

Limited Horizon Airline Scheduling using Grey Theory and GRASP framework

Palak Singh[1], Rajkumar S. Pant[2]
[1]Research Intern, Aerospace Engineering, IIT Bombay, Mumbai, India. Telephone: +91-8840588273, E-mail: ps.3030.work@gmail.com

[2]Professor, Aerospace Engineering Department, IIT Bombay, Mumbai, India. (corresponding author). Telephone: +91-9820084502, Telefax: +91(0)22-25720257, E-mail: rpant@iitb.ac.in

Abstract

Overnight cargo airline scheduling requires assignment of aircraft to routes within tight time windows under variable demand, producing a combinatorial problem where greedy local decisions consistently yield suboptimal fleet utilization and poor network coverage. This paper presents a Grey-GRASP-Lookahead scheduling methodology that integrates four mechanisms: Grey Relational Analysis for route scoring, a Greedy Randomized Adaptive Search Procedure with a Restricted Candidate List for stochastic schedule construction, a Committed Demand Discount for fleet-wide coordination, and a one-step look-ahead to mitigate myopic routing decisions. The methodology is validated against published results from Maul [7] on a six-airport Indian cargo benchmark under single-fleet, ten-times demand, and multi-fleet configurations. Without parameter tuning, the scheduler matches or exceeds GA-optimized Grey Clustering on four of five metrics under the Time objective, improving fleet utilization by up to 16.6 percentage points and cargo throughput efficiency by up to 53 percent. Extension to a 14-airport regional passenger network demonstrates generalizability beyond the cargo context.

Keywords: demand-responsive cargo scheduling, Grey Relational Analysis, GRASP algorithm, limited-horizon look-ahead, fleet utilization

1. Introduction

Overnight express cargo operations present a scheduling problem qualitatively different from passenger airline scheduling. Aircraft must complete all revenue flights within a narrow time window, serve variable point-to-point demand across a network where no single route guarantees full payload utilization, and satisfy strict turnaround and maintenance requirements. The combinatorial difficulty scales rapidly with network size: even for modest networks of six to eight airports, conventional exhaustive search is computationally untenable [1], and practical methods must balance solution quality against tractability within the domain of aircraft operation and flight management.

The present work originated as an investigation into limited-horizon look-ahead scheduling but led to a broader methodological revision. Four integrated mechanisms distinguish it from prior work. Grey Relational Analysis with relational coefficients replaces Grey Clustering, measuring each route's proximity to an ideal reference sequence across four indices. A Greedy Randomized Adaptive Search Procedure with a Restricted Candidate List replaces deterministic greedy selection, enabling stochastic schedule exploration across repeated runs. A Committed Demand Discount introduces fleet-wide coordination by discounting each route's effective remaining demand by cargo already allocated by other aircraft in the current scheduling attempt. A one-step look-ahead evaluates the fleet-competitive state following each candidate assignment before committing, reducing myopic routing decisions. The methodology is evaluated on three networks of increasing complexity and validated against Maul's [7] published results on a six-airport Indian cargo benchmark.

2. Literature survey

Grey system theory, introduced by Deng [2], is a mathematical framework for decision-making under partial information, applicable to domains where statistical distributions are unavailable or unreliable. Its most scheduling-relevant branch is Grey Clustering with whitening weight functions, in which candidate options are evaluated against a set of indices and assigned to decision categories by membership grade. Hsu and Wen [3] first applied this to airline network design, proposing route selection based on intermediate stops, route length index, and traffic concentration. Their formulation identified which routes warranted service and at what frequency but did not generate departure times or individual aircraft assignments, motivating the subsequent integration of Grey Theory into a sequential schedule generator capable of producing implementable daily flight schedules.

