare solution methods and to test which conclusions remain unchanged under 47 tested monotone transformations of the scores. Case 2 is the manufacturer's real optional-equipment catalogue with 20 option groups, in which the value of an equipment level is its share in the manufacturer's sales in ten countries. Exact fronts are computed by certified complete enumeration, by a new group-wise Pareto merge and by an ε-constraint mixed-integer linear programme (MILP) [9], [10], and two evolutionary algorithms, NSGA-II [11] and NSGA-III [12], [13], are evaluated against them under a common protocol.

The study addresses four research questions. RQ1: Does an equipment-value objective conflict with cost and reserve, and which constraints shape the decision space of a calibrated vehicle? RQ2: Can exact methods generate the complete three-dimensional Pareto front reliably and at acceptable cost? RQ3: How close do NSGA-II and NSGA-III come to the exact front, and how do population size, budget and front size affect what they return? RQ4: Which decisions are robust to the energy consumption, to the scale of the value measure and to differences in preferences between markets?

The contributions are fivefold. (i) A three-objective formulation of catalogue-based electric-bus configuration calibrated to manufacturer measurements, with exact fronts for 1,050 scenarios of base mass, HVAC load, gross-mass limit and daily duty. (ii) A sales-based equipment-value model on a real option catalogue, together with a sensitivity analysis over 47 monotone score transformations that separates ordinal-robust from scale-dependent conclusions. (iii) Three exact methods with stated and verified exactness conditions, including a group-wise Pareto merge that reduces 28.3 million option combinations to 1,633 candidate bundles. (iv) A comparison of NSGA-II and NSGA-III against known exact fronts that distinguishes what the search finds (archive) from what the final population retains, across population sizes, budgets and front sizes. (v) A public reproduction package with code, inputs, all fronts, certificates and hashes. Section 2 reviews the literature, Section 3 describes the model and methods, Section 4 reports the experiments and Section 5 concludes.

2. Literature review

2.1. Electric bus planning and vehicle sizing

Research on battery electric buses spans strategic, tactical and operational decisions [1]. At the vehicle level, energy consumption models relate route topography, speed, stops, passenger load and auxiliary demand to the energy drawn from the battery; Vepsäläinen et al. [14] proposed a computationally efficient model for city buses under varying operating conditions, and Basma et al. [4] built a battery-sizing framework showing that consumption ranges from about 2 to 4.6 kWh/km depending on the type of service. Gao et al. [3] studied battery capacity and recharging needs in city transit service. Economic studies compared the life-cycle costs of bus powertrains [15] and of electric buses with different charging methods [2]. At the system level, feasibility studies of fast charging [16] and optimisation models for charger placement and battery sizing [17] or for fleet size and mix with charging infrastructure [18] highlighted the trade-off between on-board battery capacity and infrastructure. These models, however, treat the vehicle as a small number of continuous or discrete sizing variables; they do not select a complete configuration from a catalogue that also contains optional equipment with its own cost, mass and electrical load.

Standardised on-road test cycles such as the UITP SORT cycles [19] make consumption figures of different buses comparable; their E-SORT extension reports the consumption of battery electric buses on defined urban cycles. Such a measured figure is used here to calibrate the traction part of the energy model.

2.2. Product configuration and its optimisation

Product configuration research has a long tradition in knowledge-based systems, where constraint models encode which components can be combined [5], [6]. When several configurations are admissible, optimisation selects among them. Li et al. [20] optimised product configurations with a multi-objective genetic algorithm, Wei et al. [21] used an improved NSGA-II for modular product configuration, and Song and Chan [22] applied NSGA-II to a product-extension service configuration with performance, cost and response-time objectives. A common feature

of these studies is that the evolutionary front is reported without an exact reference, so its completeness and quality cannot be measured absolutely. For small or medium catalogues, however, exact multi-objective methods can provide such a reference.

2.3. Valuing optional equipment

Multiattribute value theory justifies an additive value function when the attributes are mutually preferentially independent and their value differences are measured on an interval scale [8], [23]. Catalogue scores, star ratings and expert ranks are often only ordinal. Measurement theory then restricts meaningful statements to those whose truth is unchanged by any strictly increasing transformation of the scores [7]; a Pareto front built on summed ordinal scores is not such a statement in general. Observed choices, such as the share of buyers who select a given equipment level, provide a ratio-scale alternative that does not depend on an analyst's scoring.

2.4. Exact and evolutionary multi-objective optimisation

The ε-constraint method optimises one objective while the others are bounded [9], [24]. Mavrotas [10] discussed its effective implementation, and the augmented variant AUGMECON2 [25] generates the exact Pareto set of multi-objective integer programmes. Kirlik and Sayın [26] proposed an algorithm that generates all nondominated points of multiobjective discrete problems, and Halffmann et al. [27] surveyed exact algorithms for multiobjective integer linear optimisation. On the evolutionary side, NSGA-II [11] remains a standard baseline for two and three objectives, while NSGA-III [12] replaces crowding by niching around reference directions and has a constrained variant [13]. Both return a population of fixed size; an unbounded external archive keeps every nondominated solution found during the search [28]. Front quality is commonly measured by the hypervolume (HV) [29] and by the inverted generational distance, whose modified form IGD+ is weakly Pareto compliant [30]; Li and Yao [31] surveyed these and other indicators.

2.5. Research gap

Four gaps emerge. First, electric-bus studies size a few components but do not configure a complete vehicle, including optional equipment, from a catalogue, and rarely calibrate the vehicle model to manufacturer measurements. Second, configuration studies that value equipment by scores do not examine whether their conclusions survive admissible changes of the scale. Third, configuration studies that use evolutionary algorithms rarely have an exact front to compare against. Fourth, comparisons of evolutionary algorithms seldom separate what the search finds from what the final population retains when fronts are larger than the population. This paper addresses these gaps.

3. Methodology

Figure 1 summarises the framework. In every scenario cell the exact and the evolutionary methods solve the same problem, and every objective vector they return is re-evaluated by the production evaluator of the model before any indicator is computed.

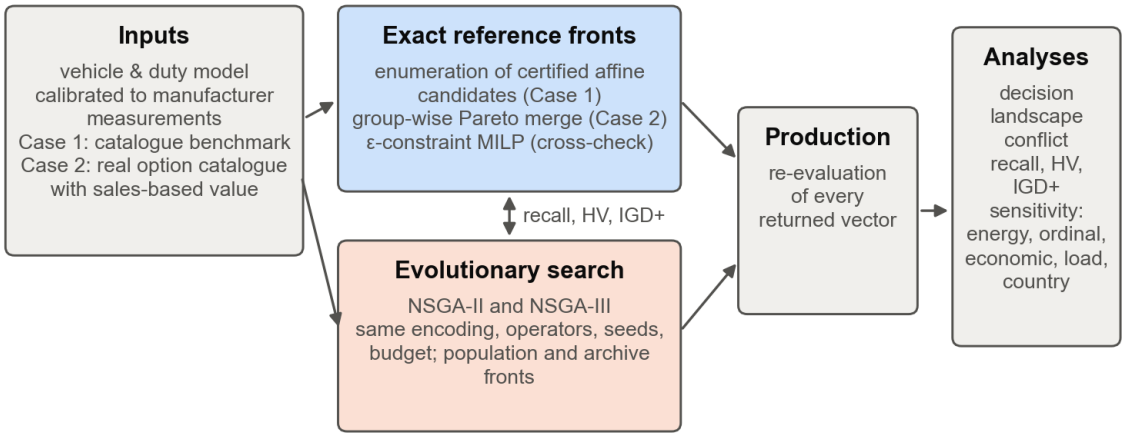


Figure 1. Solution and evaluation framework of the study.

3.1. Vehicle, duty and calibration

The vehicle is a 12 m battery-electric city bus. One item must be selected from each of nine mandatory groups: battery (220, 449, 523, 602 or 662 kWh), traction motor (150, 300 or 450 kW), gearbox, charging interface, thermal system, body layout, axle, fire-safety module and battery thermal management (Table 1). The mandatory groups generate 14,580 cores. A daily duty of L km is driven in six flat segments at the equivalent cruise speed v_c with 3 stop events per km, and the service time is L divided by the commercial speed of 16.575 km/h (253.6 km in 15.3

h, the reference duty of the instance); auxiliary and HVAC loads act over the whole service time. Table 2 lists the requirements and the scenario axes.

Table 1. Mandatory component groups (both cases).

Group	Alternatives	Key attribute	Cost index	Mass (kg)
Battery	5	220; 449; 523; 602; 662 kWh	198–595.8	1364–4104
Traction motor	3	150; 300; 450 kW	180–400	210–630
Gearbox	3	ratio 22; ratio 32	30–90	85–150
Charging interface	3	150; 450 kW	80–320	40–95
Thermal system	3	–	120–180	180–230
Body layout	6	18; 24; 60; 80; 105; 150 passengers	0–45	300–1500
Axle	1	–	120–120	900–900
Fire safety	3	–	30–150	20–60
Battery thermal	2	–	60–90	45–70

Table 2. Requirements and scenario axes.

Parameter	Value	Status
Daily duty L	200, 225, 250, 275, 300 km; 6 flat segments; 3 stops/km; v_c = 35 km/h; service time L / 16.575 km/h	Scenario; calibrated
Passengers	75 × 70 kg in the energy model; registered capacity of the body × 70 kg for mass checks	Requirement
Charging	150 kW depot charging for at most 5 h; efficiency 0.96	Case input
Battery and reserve	capacity > 450 kWh; retained buffer 20 kWh; reserve ≥ 25 kWh	Requirement
Performance and warranty	gradeability 15 % at 15 km/h; warranty ≥ 96 months	Requirement
Cost basis	12 years; r = 0.08; 330 days/year; 0.0022 index/kWh; maintenance 3.5 %; residual 12 %	Assumed index
Base mass axis m_b	6,500–8,500 kg, step 100 kg (21 levels)	Scenario; calibrated
HVAC axis H	0, 2, 4, 6, 8 kW (average over the service day)	Scenario; calibrated
Gross-mass limit G	19,000; 19,500 kg	Manufacturer

The model was calibrated to measurements communicated by a battery-electric bus manufacturer based in Bursa, Türkiye, which is not named at its request. The rules were fixed before the parameters were solved.

(i) The equivalent cruise speed reproduces the manufacturer's E-SORT consumption of 0.77 kWh/km under SORT-2-like conditions [19] (flat track, 3 stops per km, half load of 45 passengers, no HVAC, 3 kW base auxiliary load at 18.6 km/h): v_c = 35 km/h gives 0.769 kWh/km.

(ii) The HVAC axis is the average electrical HVAC and heating load over the service day; its end points are chosen so that the design-day consumption of the reference vehicle (curb mass 13,250 kg, 75 passengers) spans the measured field range from mild operation (0.85–1.00 kWh/km) to intensive HVAC use (1.10–1.35 kWh/km): 0 kW gives 0.851 and 8 kW gives 1.333 kWh/km.

(iii) The gross-mass limits are the registered 19,000 kg and the approved 19,500 kg, the base-mass axis places the curb mass of tender-compliant vehicles in the measured range of 12.5–14.0 t, and the masses of the 449 and 523 kWh batteries are the manufacturer's catalogue values. Table 3 compares the calibrated model with the measurements.

Table 3. Calibrated model and manufacturer measurements.

Quantity	Manufacturer	Model
M1 SORT-2-like consumption (kWh/km)	0.77	0.769
M2/M3 design-day consumption, HVAC axis lower end (kWh/km)	0.85-1.00 (mild)	0.851 at 0 kW
M2/M3 design-day consumption, HVAC axis upper end (kWh/km)	1.10-1.35 (intensive)	1.333 at 8 kW
M8 total average auxiliary load, mild weather (kW)	4-8	3-11 over the HVAC axis (3 kW base + HVAC)
M5 curb mass of the reference vehicle over the base-mass axis (kg)	12,500-14,000	12,300-14,300 (523 kWh); 12,904-14,904 (662 kWh)
M4 gross-mass limits (kg)	19,000 (19,500 approved)	19,000 and 19,500
M6 battery mass 449 / 523 kWh (kg)	2,800 / 3,500	2,800 / 3,500

The scenario grid combines 21 base masses, 5 HVAC loads, 2 gross-mass limits and 5 daily duties into 1,050 cells. Four anchor cells are used for detailed analysis: A = (6,800 kg, 4 kW, 19,000 kg, 250 km), B = (7,200 kg, 8 kW, 19,000 kg, 300 km), C = (7,800 kg, 6 kW, 19,500 kg, 300 km) and the infeasible control D = (7,400 kg, 8 kW, 19,000 kg, 300 km).

3.2. Two equipment catalogues

3.2.1. Case 1 (catalogue benchmark)

Between two and six of 80 optional items are selected. Seventy-four plain accessories are identical (cost index 8, 10 kg, no electrical load, score 1); six named items differ in cost, mass, load and score (Table 4). Case 1 is a constructed benchmark instance, designed to test solution methods rather than to measure preferences. The six named items carry ordinal levels taken from the safety-relevance attribute of the catalogue from which the instance was built, and the 74 interchangeable accessories were introduced deliberately: they create a large symmetric decision space (326 million option subsets per core), which is a known difficulty for evolutionary search and a demanding test for exact methods. Only conclusions that held under all 47 tested monotone transformations of the scores are drawn from Case 1 (Section 4.9).

Table 4. Case 1 optional items (two to six are selected).

Item	Count	Cost index	Mass (kg)	Load (kW)	Score q_o
Plain accessory	74	8	10	0	1
Passenger screens	1	30	30	0.6	1
Speakers	1	12	12	0.2	1
Antenna	1	15	8	0.1	2
Heated glass	1	35	22	1.5	3
Camera system	1	40	25	0.4	5
Sensor package	1	60	20	0.5	9

3.2.2. Case 2 (real catalogue)

The manufacturer's optional-equipment catalogue has 20 groups with 50 levels, 49 of which are selectable because the first-aid kit is fixed. Exactly one level is selected per group (Table 5). Costs were confirmed level by level by the manufacturer and are converted to the cost index of the mandatory groups (1 index unit = 100 EUR, the conversion that reproduces the battery entries of the same cost table). Mass differences are the manufacturer's values. Continuous electrical loads come from the manufacturer's ranges (18 levels); the remaining 31 levels carry explicit assumptions, varied by ×0.5 and ×1.5 in Section 4.9. The first-aid kit is mandatory in every market and fixed. Two hard rules forbid the suppression-type fire systems with the diesel pre-heater. Product codes are anonymised.

Table 5. Case 2 option groups (cost index; mass difference to the lightest level; continuous load).

Category	Group	Levels	Cost index	Max. mass (kg)	Max. load (kW)
Safety	Camera system	6	0–13	5	0.1
Safety	Driver assistance (ADAS)	2	30–72	5	0.1
Safety	Fire detection and suppression	3	12–40	9	0.02
Safety	Tyre-pressure monitoring	2	0–0.7	0.5	0.005
Safety	Mirrors	2	10–35	15	0.175
Safety	Headlights	2	7–12	2	0.11
Safety	Alcohol interlock	2	0–7	1	0.01
Safety	First-aid kit	1	0.5–0.5	3.5	0
Passenger comfort	Glazing	4	20–28	76	0
Passenger comfort	Sun blinds	2	1–4.5	0.7	0.01
Passenger comfort	Passenger USB	2	0–5	2.5	0.15
Passenger comfort	Wi-Fi	2	0–0.45	1.3	0.03
Driver ergonomics	Driver seat	4	7–21	12	0.1
Driver ergonomics	Driver cabin	3	5–20	30	0
Accessibility and information	Ramp	2	10–40	3	0.01
Accessibility and information	Interior information display	2	0–1	3	0.1
Accessibility and information	Announcement system	2	0–5	7	0.05
Operations and compliance	Ticketing	2	3–5	10	0.15
Operations and compliance	Tachograph	2	0.1–8	5	0.01
Operations and compliance	Pre-heater	2	20–32	20	0

3.3. Energy, reserve and cost model

In Case 1 the structural constraints are Eq. (1); in Case 2 exactly one level is chosen in every option group (Eq. (2)):

$$\sum_{i\in I_g} x_i = 1, \forall g \in G, 2 \le \sum_{o\in O} x_o \le 6 \tag{1}$$

$$\sum_{i\in I_g} x_i = 1, \forall g \in G, \sum_{\ell\in L_g} x_{g\ell} = 1, \forall g \in \Gamma \tag{2}$$

The curb mass m_c adds the selected modules to the base mass. The energy model uses the route mass m_r with 75 passengers of m_p = 70 kg; the gross-mass, axle and gradeability checks use the design mass m_f with the registered capacity n_c of the body:

$$m_c = m_b + \sum_i m_i x_i, m_r = m_c + 75m_p, m_f = m_c + n_c m_p \tag{3}$$

Table 6 lists the notation.

Table 6. Notation.

Symbol	Meaning	Symbol	Meaning
G, I_g	mandatory groups; alternatives of group g	x_i, x_o, $x_{g\ell}$	selection of item i, option o, level ℓ of option group g
O; Γ, L_g	Case 1 options; Case 2 option groups and their levels	K, y_k	screened cores; selection of core k
m_b, m_c, m_r, m_f	base, curb, route and design mass (kg)	m_i, a_o, κ_o	item mass (kg), electrical load (kW), cost index
L, d_s, t_s, n_s	daily duty (km); segment distance (km), duration (h), stops	F_s, W_s	road-load force (N); stop energy (J)
v_s	segment speed (m/s) = v_c (km/h) / 3.6	c_{rr}, C_dA_f, ρ	rolling resistance; aerodynamic area (m²); air density (kg/m³)
η_m, η_g, η_r, η_c	motor, gear, regeneration, charging efficiency	P_a, H, P_r	auxiliary, HVAC load; regeneration limit (kW)
T	service time (h)	E	daily energy consumption (kWh)
Q, U, B	nominal capacity; usable energy; retained buffer (kWh)	K, K_B	capital cost index; battery part
A_F, ψ	annuity factor; capital-cost multiplier	δ, ρ_B	battery-replacement indicator (0/1); replacement cost fraction
f_1, f_2, f_3	cost index, reserve (kWh), equipment value	q_o; $v_{g\ell}$	Case 1 score; Case 2 sales share of level ℓ
M, A, C	option mass, load and cost totals	β_j, γ_j	slopes of output j per kg and per kW
ε_R, ε_S	reserve and value bounds	P^*, $\tilde{z}$	exact Pareto front; normalised objective vector

For segment s with distance d_s (km), speed v_s (m/s), grade angle θ_s (zero on the calibrated route) and duration $t_s = d_s/v_c$ (h), the road-load force (N) and the stop energy (J) of n_s stops from rest to v_s are

$$F_s = m_r g(c_{rr}\cos\theta_s + \sin\theta_s) + \frac{1}{2}\rho C_d A_f v_s^2, W_s = \frac{1}{2} m_r v_s^2 n_s \tag{4}$$

with g = 9.80665 m/s², ρ = 1.225 kg/m³, c_{rr} = 0.0062 and C_dA_f = 7.1 m². Road work is $F_s \cdot 1{,}000\, d_s$ (J), and 1 kWh = $3.6\cdot10^6$ J. With η_m = 0.94, η_g = 0.97, η_r = 0.82, the regeneration power limit P_r (the smaller of the battery charge acceptance and the motor regeneration power, kW), the auxiliary load P_a (3 kW base plus the loads of the selected options, kW), the HVAC load H (kW) and the service time T (h), the daily energy (kWh) is

$$E = \sum_s \left[\frac{max(10^3 F_s d_s, 0) + W_s}{3.6\cdot 10^6 \eta_m \eta_g} - min\left(\frac{\eta_r (max(-10^3 F_s d_s, 0) + W_s)}{3.6\cdot 10^6}, P_r t_s\right)\right] + (P_a + H)T \tag{5}$$

The usable energy is $U = Q \cdot SOH \cdot DoD$ with end-of-warranty state of health 0.92 and depth of discharge 0.90. The second objective is the deterministic end-of-duty reserve above the retained buffer B = 20 kWh:

$$f_2 = U - E - B \tag{6}$$

With capital cost index K, battery part K_B, horizon N = 12 years, discount rate r = 0.08, D = 330 operating days, energy price p = 0.0022 index/kWh, maintenance fraction μ = 0.035, residual fraction s = 0.12, charging efficiency η_c = 0.96 and annuity factor A_F, the first objective is the present-value cost index

$$f_1 = K(1 + \mu A_F - s(1+r)^{-N}) + A_F \frac{DpE}{\eta_c} + \delta \rho_B K_B (1+r)^{-8}, A_F = \sum_{t=1}^{N} (1+r)^{-t} \quad (7)$$

where $\delta \in \{0, 1\}$ indicates a battery replacement in year 8 at ρ_B = 0.55 of the battery cost (δ = 1 only in the economic sensitivity analysis). The index is dimensionless. Appendix A gives a worked calculation for one reported configuration that reproduces the evaluator exactly.