Singh and Pant [4, 5] reformulated the problem as selecting the next destination airport from the aircraft's current location at each scheduling step. Four Grey indices were defined: a Demand Index measuring route demand as a fraction of total outgoing demand from the source, a Cost Index measuring route operating cost as a fraction of total outgoing cost, a Time Index measuring fly time as a fraction of total outgoing fly time where fly time includes block time, slot waiting, and turnaround, and a planner-specified Route Priority Index in the range [-1, +1]. These indices are whitened using triangular functions with reject and select thresholds at the 20th and 80th percentiles of each index's observed range [4]. Aircraft are processed serially by next available take-off time, with the Grey Clustering decision weight selecting the highest-scoring feasible destination at each step. Grey index weights serve as design variables for the GADO optimizer [6], which iterates over 15 or more generations to maximize a specified objective function. The framework was demonstrated on three and five airport cargo networks. Singh identified three limitations of direct relevance to the present work: no mechanism for one aircraft's assignment to account for demand committed by others in the same scheduling run, no anticipation of downstream routing consequences from the greedy decision rule, and uniform fixed weights applied regardless of changing network state [4].

Maul [7] extended this framework to a six-airport Indian network with a heterogeneous fleet of Boeing 737-200 and Boeing 757-200 aircraft, introducing aircraft-type-specific Break-Even Load Factor thresholds and 10-minute slot windows at each airport to increase operational fidelity. The Grey Clustering architecture, serial scheduling loop, and GADO optimizer remained structurally unchanged. Two fundamental limitations persisted: the absence of inter-aircraft coordination allowed multiple aircraft to independently converge on the same high-scoring routes, producing demand pile-up while leaving viable alternatives underserved, and the greedy local decision rule selected the best-scoring next route without evaluating downstream routing consequences, producing behaviour that was locally rational but potentially globally suboptimal as the operating window narrowed. These observations directly motivated the extensions described in the present work.

3. Methodology

The scheduler presented in this paper integrates four mechanisms: Grey Relational Analysis for route scoring, a Greedy Randomized Adaptive Search Procedure for stochastic candidate selection, a Committed Demand Discount for fleet coordination, and a one-step look-ahead for myopia prevention. This section describes each mechanism, the constraint set that governs all assignment decisions, the Quality-of-Service metric used for evaluation, and the objective weight configurations applied throughout the experiments.

3.1 Grey Relational Analysis Route Scoring

Grey Relational Analysis, introduced by Deng [2] and extended to airline network problems by Hsu and Wen [3], evaluates each candidate route against an ideal reference sequence and produces a scalar score reflecting its relative desirability under a given objective. The scoring is built from a decision matrix whose columns correspond to four route indices.

1. **Block time (BT)** is the total door-to-door duration for a given origin and destination pair, computed as the great circle distance divided by cruise speed plus turnaround time

https://doi.org/10.71945/icas2026_0816

expressed in hours. Shorter block times receive higher normalized scores, so BT is treated as a cost index.

2. **Demand Intensity (DI)** is route demand in tons divided by block time in hours. This index, absent from the Grey Clustering formulations of Singh and Pant [4, 5] and Maul [7], captures throughput efficiency on a per-hour basis and is treated as a benefit index.
3. **Route Risk (Risk)** is a composite of three sub-components: demand volatility relative to the network mean, distance exposure relative to the longest route in the network, and a connectivity penalty that down-scores routes linking airports with low aggregate outgoing demand. It is treated as a cost index.
4. **Raw Demand (RawD)** is the total daily cargo demand on the route in tons, treated as a benefit index. In the multi-fleet configuration, a fifth index is added: **Load Factor (LF)**, defined as route demand divided by the product of block time and the maximum fleet payload, capped at unity, which measures how effectively the highest-capacity aircraft in the fleet would be utilized on that route.