3.4. Third objective: equipment value

In Case 1 the value of the selected equipment is the sum of the ordinal scores q_o:

$$f_3 = \sum_{o \in O} q_o \, x_o \quad (8)$$

Because q_o is ordinal, only conclusions that hold under every strictly increasing transformation φ of the scores are meaningful [7]; Section 4.9 therefore re-solves Case 1 with $f_3 = \Sigma \varphi(q_o) \, x_o$ for 47 transformations, which provides sensitivity evidence rather than a proof of invariance under every such transformation. In Case 2 the value of level ℓ of group g is its share in the manufacturer's sales:

$$f_3 = \sum_{g \in \Gamma} \sum_{\ell \in L_g} v_{g\ell} \, x_{g\ell}, v_{g\ell} = \frac{n_{g\ell}}{\sum_{\ell' \in L_g} n_{g\ell'}} \quad (9)$$

where n_{gl} is the number of vehicles delivered with level ℓ in ten countries (pooled, or one country at a time). The data (market_counts.csv in the deposit) record, for every level and country, the number of buses that the manufacturer delivered with that level in four consecutive quarters, over its whole bus range rather than the 12 m class alone. The number of vehicles delivered, N_c, is 100 (DE), 80 (UK), 28 (NO), 82 (ES), 112 (FR), 105 (IT), 50 (PL), 30 (RO), 115 (US), 138 (TR), 840 in total. In 16 of the 19 groups the levels are mutually exclusive and their counts add up to N_c, except for at most 7 vehicles per group and country without a recorded level and 3 group–country pairs with 1–2 records more than N_c. Three groups differ by construction: vehicles without driver assistance (G02) or without a pre-heater (G20) are not listed (in Germany, 98 and 57 of the 100 delivered vehicles are listed), and in the driver-seat group (G13) the two swivel levels are subsets of the standard and premium levels, so that its counts add up to more than N_c (123 in Germany).

In Eq. (9) every share is taken within the recorded counts of its group; for G02 and G20 it is therefore the share among vehicles fitted with the equipment, which the requirements prescribe. f_3 is a sum of marginal sales shares, i.e. a sales-share-based proxy of equipment value on a ratio scale, not a measured utility, and it does not represent complementarities between groups. It decomposes into five categories: safety, passenger comfort, driver ergonomics, accessibility and information, and operations and compliance. Section 4.9 re-solves Case 2 with a delivery-based variant in which the G13 levels are made mutually exclusive (standard or premium, each with or without swivel) and every share is divided by the number of vehicles delivered; f_3 is then exactly the expected number of option groups in which the configuration matches a randomly drawn delivered vehicle, which follows from linearity of expectation and requires no independence between groups.

3.5. Feasibility constraints

A configuration is feasible if the reserve is at least 25 kWh; the design mass does not exceed the gross-mass limit G; the front and rear axle loads (40 % and 60 % of the design mass) do not exceed 13,000 kg and 23,000 kg; the motor and gear provide the power and wheel torque needed for 15 % gradeability at 15 km/h; the motor–gear speed limit is respected; the battery discharge power covers traction, auxiliary and HVAC demand; the overnight charge of 150 kW × 5 h × 0.96 covers the daily energy; the battery exceeds 450 kWh and meets the 96-month warranty; the body carries at least 75 passengers and fits the 12.3 m × 2.63 m envelope; the homologation class is M3; a mandatory safety system is present; and the compatibility rules hold (Case 1: a suppression-type fire module requires an active battery thermal system; Case 2: the manufacturer's two rules). All constraints are evaluated by a production evaluator and, independently, by an arithmetic oracle that shares no code with it. The problem is min f_1, max f_2, max f_3 subject to these constraints.

3.6. Exact method 1: certified enumeration (Case 1)

Identical optional items are interchangeable, so the 326,207,116 option subsets of a core fall into 248 symmetry classes. A class whose capital cost, mass, electrical load and direct energy are all no larger, and whose score is no smaller, than those of another class can neither yield a better objective vector nor be more often feasible, because every output is monotone in these totals (Appendix B, Propositions 1–2). Removing dominated classes leaves 41 option signatures. For each daily duty, each of the 14,580 cores is combined with every signature (597,780 evaluations); combinations that are infeasible even in the most favourable scenario are discarded, which leaves 13,284 candidates on 324 cores. For every candidate a certificate proves that no force-sign or regeneration-limit

change occurs inside the scenario domain, so every output is affine in base mass and HVAC load. The exact front of a cell is the three-dimensional nondominated filter of the feasible candidates (Algorithm 1). A vector z dominates z' if $f_1 \leq f'_1$, $f_2 \geq f'_2$ and $f_3 \geq f'_3$ with at least one strict inequality; equal vectors are counted once, so all front sizes are numbers of distinct Pareto objective vectors.

Algorithm 1. Exact front by certified enumeration (Case 1).

```
Input: catalogue, requirements, daily duty L, scenario domain
1: group optional subsets into symmetry classes of identical items            (248 classes)
2: keep classes not dominated in (capex, mass, load, direct energy, -score)   (41 signatures)
3: for each core k and signature s do
4:     evaluate (k,s) in the most favourable scenario; discard if infeasible
5:     certify a fixed energy branch; store affine coefficients in base mass and HVAC
6: for each cell (m_b, H, G) do
7:     evaluate all candidates, keep the feasible ones, apply the 3-D Pareto filter
8:     re-evaluate every Pareto vector with the production evaluator
Output: exact front of every cell
```

3.7. Exact method 2: group-wise Pareto merge (Case 2)

In Case 2 the 19 free option groups admit 28.3 million level combinations per core, which rules out enumeration. Each level contributes additively to the capital cost C, the mass M, the electrical load A and the value V of the bundle. Let $\mathfrak{B}_g$ be the bundles over the first g groups that are not dominated in $(C, M, A, -V)$. Then $\mathfrak{B}_{g+1}$ is the nondominated subset of $\mathfrak{B}_g$ extended by the levels of group g+1, because a partial bundle dominated by another remains dominated after any common completion; the two groups linked by the compatibility rules are merged as one joint group. By the monotonicity of Proposition 1, a bundle outside the final set $\mathfrak{B}$ is dominated, or infeasible, for every core and scenario (Appendix B, Proposition 5). The merge leaves 1,633 bundles. For each core the outputs are affine in $m_b + M$ and $H + A$; intercepts and slopes are measured with the lightest bundle and certified on 1,000 random (core, bundle, scenario) triples per duty (Algorithm 2).

Algorithm 2. Exact front by group-wise Pareto merge (Case 2).

```
Input: option groups with (cost, mass, load, value) per level; compatibility rules
1: B <- {empty bundle}; merge the groups linked by a rule into one joint group
2: for each option group g do
3:     B <- nondominated subset of {b + level : b in B, level in g} in (C,M,A,−V)
4: for each daily duty and core k: measure affine intercepts and slopes in (m_b + M,H + A);
certify
5: for each cell: evaluate every (k,b), keep feasible, 3-D filter per core, then across cores
6: re-evaluate every Pareto vector with the production evaluator
Output: exact front of every cell
```

3.8. Exact method 3: ε −constraint mixed-integer linear programme

The MILP is an independent cross-check for Case 1 and the only reference for enlarged catalogues. It has one binary variable x_o per optional item and one binary y_k per core that satisfies the option-independent requirements of the cell. With the option totals

$$M = \sum_{o \in O} m_o \, x_o, A = \sum_{o \in O} a_o \, x_o, C = \sum_{o \in O} \kappa_o \, x_o \tag{10}$$

and, for every core, the cost c^0_k, reserve r^0_k and option-dependent margins g^0_{jk} at a reference option set (M_0, A_0, C_0), the objectives and constraints are linear because the slopes β and γ of all outputs with respect to option mass and load are identical for all cores; this property is checked with an additional probe in every cell, and the model is rejected otherwise:

$$min f_1 = \sum_{k \in K} c^0_k \, y_k + \beta_c (M - M_0) + \gamma_c (A - A_0) + \psi (C - C_0) \tag{11}$$

$$f_2 = \sum_{k \in K} r^0_k \, y_k + \beta_r (M - M_0) + \gamma_r (A - A_0) \tag{12}$$

$$\sum_{k \in K} g^0_{jk} \, y_k + \beta_j (M - M_0) + \gamma_j (A - A_0) \geq 0, \forall j \in J \tag{13}$$

$$\sum_{k \in K} y_k = 1{,}2 \leq \sum_{o \in O} x_o \leq 6, f_2 \geq \varepsilon_R, f_3 \geq \varepsilon_S, x_o, y_k \in \{0, 1\} \tag{14}$$

Each subproblem, Eqs. (11)–(14) with Eq. (8), is solved lexicographically: cost is minimised, then reserve is maximised and finally value is maximised. For every integer score level ε_S the (cost, reserve) front is traced by raising ε_R to the last reserve plus $\delta_R = 10^{-6}$ kWh (Algorithm 3), and the union is filtered in three dimensions. The procedure cannot skip a nondominated vector as long as the smallest positive reserve gap between front vectors exceeds δ_R plus the solver tolerances (Appendix B, Proposition 4). Gurobi 13.0.2 [32] was used with zero optimality gap, feasibility, integrality and optimality tolerances of 10^{-9} and one thread, in the spirit of the augmented ε −constraint method [25].

Algorithm 3. Exact front by the ε-constraint MILP.

```
Input: scenario cell
1: screen cores on option-independent constraints; compute intercepts and slopes; verify
    catalogue-wide slopes
2: build the MILP with one binary per optional item and per retained core
3: for eps_S = smallest, ..., largest attainable score do
4:     eps_R <- -infinity
5:     repeat: solve lexmin (f1, -f2, -f3) s.t. f3 >= eps_S, f2 >= eps_R; if infeasible then
            break
6:             record the solution; eps_R <- f2* + 1e-6
7: return the 3-D nondominated union, re-evaluated by the production evaluator
```

3.9. Evolutionary methods: NSGA-II and NSGA-III

Both algorithms were run with pymoo 0.6.1.6 [33] on the same problem object. In Case 1 a genome contains nine integer genes (the mandatory groups) and 80 bits (the optional items); uniform crossover is applied with probability 0.9, mutation resamples each group gene with probability 1/9 and, with probability 0.5, adds, removes or swaps one item, and a repair keeps two to six items. In Case 2 a genome contains one integer gene per mandatory and per option group, with uniform crossover and per-gene resampling with probability $1/n_{genes}$; the compatibility rules act as constraints. The objectives ($f_1, -f_2, -f_3$) are minimised and the constraint margins, scaled to comparable magnitudes, enter as inequality constraints. NSGA-II [11] uses constrained domination with rank and crowding. NSGA-III [12], [13] uses Das–Dennis reference directions ($p = 18$, $H = 190$, population 200 unless stated) with constraint-violation-first mating and feasibility-first survival. In the library version used, the default NSGA-III tournament breaks ties in constraint violation with an unseeded generator; it was replaced by an equivalent tournament that uses the run's seeded generator, so every run is reproducible. The initial population is random; no exact solution is injected. Seeds 1–10 are used throughout.

Every run returns two fronts, both re-evaluated by the production evaluator: the *population front*, the feasible nondominated members of the final population; and the *archive front*, the distinct nondominated vectors among all feasible configurations evaluated during the run, i.e. an unbounded external archive [28]. The difference between the two separates what the search finds from what a population of fixed size can hold (Algorithm 4).

Algorithm 4. Common protocol for NSGA-II and NSGA-III.

```
Input: cell, algorithm, seed, population N, generations, (reference directions)
1: seed the generator; initialise N random genomes; evaluate (f1, -f2, -f3) and scaled margins
2: for each generation do
3:     select parents (NSGA-II: rank/crowding tournament; NSGA-III: constraint-violation
        tournament)
4:     crossover, mutation, repair; eliminate duplicates; evaluate offspring; add feasible ones
        to the archive
5:     survival (NSGA-II: rank + crowding; NSGA-III: rank + reference-direction niching)
6: population front <- feasible nondominated members of the final population
7: archive front <- nondominated subset of all feasible evaluated configurations
Output: both fronts, re-evaluated by the production evaluator
```

3.10. Performance indicators

Because the exact front P^* is known, every indicator measures absolute quality. Objective vectors in minimisation form are normalised with the ideal and nadir points of P^*, giving $\tilde{z}$. The hypervolume of a set P is the volume it dominates up to the reference point $r = (1.1, 1.1, 1.1)$ [29], and IGD+ is the average dominance-corrected distance from P^* to P [30], [31]:

$$HV(P) = \lambda\left(\bigcup_{p\in P}\left[\tilde{z}_p, r\right]\right), IGD^+(P) = \frac{1}{|P^*|}\sum_{z\in P^*} min_{p\in P}\left\|\max\left(\tilde{z}_p - \tilde{z}, 0\right)\right\| \quad (15)$$

where λ is the Lebesgue measure. Recall is the share of P^* vectors reproduced (cost within 10^{-6}, reserve within 10^{-6} kWh and equal value), and a run has full recall if it reproduces all of them. The HV ratio divides $HV(P)$ by

$HV(P^*)$. Results over seeds are reported descriptively: the number of full-recall runs, the mean, minimum and maximum recall and HV ratio, and the HV gap of every seed.

3.11. Verification protocol

The implementation is covered by 51 automated tests: calibration targets, duty construction, a brute-force check of the three-dimensional Pareto filter, the reduction certificate, the anchor fronts, equality of MILP and enumeration fronts, embedding of the two-objective front, seed determinism of both algorithms, the archive–population relation, the Case 2 data table, the group-wise merge against brute force on a sub-catalogue and the Case 2 fronts against the production evaluator. In Case 1 all 33,495 front vectors were re-evaluated by the production evaluator (maximum absolute error $2.3 \cdot 10^{-12}$) and 5,000 random comparisons of the affine model with the independent oracle gave a maximum error of $7.3 \cdot 10^{-12}$; in Case 2 the corresponding errors were $2.7 \cdot 10^{-12}$ over all 384,240 vectors and $9.1 \cdot 10^{-12}$ in the core certification. Code, inputs and outputs are bound by SHA-256 hashes.

4. Experimental tests

4.1. Experimental setup

All experiments were run on a Windows 11 computer with an Intel Core i5-13500HX processor (20 logical cores) and 32 GB of memory, using Python 3.12.13, NumPy 2.3.5, SciPy 1.17.1, pymoo 0.6.1.6 and Gurobi 13.0.2. The design, fixed before the runs, is given in Table 7. The comparison cells are the four anchors and nine feasible cells drawn with seed 20261016 (S1–S9). Parallel runs used up to 18 processes; the computing times in Table 9 were measured sequentially in a single process.

Table 7. Experimental design.

Experiment	Scope	Settings
Exact fronts, Case 1	all 1,050 cells	Algorithm 1; MILP (Algorithm 3) in every cell
Exact fronts, Case 2	all 1,050 cells	Algorithm 2, pooled sales value
Algorithm comparison	$A - D + S1 - S9$; every feasible cell	NSGA-II and NSGA-III, N = 200, 250 generations, seeds 1–10
Population, budget, front size	anchor A, catalogue replicated k = 1, 2, 4, 8	N = 100, 200, 400, 800; 50,000 or 250 N evaluations
Reference-direction density	anchors A–C	NSGA-III with p = 12, 18, 24
Algorithms on the real catalogue	Case 2: $A, B, C, S1 - S3$	N = 200 and 800, 250 generations, seeds 1–10
Energy sensitivity	all cells	E × 0.85–1.15; traction or auxiliary part × 0.85, 1.15
Ordinal sensitivity	all cells, Case 1	47 strictly increasing score transformations
Economic sensitivity	anchors $A - C$	battery price × tariff × replacement, 50 cases each
Preference heterogeneity, load	Case 2, comparison cells	10 country value models; option loads × 0.5, 1.5

4.2. Decision landscape (RQ1)

Figure 2 maps the number of distinct Pareto objective vectors in Case 1. Of the 1,050 cells, 719 are feasible; every feasible cell has a genuine trade-off with 2 to 63 vectors (median 61), and none is a singleton. The feasible domain is bounded by the gross-mass limit: no configuration is feasible above a base mass of 7,600 kg at 19,000 kg or 8,100 kg at 19,500 kg, because the design mass with the registered capacity then exceeds the limit. Within the measured consumption range the energy constraints bind only on long duties: the front of a cell is smaller than at 200 km and 0 kW in 83 cells, all with a daily duty of 275–300 km and an HVAC load of 6–8 kW. At anchor B (300 km, 8 kW) the 523 kWh battery is no longer feasible and only the 602 kWh battery remains. A calibrated 12 m bus is therefore primarily mass-limited, and energy decides the battery only for long, HVAC-intensive duties. Because the 74 plain accessories are interchangeable, one distinct vector can stand for many item selections, up to $1.9 \cdot 10^8$ in Case 1; the published result files list this number for every vector.

4.3. Objective conflict

The conflict between the objectives was measured by Spearman rank correlations on every front with at least four vectors (Figure 3). In Case 1, cost and value are positively correlated on every front (median ρ = 0.79) and reserve and value negatively on every front (median −0.33): more equipment always costs money and energy. In Case 2 the same signs hold (medians 0.56 and −0.37). The cost–reserve correlation is weaker in Case 1 (median 0.26, positive on 56 % of the fronts) than in Case 2 (0.55); it is positive across battery tiers, where a larger battery costs more and stores more, and negative within a tier, where added equipment raises cost and lowers reserve. The answer to RQ1 is affirmative: the value objective conflicts with both other objectives in both cases.

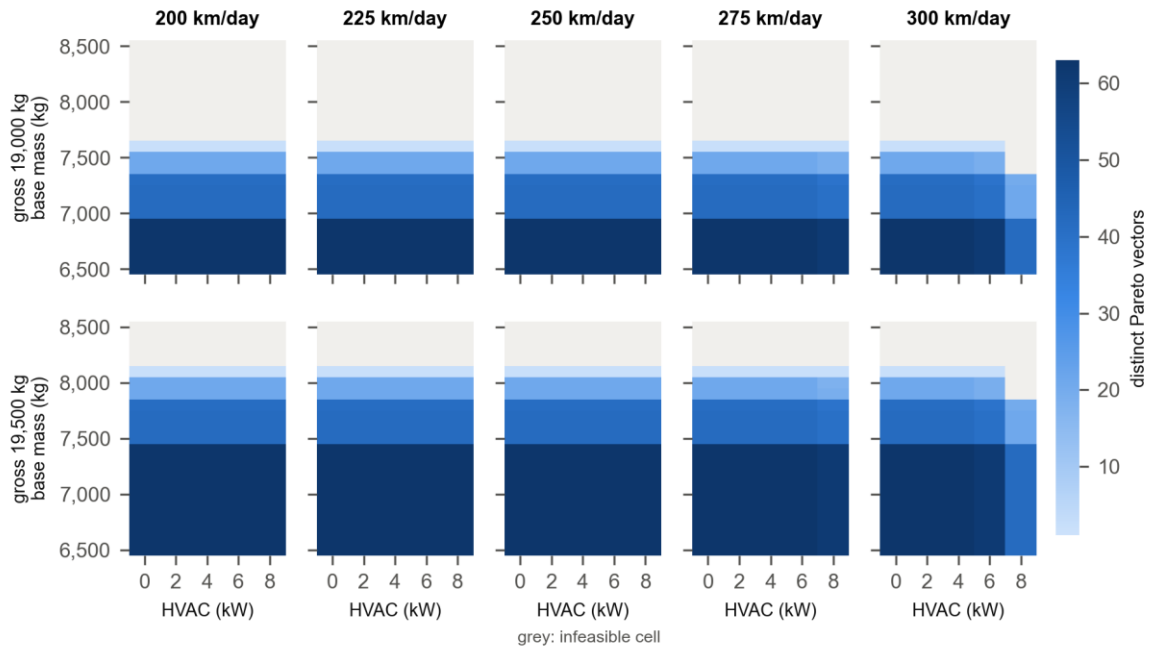


Figure 2. Case 1: number of distinct Pareto objective vectors per cell by daily duty (columns) and gross-mass limit (rows); grey cells are infeasible.

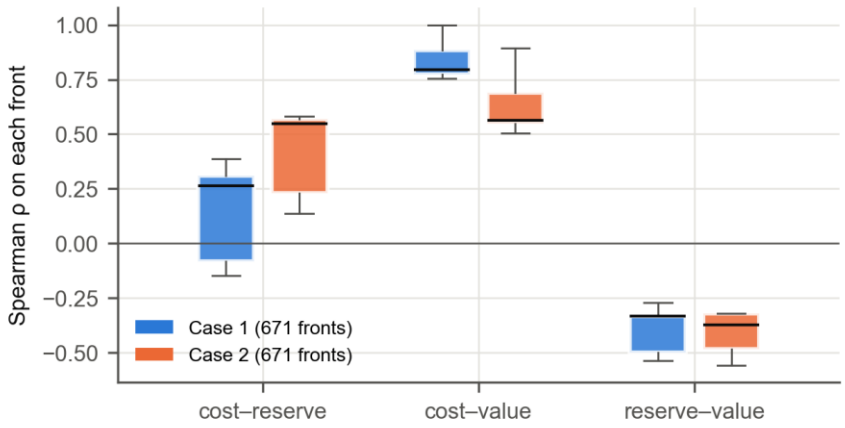


Figure 3. Spearman rank correlations between the objectives on every front with at least four vectors (boxes: quartiles and median; whiskers: 1.5 × interquartile range; outliers not shown).