The decision matrix X has dimensions *m* by *n*, where *m* is the number of candidate routes and *n* is the number of indices. Each column is normalized independently. For benefit indices: where: x^*_{ij} - normalized value of criterion j for alternative I, x_{ij} – original performance value, $\max_j$ – maximum value of criterion j, $\min_j$ – minimum value of criterion j, $\gamma(x^*_{ij})$ – grey relational coefficient, Γ_i – grey relational grade for alternative I, w_j - weighting coefficient for criterion j, ρ - distinguishing coefficient (typically 0.5), Δ_{min} – minimum deviation value , Δ_{max} – maximum deviation value

$$x^*_{ij} = \frac{x_{ij} - \min_j}{\max_j - \min_j} \tag{1}$$

For cost indices the direction is inverted:

$$x^*_{ij} = \frac{\max_j - x_{ij}}{\max_j - \min_j} \tag{2}$$

The Grey Relational coefficient for route *i* on index *j* measures proximity to the ideal reference, which is unity in normalized space:

$$\gamma(x^*_{ij}) = \frac{\Delta_{\min} + \rho\Delta_{\max}}{| 1 - x^*_{ij} | + \rho\Delta_{\max}} \tag{3}$$

where Δ_{min} and Δ_{max} are the global minimum and maximum absolute deviations from the reference across all routes and all indices, and $\rho = 0.5$ is the distinguishing coefficient. The weighted Grey Relational Grade for route *i* is then:

$$\Gamma_i = \sum_j w_j \cdot \gamma(x^*_{ij}) \tag{4}$$

where *wj* are objective-dependent weights that sum to unity. Routes with a higher grade Γi are preferred under the active objective configuration.

3.2 GRASP with Restricted Candidate List

The scheduler adopts the Greedy Randomized Adaptive Search Procedure (GRASP) framework [8], replacing deterministic greedy selection with a controlled stochastic selection from a Restricted Candidate List (RCL). At each assignment step, all feasible routes from the current aircraft position are scored using the GRA grade adjusted by the Committed Demand Discount described in Section 3.3. The candidates are sorted in descending order of this adjusted score, and the top fraction governed by parameter α forms the RCL. One route is then chosen from the RCL by weighted random sampling, with selection probability proportional to the combined score defined in Section 3.4.

The parameter α is sampled uniformly from the interval [0.25, 0.60] once at the start of each independent run, rather than being fixed across all runs. As a result, each of the 50 runs per objective mode explores a different level of selectivity: smaller values of α restrict the list to a narrow elite, while larger values admit a broader and more diverse candidate pool. This produces a

https://doi.org/10.71945/icas2026_0816

distribution of feasible schedules from which the best can be identified. The GRASP expansion loop continues assigning flights to each aircraft in turn, sorted by earliest available time, until no aircraft can be assigned any further flight with a positive combined score.

3.3 Committed Demand Discount

When multiple aircraft are scheduled concurrently, a naive scoring approach allows several aircraft to independently identify the same high-demand route as optimal. This demand pile-up leaves low-demand or distant routes consistently unserved while wasting capacity on routes that are already saturated. The Committed Demand Discount (CDD) corrects this by discounting, from the scoring demand of any route, the cargo that has already been committed to that route by other aircraft earlier in the same scheduling iteration.

If aircraft *k* is being considered for assignment and route *r* already carries committed cargo C_r from other aircraft in the current schedule, the adjusted demand used for scoring is: where, $\dot{D}_r$ = residual unmet demand on route r in tons, D_r = total demand on route r, C_r = carried cargo/passenger demand on route r

$$\bar{D}_r = \max(0,\ D_r - C_r) \tag{5}$$

The demand factor that scales the GRA grade in Step A of the GRASP expansion loop is $\dot{D}_r$ / D_r, which decays toward zero as committed cargo approaches the full route demand. Any route whose adjusted demand reaches zero is excluded from the candidate set entirely for that aircraft. This concurrent awareness has no direct equivalent in the sequential scheduling architectures of Singh and Pant [5] or Maul [7].

3.4 One-Step Look-Ahead

Greedy schedulers are susceptible to myopic routing decisions, where a locally attractive assignment leaves the aircraft stranded at an airport with no viable onward flights before the end of the scheduling window. The one-step look-ahead evaluates the downstream consequences of each candidate assignment before committing to it.