4.4. Anchor fronts and aspiration-based selection

Figure 4 shows the Case 1 fronts at anchors $A - C$. At A all three tender-compliant batteries appear, the cost index ranges from 2,884 to 3,462 and the reserve from 107 to 256 kWh. At B only the 602 kWh battery is feasible and the reserve falls to 33–79 kWh. A decision maker can select from a front with aspiration levels: the least-cost configuration with reserve $\geq R$ and value $\geq Q$. Table 8 shows that at A any reserve up to 100 kWh is reached with the 523 kWh battery, a reserve of 150 kWh requires the 602 kWh battery at every value level, and a reserve of 200 kWh with a value of at least 15 requires the 662 kWh battery.

4.5. Exact methods (RQ2)

The ε-constraint MILP and the certified enumeration agree vector by vector in all 1,050 cells of Case 1: 33,495 of 33,495 vectors, and empty fronts in the 331 infeasible cells. The smallest positive reserve gap between front vectors of a cell is 0.0588 kWh, 58,844 times the ε step, so the step cannot skip a vector. Table 9 compares the computing times. Enumeration needs one compilation per duty (about 41 s), after which a cell is answered in milliseconds; the MILP needs 4–5 s per cell without any compilation; one evolutionary run takes about 9–10 s. In Case 2 the merge and the core compilation take well under a second and an exact front 0.5–4.6 s, although the front has hundreds of vectors.

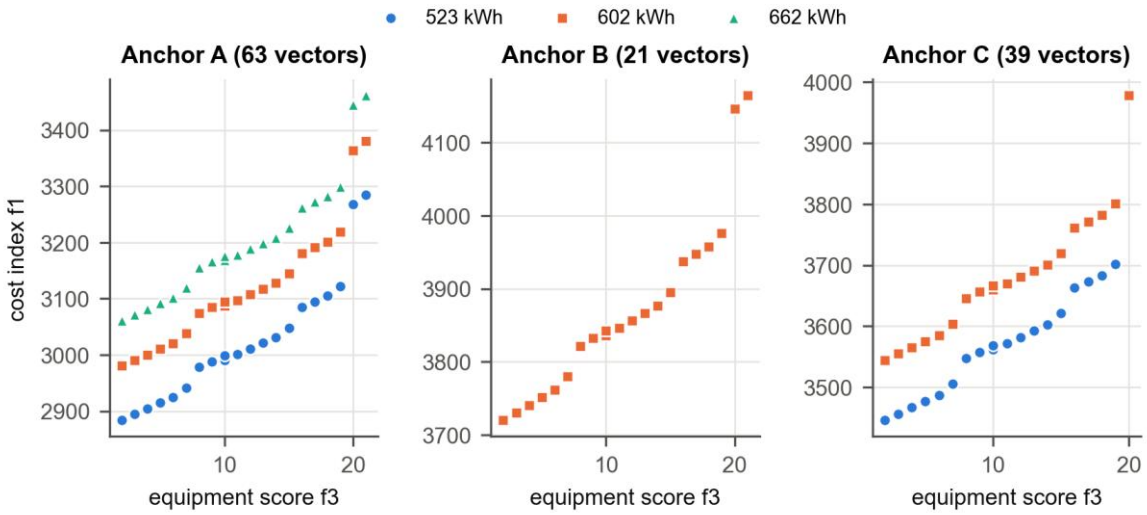


Figure 4. Case 1 exact fronts at anchors $A - C$ (cost index against equipment score; marker = battery).

Table 8. Aspiration queries at anchor A (Case 1): least-cost configuration with reserve $\geq R$ and score $\geq Q$.

R (kWh)	Q	Cost index	Reserve (kWh)	Score	Battery (kWh)
25	2	2,884.00	144.86	2	523
25	10	2,990.64	137.25	10	523
25	15	3,048.25	135.46	15	523
25	21	3,284.46	106.60	21	523
150	2	2,980.19	208.57	2	602
150	10	3,086.83	200.95	10	602
150	15	3,144.45	199.17	15	602
150	21	3,380.65	170.31	21	602
200	2	2,980.19	208.57	2	602
200	10	3,086.83	200.95	10	602
200	15	3,225.71	246.11	15	662
200	21	3,461.91	217.25	21	662

Table 9. Sequential single-process computing times (s) at anchors $A - C$.

Case	Anchor	$\lvert P \rvert$	Preparation	Exact front	MILP	NSGA-II	NSGA-III
Case 1	A	63	41.36	0.006	4.7	9.0	10.0
Case 1	B	21	43.43	0.004	4.9	9.7	9.6
Case 1	C	39	40.51	0.004	4.0	8.7	9.6
Case 2	A	734	0.73	4.563	–	9.8	10.2
Case 2	B	228	0.74	0.537	–	8.0	8.6
Case 2	C	416	0.73	0.940	–	8.4	9.1

Preparation is the per-duty compilation (Case 1) or the bundle merge and core compilation (Case 2). The answer to RQ2 is affirmative: both exact references are complete under stated and verified conditions (Appendix B), and the MILP provides an independent confirmation in every Case 1 cell.

4.6. Evolutionary versus exact fronts (RQ3)

Table 10 compares the algorithms in the twelve feasible comparison cells (anchor D is infeasible and both algorithms return no feasible point there). NSGA-II found and retained the complete exact front in 106 of 120 runs. NSGA-III found it in 107 runs but retained it in its final population in only 60: at fronts of 40–63 vectors, its niching around 190 reference directions drops individual vectors that its search had visited. Over every feasible cell of the grid (7,190 runs per algorithm), NSGA-II reached full recall in 97.0 % of the runs in archive and population alike, and NSGA-III in 97.3 % (archive) and 32.4 % (population). The misses concentrate in cells just below the gross-mass limit, where only two configurations are feasible and a few runs never reach them. A denser set of reference directions improves the retention of NSGA-III: at anchor A, full population recall rises from 0/10 (p = 12) to 1/10 (p = 18) and 7/10 (p = 24).

Table 10. NSGA-II and NSGA-III against the exact Case 1 fronts (10 seeds per cell; full recall: population/archive).

Cell	$\|P\|$	NSGA-II full	NSGA-III full	NSGA-II min recall	NSGA-III min recall	NSGA-II min HV	NSGA-III min HV
A	63	10/10	1/10	1.000	0.937	1.0000	0.9976
B	21	10/10	10/10	1.000	1.000	1.0000	1.0000
C	39	2/2	1/1	0.974	0.923	0.9962	0.9929
S1	63	10/10	1/10	1.000	0.937	1.0000	0.9966
S2	42	10/10	3/10	1.000	0.952	1.0000	0.9988
S3	21	8/8	10/10	0.952	1.000	0.9993	1.0000
S4	21	8/8	10/10	0.905	1.000	0.9980	1.0000
S5	2	10/10	10/10	1.000	1.000	1.0000	1.0000
S6	41	9/9	4/7	0.976	0.927	0.9960	0.9821
S7	2	9/9	9/9	0.000	0.000	0.0000	0.0000
S8	42	10/10	1/10	1.000	0.952	1.0000	0.9861
S9	63	10/10	0/10	1.000	0.905	1.0000	0.9966

4.7. Population size, budget and front size

To vary the front size, the six named options of Case 1 were replicated k times with jittered attributes; the exact reference for $k \geq 2$ is the MILP. Figure 5 shows the mean recall at a fixed budget of 50,000 evaluations. At $k = 8$ (333 vectors), NSGA-II retains 0.11, 0.28, 0.56 and 0.77 of the front with 100, 200, 400 and 800 individuals, while its archive holds 0.84 with 200 individuals. With 100 and 200 individuals the population is smaller than the front, so its recall cannot exceed N divided by the front size. Doubling the budget with 400 individuals leaves the population recall unchanged (0.56) and raises the archive recall to 0.95; 800 individuals and 200,000 evaluations reach 0.94 (archive 0.98). NSGA-III shows the same pattern at lower levels (0.25 to 0.65).

The populations of 400 and 800 exceed the front, so population capacity alone cannot explain their loss of recall. In these runs only 0.53 (N = 400) and 0.45 (N = 800) of the nondominated vectors of the final population are Pareto-optimal, i.e. the population has not converged; Pareto-optimal vectors that the search evaluated (archive recall 0.88 with 400 individuals) are not retained by the selection and survival steps; and at a fixed budget 800 individuals have only 62 generations, so that the archive recall itself falls to 0.78. The decline of the population recall with catalogue size is therefore a lower bound on what the algorithms find. It reflects the retention mechanism of the final population, bounded by capacity when the population is smaller than the front and by incomplete convergence and survival when it is larger, rather than a failure of the search to visit the front; this study does not separate these mechanisms further. Because the replicated catalogue is synthetic, the same question was examined on the real catalogue (Section 4.8).

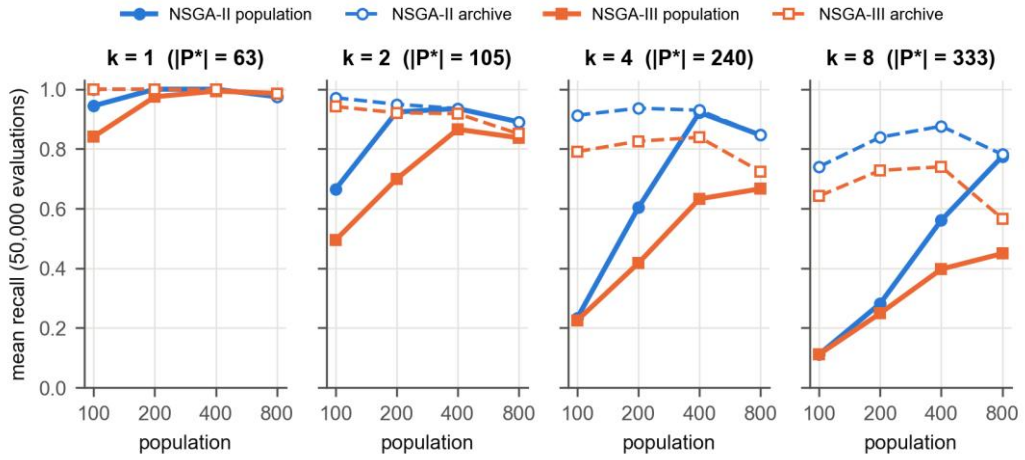


Figure 5. Mean recall of the population front (solid) and of the archive front (dashed) against population size at 50,000 evaluations, catalogue replicated k times (anchor A, 10 seeds).

4.8. Case 2: the real catalogue

With the manufacturer's catalogue and the pooled sales value, 671 of the 1,050 cells are feasible and the exact fronts contain 192–748 distinct vectors (median 672; 384,240 in total). Fewer cells are feasible than in Case 1 because the real options are heavier. At anchor A (Figure 6) the front has 734 vectors; the value ranges from 5.34 to 13.28 and the cost index from 3,038 to 3,502. Safety equipment contributes the largest share of value on the whole front, whereas accessibility and information equipment and passenger-comfort equipment account for most of the value gained towards the high-value end (Figure 6b).

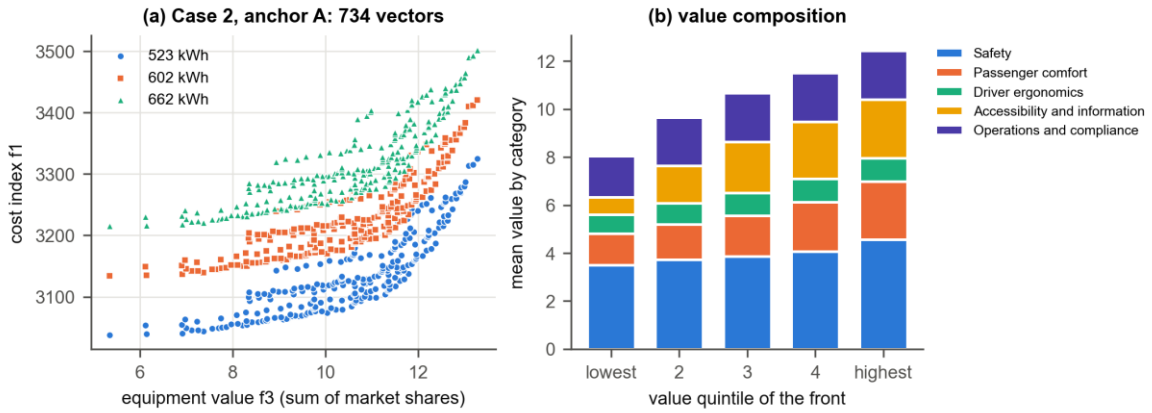


Figure 6. Case 2 at anchor A: (a) exact front, cost index against equipment value (marker = battery); (b) mean value by category along the value quintiles of the front.

Table 11. Case 2: mean recall (population / archive) of NSGA-II and NSGA-III against the exact fronts (10 seeds; N = 200 and 800, 250 generations).

Cell	$\lvert P \rvert$	NSGA-II, N = 200	NSGA-III, N = 200	NSGA-II, N = 800	NSGA-III, N = 800
A	734	0.21 / 0.95	0.23 / 0.98	1.000 / 1.000	0.999 / 0.999
B	228	0.85 / 1.00	0.87 / 1.00	1.000 / 1.000	1.000 / 1.000
C	416	0.43 / 0.99	0.44 / 0.99	1.000 / 1.000	1.000 / 1.000
S1	744	0.21 / 0.95	0.23 / 0.97	1.000 / 1.000	1.000 / 1.000
S2	494	0.35 / 0.99	0.36 / 0.99	1.000 / 1.000	1.000 / 1.000
S3	223	0.88 / 1.00	0.89 / 1.00	1.000 / 1.000	1.000 / 1.000

Table 11 confirms the replication experiment on real data. With 200 individuals both algorithms retain 0.21–0.89 of fronts with 223–744 vectors in their final population, although their archives contain 0.95–1.00 of them (HV ratio of the archive ≥ 0.9993). With 800 individuals both reproduce nearly complete fronts (mean recall ≥ 0.999); the front was complete in 60 of 60 NSGA-II runs and 55 of 60 NSGA-III runs.

4.9. Robustness of the decisions (RQ4)

Scaling the daily energy of every configuration by 0.85–1.15 (Figure 7) changes the number of feasible cells from 725 to 696 (baseline 719); the least-cost battery changes in at most 59 cells, and the effect concentrates on the long duties. At anchor B the front has 42 vectors at −15 %, 21 at baseline and 0 at +15 %. Scaling only the auxiliary part has a larger effect than scaling only the traction part (Table 12). The signs of the conflict are unchanged in every case (median ρ cost–value 0.77–0.85, reserve–value −0.38 to −0.33). Because the baseline is calibrated to measured consumption, the ±15 % band covers the spread of the measured field values.

Table 12. Energy sensitivity over all 1,050 cells (Case 1).

Perturbation	Feasible cells	Vectors	Median $\lvert P \rvert$	Least-cost battery changed	$\lvert P \rvert$ at A / B / C
baseline	719	33,495	61	0	63 / 21 / 39
total × 0.85	725	34,175	63	23	63 / 42 / 41
total × 0.90	725	34,121	61	23	63 / 40 / 41
total × 0.95	725	33,918	57	23	63 / 35 / 39
total × 1.05	716	32,812	48	1	63 / 19 / 26
total × 1.10	707	32,144	42	46	63 / 15 / 19
total × 1.15	696	30,945	42	59	63 / 0 / 19
traction × 0.85	725	34,102	61	23	63 / 40 / 41
traction × 1.15	707	32,332	42	32	63 / 19 / 20
auxiliary × 0.85	725	34,111	61	23	63 / 40 / 41
auxiliary × 1.15	708	32,381	42	23	63 / 19 / 20

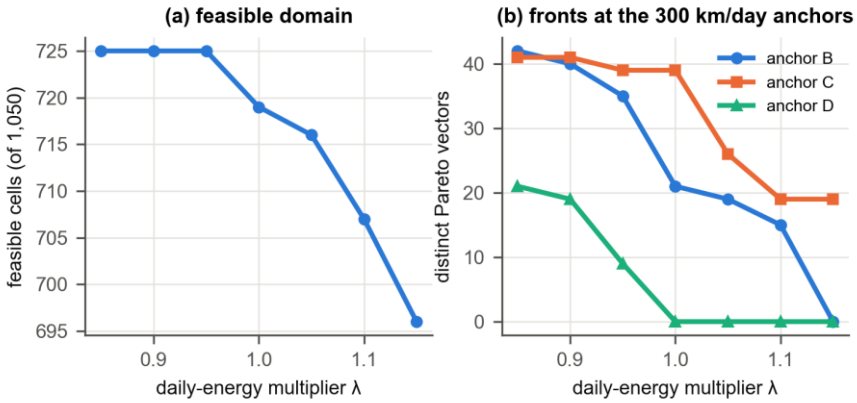


Figure 7. Energy sensitivity: (a) feasible cells and (b) front size at the 300 km/day anchors against the multiplier λ of the daily energy.

Case 1 was re-solved exactly in all cells under 47 strictly increasing transformations of the five score levels {1, 2, 3, 5, 9}: equal-interval rank coding, the powers 0.25, 0.5, 2 and 3, a logarithm and 40 random monotone maps (Table 13). In every transformation and every feasible cell the battery–motor choices on the front (719/719 cells) and the least-cost configuration (719/719) are unchanged. The equipment packages on the front are not: the median share of the original front configurations that remain on the transformed front is 0.81 over the random maps (minimum 0.57), and the number of vectors varies from 22,259 to 68,328. Conclusions about hardware therefore held under all 47 tested transformations, which is sensitivity evidence rather than a proof for every strictly increasing transformation, whereas recommendations of specific equipment packages require a cardinal value measure such as the sales shares of Case 2.

Table 13. Results under the 47 tested strictly increasing transformations of the Case 1 score (all feasible cells).

Transformation	Vectors	Median $\|P\|$	Battery–motor set unchanged	Least-cost unchanged	Median share of front kept
identity	33,495	61	719	719	1.00
rank	25,505	46	719	719	0.71
power 0.25	22,259	40	719	719	0.62
power 0.5	22,259	40	719	719	0.62
power 2	68,328	122	719	719	1.00
power 3	68,328	122	719	719	1.00
log	22,309	40	719	719	0.62
40 random maps	22,259–68,328	40–122	719	719	0.57–1.00

At anchors $A - C$ the battery price was scaled by 0.5–1.5, the tariff by 0.5–2 and a battery replacement in year 8 was added or not (50 cases per anchor). The front sizes stayed within 63–63 (A), 20–21 (B) and 37–39 (C), and the battery sets on the fronts did not change; only the cost index levels moved.

With the ten country-specific value models the fronts at anchor A contain 189–1483 vectors, and only 7.3 %–12.6 % of the pooled front configurations of the feasible comparison cells lie on all ten country fronts (Case 2). The least cost of reaching 90 % of the attainable value differs by 4.3 % between countries (Figure 8). Scaling the option loads by 0.5 or 1.5 leaves the battery choices unchanged, and 77 %–100 % of the front configurations are shared with the baseline. With the delivery-based variant of the value model (Section 3.4), the same 671 cells are feasible and the fronts contain 194–722 vectors (median 658; 372,896 in total); the battery–motor sets and the least-cost configuration are unchanged in all 671 cells, at least 95.1 % of the front configurations of every cell remain on the front, and 7.6 %–12.5 % of the pooled front configurations lie on all ten country fronts. The conclusions drawn from Case 2 therefore do not depend on how the driver-seat levels and the unlisted vehicles are counted.

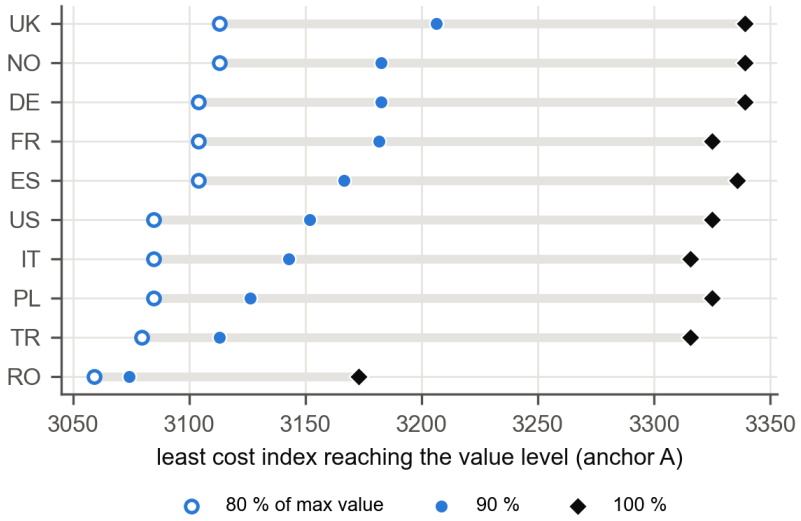


Figure 8. Case 2, anchor A: least cost index reaching 80 %, 90 % and 100 % of the attainable equipment value under the value model of each country.

4.10. Two-objective special case

Without the third objective the model reduces to a cost–reserve problem. Solved with the same compiler and grid, it has 719 feasible cells with only 1–3 distinct Pareto objective vectors; 152 cells have a single Pareto objective vector, and no configuration on these fronts carries more than the minimum two items. In every cell the two-objective front is embedded in the three-objective front. The value objective thus turns small, often single-point fronts into trade-off surfaces with up to 63 (Case 1) and 748 (Case 2) vectors.

4.11. Discussion and managerial implications

The calibration changed what limits the vehicle. With consumption tied to the measured field range, the registered gross-mass limit rather than the battery energy bounds the feasible configurations, and the battery choice is energy-driven only on long, HVAC-intensive duties; heavier vehicles cannot carry the required capacity within the limit, which agrees with the manufacturer's observation that standing capacity shrinks as batteries grow. Exact references were available for every cell, which turned the evaluation of NSGA-II and NSGA-III into an absolute measurement. Two findings stand out. First, both algorithms searched well: their archives contained the complete front in about 97 % of all runs. Second, what they return depends on how well the final population retains the front. NSGA-III, whose niching spreads a fixed population over reference directions, lost vectors at front sizes that NSGA-II retained; both lost most of the front when it exceeded the population size, and part of it even with populations larger than the front. Doubling the evaluations did not help at 400 individuals; more individuals with a larger budget, or an archive, did.

Three managerial implications follow, each conditional on the calibrated duty model and the stated requirements. (i) For 12 m buses with a registered gross mass of 19,000–19,500 kg and daily duties up to about 250 km, the mass budget and not the energy reserve decides which battery–equipment packages are admissible; for 275–300 km duties with intensive HVAC, the smallest tender-compliant battery may become infeasible. (ii) Hardware choices on the front were unchanged under all 47 tested monotone rescorings of equipment, so in this instance they did not depend on the choice of a cardinal value scale; equipment packages did. (iii) Preferred equipment packages differ between markets: only 7–13 % of the pooled front configurations (median about 8 %) are efficient in all ten markets studied, so a configurator should carry market-specific value models. When an evolutionary algorithm is used, its population should be sized to the expected front or an archive should be kept.

Internal validity is supported by independent evaluators, exact agreement of enumeration and MILP, certification of the affine models, production re-evaluation of every reported vector and automated tests. Construct validity is limited in three ways. The duty model is deterministic and flat with an equivalent cruise speed calibrated to one standard cycle; it reproduces the measured consumption range but not route-specific topography or traffic. The Case 1 scores are ordinal assumptions, which is why only conclusions that held under all 47 tested transformations are drawn from them. The Case 2 value is a sales-share-based proxy, i.e. revealed choices of past buyers rather than a measured utility of individual operators; as a sum of marginal shares it does not represent complementarities between option groups, and the delivery records cover the whole bus range, not only 12 m buses. External validity is limited to one vehicle class, one manufacturer and the stated requirements; the cost index is not a monetary price.

5. Conclusions and future research

This paper formulated the catalogue-based configuration of battery electric buses as a three-objective problem, calibrated it to manufacturer measurements and solved two cases exactly over 1,050 scenarios. Regarding RQ1, the equipment-value objective conflicts with cost and reserve in both cases, and the calibrated decision space is bounded mainly by the gross-mass limit. Regarding RQ2, certified enumeration, the new group-wise Pareto merge and the ε-constraint MILP provide complete fronts under stated conditions, and the MILP reproduced the enumeration in every cell. Regarding RQ3, NSGA-II and NSGA-III found the complete front in about 97 % of the runs, but the final population does not always retain it: the decline of recall with catalogue size reflects the population capacity when the front is larger than the population and incomplete convergence and survival when it is smaller. Regarding RQ4, hardware decisions were unchanged under all 47 tested monotone rescorings of equipment and under the economic parameters, energy matters on long duties, and equipment packages depend on the value measure and on the market. Future work should add stochastic duty cycles with route topography, value models elicited from operators and manufacturers, and exact methods for larger catalogues with more objectives.

Acknowledgments

The authors thank a battery-electric bus manufacturer based in Bursa, Türkiye, for providing measurement, catalogue and sales data; at its request the manufacturer is not named and product codes are anonymised.

Funding

None.

Data availability

The calibrated instance, both catalogues (anonymised), the sales-share value models, all computed fronts of both cases, the algorithm runs, the certificates, the test suite, the exact software environment and one-command run instructions are available at https://doi.org/10.5281/zenodo.23527139; the main result tables are included in the deposit as Supplementary_Data.xlsx. Version 1.1.0 of the deposit adds the delivery-based variant of the Case 2 value model and documents the delivery data. All results except the ε-constraint MILP fronts can be reproduced with open-source software; without a Gurobi licence, the published MILP fronts are compared with the recomputed enumeration.

Conflict of interest

The authors declare that there is no conflict of interest to disclose.

Author contributions

Ali Yiğit Sabaner: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – Original draft. **Ece Nur Polat:** Resources, Data curation, Investigation, Validation, Writing – Review & Editing. **Ramazan Yaman:** Conceptualization, Methodology, Supervision, Project administration, Writing – Review & Editing.

Declaration of using AI tools

During the preparation of this manuscript, the authors used Claude (Opus 5.5, Anthropic) exclusively for language refinement, grammatical corrections, and document formatting. The tool was not used for research design, scientific content generation, data analysis, or interpretation of results. All AI-assisted revisions were critically reviewed and verified by the authors, who assume full responsibility for the scientific integrity and final content of the manuscript.

Appendix A. Energy model: units, parameters and a worked calculation

Units: masses in kg, distances in km (road work uses metres: 1,000 d_s), speeds in km/h in the input and in m/s in the force and stop-energy terms ($v_s = v_c/3.6$), forces in N, energies in J converted to kWh by 1 kWh = $3.6 \cdot 10^6$ J, powers in kW and times in h. The route mass (75 passengers) enters the energy terms; the design mass (registered capacity) enters the gross-mass, axle and gradeability constraints. Auxiliary and HVAC loads act over the whole service time $T = L$ / 16.575 km/h, which consists of the driving time $\Sigma d_s/v_c$ and the dwell time at stops and terminals. Table A1 lists every parameter with its provenance, and Table A2 reproduces the least-cost configuration at anchor A (523 kWh battery, two plain accessories) step by step; every quantity equals the value of both evaluators (largest difference 0).

Table A1. Parameters of the energy, reserve and cost model with their provenance.

Parameter	Value	Provenance
Gross-mass limit G	19,000; 19,500 kg	Manufacturer (registered; approved tolerance)
Battery mass 449 / 523 kWh	2,800 / 3,500 kg	Manufacturer catalogue
Consumption targets	SORT-2 0.77; field 0.85–1.00 and 1.10–1.35 kWh/km	Manufacturer measurements (calibration targets)
Equivalent cruise speed v_c	35 km/h	Calibrated to the SORT-2 consumption
HVAC axis H	0–8 kW	Calibrated to the field consumption range
Base-mass axis m_b	6,500–8,500 kg	Calibrated to the curb-mass range 12.5–14.0 t
Stop density	3 stops per km	UITP SORT-2 [19]
Commercial speed (service time $T = L / v$)	253.6 km / 15.3 h = 16.575 km/h	Reference duty of the instance (scenario assumption)
g, ρ	9.80665 m/s², 1.225 kg/m³	Physical constants
c_{rr}, $C_d A_f$	0.0062; 7.1 m²	Catalogue assumption (typical 12 m city bus)
η_m, η_g, η_r	0.94, 0.97, 0.82	Catalogue assumption
Regeneration limit P_r	min(180, 300) = 180 kW	Catalogue: battery acceptance, motor regeneration power
SOH, DoD	0.92, 0.90	Catalogue assumption (end of warranty)
Base auxiliary load	3 kW	Catalogue assumption; manufacturer: 4–8 kW total in mild weather
Option loads a_o	see Tables 4–5	Case 1 catalogue; Case 2 manufacturer ranges or flagged assumptions
Passengers, mass per passenger	75 (energy), registered capacity (mass checks); 70 kg	Tender requirement
Charging	150 kW × 5 h, η_c = 0.96	Tender requirement; catalogue efficiency
Reserve floor, buffer, battery size	25 kWh, 20 kWh, > 450 kWh	Tender requirement
Economic parameters	N = 12, r = 0.08, D = 330, p = 0.0022, μ = 0.035, s = 0.12, ρ_B = 0.55	Assumed index (sensitivity in Section 4.9)

Table A2. Worked calculation at anchor A (6,800 kg, 4 kW, 19,000 kg, 250 km).

Quantity	Value	Quantity	Value
module mass	5,800.0000	regenerated energy	144.1159
curb mass	12,600.0000	driving time	7.1429
route mass	17,850.0000	dwell time	7.9400
design mass	18,200.0000	auxiliary energy	105.5797
drive efficiency	0.9118	daily energy E	268.1807
auxiliary power	7.0000	daily energy per km	1.0727
regeneration power limit	180.0000	usable energy	433.0440
segments x distance	250.0000	reserve f_2	144.8633
stop events per day	750.0000	capital cost index	1,114.7000
road-load force per segment	1,496.3528	annuity factor	7.5361
kinetic energy per stop	0.2343	cost index f_1	2,884.0011
traction energy (battery side)	306.7169		

Appendix B. Exactness of the reference fronts

Proposition 1 (monotone response). For a fixed core and duty, if the branch certificate holds over the mass interval of the domain, f_1 is non-decreasing in the option totals C, M and A, f_2 and every constraint margin are non-increasing in M and A and independent of C, and f_3 depends only on the value. *Proof.* With fixed force signs and a fixed side of the regeneration cap in every segment, E is affine in the route mass with a non-negative slope (verified by the certificate) and linear in A with slope $T > 0$; $f_2 = U - E - B$; $f_1 = \psi K + A_F D_p E/\eta_c$ with $\psi > 0$; the energy, mass, axle, gradeability and battery-power margins are non-increasing in E, the design mass or the load, and the remaining margins do not depend on (C, M, A).

Proposition 2 (option-class reduction). Removing option classes dominated in (C, M, A, direct energy, −value) preserves feasibility of every cell and every distinct nondominated objective vector. *Proof.* If class t dominates class s, then by Proposition 1, for every core and cell, (k, t) is feasible whenever (k, s) is and its objective vector weakly dominates that of (k, s); identical items are interchangeable and the reduction stops if an option appears in a compatibility rule.

Proposition 3 (affine compilation). Every retained candidate is affine in base mass and HVAC load on the domain (by the branch certificate). Verified by 5,000 random checks against the independent oracle (maximum error $7.3 \cdot 10^{-12}$).

Proposition 4 (ε −constraint MILP). With integer score levels, lexicographic solves and reserve step δ_R, the procedure returns every distinct nondominated vector if the smallest positive reserve gap between front vectors of equal score exceeds δ_R plus the solver tolerances. *Proof sketch.* At a fixed score level the lexicographic solution is the least-cost, then largest-reserve configuration with reserve $\geq \varepsilon_R$; raising ε_R by δ_R excludes only vectors with smaller reserve, and the next nondominated vector lies at least one gap higher. Evidence: the smallest gap is 0.0588 kWh against $\delta_R = 10^{-6}$ kWh, and the MILP equals the enumeration in all 1,050 cells.

Proposition 5 (group-wise merge). The final bundle set $\mathfrak{B}$ contains every option bundle that appears on the front of any core and cell. *Proof.* The totals are additive over groups; a partial bundle dominated in $(C, M, A, -V)$ stays dominated after any common completion, and Proposition 1 transfers dominance of totals to the objectives and margins. Evidence: equality with brute force on a six-group sub-catalogue, certification of the affine core models (maximum error $9.1 \cdot 10^{-12}$) and production re-evaluation of all 384,240 vectors.

Table A3 states, for every front set, the exact reference and the independent checks applied.

Table A3. Reference used for each front set and the independent checks applied.

Front set	Exact reference	Independent check	Production re-evaluation
Case 1, 1,050 cells	Enumeration (Props 2–3)	MILP identical in all cells	all 33,495 vectors
Case 1, energy ±15 %	Enumeration, perturbed compiler	baseline equals Case 1; independent re-derivation	–
Case 1, 47 score transforms	Enumeration	identity equals Case 1	–
Case 1, catalogue × k (k ≥ 2)	MILP only	k = 1 equals enumeration	all MILP vectors
Case 2, all cells and variants	Group-wise merge (Prop 5)	brute force on a sub-catalogue; certification	all vectors

References

[1] S. S. G. Perumal, R. M. Lusby, and J. Larsen, "Electric bus planning & scheduling: A review of related problems and methodologies," *Eur. J. Oper. Res.*, vol. 301, no. 2, pp. 395–413, 2022, doi: 10.1016/j.ejor.2021.10.058.

[2] A. Lajunen, "Lifecycle costs and charging requirements of electric buses with different charging methods," *J. Clean. Prod.*, vol. 172, pp. 56–67, 2018, doi: 10.1016/j.jclepro.2017.10.066.

[3] Z. Gao et al., "Battery capacity and recharging needs for electric buses in city transit service," *Energy*, vol. 122, pp. 588–600, 2017, doi: 10.1016/j.energy.2017.01.101.

[4] H. Basma, C. Mansour, M. Haddad, M. Nemer, and P. Stabat, "Energy consumption and battery sizing for different types of electric bus service," *Energy*, vol. 239, Art. no. 122454, 2022, doi: 10.1016/j.energy.2021.122454.

[5] L. L. Zhang, "Product configuration: A review of the state-of-the-art and future research," *Int. J. Prod. Res.*, vol. 52, no. 21, pp. 6381–6398, 2014, doi: 10.1080/00207543.2014.942012.

[6] A. Felfernig, L. Hotz, C. Bagley, and J. Tiihonen, *Knowledge-Based Configuration: From Research to Business Cases*. Waltham, MA, USA: Morgan Kaufmann, 2014, doi: 10.1016/C2011-0-69705-4.

[7] F. S. Roberts, *Measurement Theory: With Applications to Decisionmaking, Utility, and the Social Sciences*. Cambridge, U.K.: Cambridge Univ. Press, 1984 (first published Reading, MA, USA: Addison-Wesley, 1979), doi: 10.1017/CBO9780511759871.

[8] R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. Cambridge, U.K.: Cambridge Univ. Press, 1993 (first published New York, NY, USA: Wiley, 1976), doi: 10.1017/CBO9781139174084.

[9] Y. Y. Haimes, L. S. Lasdon, and D. A. Wismer, "On a bicriterion formulation of the problems of integrated system identification and system optimization," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-1, no. 3, pp. 296–297, 1971, doi: 10.1109/TSMC.1971.4308298.

[10] G. Mavrotas, "Effective implementation of the ε-constraint method in multi-objective mathematical programming problems," *Appl. Math. Comput.*, vol. 213, no. 2, pp. 455–465, 2009, doi: 10.1016/j.amc.2009.03.037.

[11] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002, doi: 10.1109/4235.996017.

[12] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, 2014, doi: 10.1109/TEVC.2013.2281535.

[13] H. Jain and K. Deb, "An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, Part II: Handling constraints and extending to an adaptive approach," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 602–622, 2014, doi: 10.1109/TEVC.2013.2281534.

[14] J. Vepsäläinen, K. Otto, A. Lajunen, and K. Tammi, "Computationally efficient model for energy demand prediction of electric city bus in varying operating conditions," *Energy*, vol. 169, pp. 433–443, 2019, doi: 10.1016/j.energy.2018.12.064.

[15] A. Lajunen and T. Lipman, "Lifecycle cost assessment and carbon dioxide emissions of diesel, natural gas, hybrid electric, fuel cell hybrid and electric transit buses," *Energy*, vol. 106, pp. 329–342, 2016, doi: 10.1016/j.energy.2016.03.075.

[16] M. Rogge, S. Wollny, and D. U. Sauer, "Fast charging battery buses for the electrification of urban public transport—A feasibility study focusing on charging infrastructure and energy storage requirements," *Energies*, vol. 8, no. 5, pp. 4587–4606, 2015, doi: 10.3390/en8054587.

[17] A. Kunith, R. Mendelevitch, and D. Goehlich, "Electrification of a city bus network—An optimization model for cost-effective placing of charging infrastructure and battery sizing of fast-charging electric bus systems," *Int. J. Sustain. Transp.*, vol. 11, no. 10, pp. 707–720, 2017, doi: 10.1080/15568318.2017.1310962.

[18] M. Rogge, E. van der Hurk, A. Larsen, and D. U. Sauer, "Electric bus fleet size and mix problem with optimization of charging infrastructure," *Appl. Energy*, vol. 211, pp. 282–295, 2018, doi: 10.1016/j.apenergy.2017.11.051.

[19] International Association of Public Transport (UITP), "UITP SORT/E-SORT brochures," Brussels, Belgium. Accessed: Oct. 17, 2026. [Online]. Available: https://www.uitp.org/publications/uitp-sort-e-sort-brochures/

[20] B. Li, L. Chen, Z. Huang, and Y. Zhong, "Product configuration optimization using a multiobjective genetic algorithm," *Int. J. Adv. Manuf. Technol.*, vol. 30, no. 1–2, pp. 20–29, 2006, doi: 10.1007/s00170-005-0035-8.

[21] W. Wei, W. Fan, and Z. Li, "Multi-objective optimization and evaluation method of modular product configuration design scheme," *Int. J. Adv. Manuf. Technol.*, vol. 75, no. 9–12, pp. 1527–1536, 2014, doi: 10.1007/s00170-014-6240-6.

[22] W. Song and F. T. S. Chan, "Multi-objective configuration optimization for product-extension service," *J. Manuf. Syst.*, vol. 37, pp. 113–125, 2015, doi: 10.1016/j.jmsy.2015.09.001.

[23] J. S. Dyer and R. K. Sarin, "Measurable multiattribute value functions," *Oper. Res.*, vol. 27, no. 4, pp. 810–822, 1979, doi: 10.1287/opre.27.4.810.

[24] M. Ehrgott, *Multicriteria Optimization*, 2nd ed. Berlin, Germany: Springer, 2005, doi: 10.1007/3-540-27659-9.

[25] G. Mavrotas and K. Florios, "An improved version of the augmented ε-constraint method (AUGMECON2) for finding the exact Pareto set in multi-objective integer programming problems," *Appl. Math. Comput.*, vol. 219, no. 18, pp. 9652–9669, 2013, doi: 10.1016/j.amc.2013.03.002.

[26] G. Kirlik and S. Sayın, "A new algorithm for generating all nondominated solutions of multiobjective discrete optimization problems," *Eur. J. Oper. Res.*, vol. 232, no. 3, pp. 479–488, 2014, doi: 10.1016/j.ejor.2013.08.001.

[27] P. Halffmann, L. E. Schäfer, K. Dächert, K. Klamroth, and S. Ruzika, "Exact algorithms for multiobjective linear optimization problems with integer variables: A state of the art survey," *J. Multi-Criteria Decis. Anal.*, vol. 29, no. 5–6, pp. 341–363, 2022, doi: 10.1002/mcda.1780.

[28] J. E. Fieldsend, R. M. Everson, and S. Singh, "Using unconstrained elite archives for multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 7, no. 3, pp. 305–323, 2003, doi: 10.1109/TEVC.2003.810733.

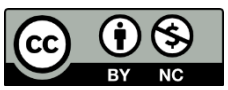
[29] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, 1999, doi: 10.1109/4235.797969.

[30] H. Ishibuchi, H. Masuda, Y. Tanigaki, and Y. Nojima, "Modified distance calculation in generational distance and inverted generational distance," in *Evolutionary Multi-Criterion Optimization (EMO 2015)*, Lecture Notes in Computer Science, vol. 9019. Cham, Switzerland: Springer, 2015, pp. 110–125, doi: 10.1007/978-3-319-15892-1_8.

[31] M. Li and X. Yao, "Quality evaluation of solution sets in multiobjective optimisation: A survey," *ACM Comput. Surv.*, vol. 52, no. 2, pp. 1–38, 2019, doi: 10.1145/3300148.

[32] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," version 13.0.2. Accessed: Oct. 17, 2026. [Online]. Available: https://docs.gurobi.com/

[33] J. Blank and K. Deb, "pymoo: Multi-objective optimization in Python," *IEEE Access*, vol. 8, pp. 89497–89509, 2020, doi: 10.1109/ACCESS.2020.2990567.



All open access articles published in Transactions on Computational Modeling and Intelligent Systems (http://tcmis.org) are distributed under the terms of the CC BY-NC 4.0 license (Creative Commons Attribution Non-Commercial 4.0 International Public License as currently displayed at http://creativecommons.org/licenses/by-nc/4.0/legalcode) which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original work is properly cited.