For a candidate route *r* under consideration for aircraft *k*, the procedure is as follows. A copy of the aircraft state and remaining demand are created without modifying the originals. The hypothetical assignment of *r* is executed on these copies. The top-*k* routes available to the aircraft from its resulting position are then scored using GRA grades with CDD applied, and their average constitutes the forward-looking score $\bar{\Gamma}_{\text{future}}$. The combined score used for RCL construction is: where, S_r = look-ahead scheduling score for route r, Γ_r = immediate Grey relational grade of route r, $\bar{\Gamma}_{\text{future}}$ = discounted estimate of future routing reward, γ = look-ahead discount factor

$$S_r = \Gamma_r + \gamma\, \bar{\Gamma}_{\text{future}} \tag{6}$$

where γ = 0.3 is the look-ahead weight, and *k* = 3 top candidates are used to compute the future average. To keep computation tractable, the look-ahead is applied only to the top *N* = 10 candidates ranked by immediate score before any combined scoring takes place. The procedure does not recurse; it evaluates exactly one step forward.

3.5 Constraint Set

Every assignment decision is governed by the following hard constraints, which correspond to the operational rules used in Maul [7].

C1.1 Operating window. Every flight must land on or before DAY_END = 615 minutes. A candidate flight is rejected if its projected landing time exceeds this limit.

C1.2 Turnaround time. A minimum ground time of 45 minutes must elapse between the landing of one flight and the departure of the next for the same aircraft. This is enforced structurally: after each assignment the aircraft available time is set to landing time plus 45 minutes, so no separate check is required at the candidate stage.

C1.3 Daily block time cap. The sum of block times across all flights assigned to a single

https://doi.org/10.71945/icas2026_0816

aircraft cannot exceed the length of the scheduling window, 615 minutes.

C2.1 Location continuity. An aircraft can only depart from the airport at which it currently sits. Every candidate flight whose origin does not match the aircraft's current location is excluded.

C2.3 Range. Route distance must not exceed the maximum ferry range of the aircraft type. For the single-fleet configuration this constraint is non-binding, as all Indian domestic routes fall within the operating range. For the multi-fleet configuration, range limits are applied per aircraft type.

No base-end location requirement is imposed. Aircraft are free to end the scheduling day at any airport. This reflects overnight cargo operations, where repositioning for the next cycle is handled as a separate planning problem, and it matches the free-end positioning used in Maul [7].

3.6 Quality of Service Metric

The Quality of Service (QOS) metric quantifies how evenly the scheduler distributes service across routes relative to what each route demands. A value of zero would indicate that every route is served in exact proportion to its demand; higher values reflect an increasingly uneven distribution, where high-demand routes receive multiple services while thin routes receive none.

The metric is expressed as the root mean square of the gap between required and actual service frequency on each route. Frequency is measured in aircraft loads by dividing cargo volume by a reference payload capacity. For the single-fleet configuration, where all aircraft carry the same payload P = 15 tons, the required and actual frequencies per route are D_r / P and S_r / P respectively. The payload cancels in the squared difference, giving: where, QOS – Quality of Service metric, N – total number of evaluated routes, D_r – demand on route r, S_r – served demand on route r, P – reference payload capacity, C_{max} – maximum aircraft carrying capacity, (D_r – S_r) – unmet demand on route r

$$QOS = \sqrt{\frac{1}{N}\sum_{r}\frac{(D_r - S_r)^2}{P^2}} \tag{7}$$

For the multi-fleet configuration, aircraft payloads differ across types and no single aircraft payload applies uniformly. Following the convention of Maul [7], the maximum fleet payload C_{max} = 39 tons, corresponding to the Boeing 757-200, is used as the common reference denominator. This normalizes frequency gaps relative to the highest-capacity aircraft in the fleet and ensures that QOS values are internally consistent across all routes regardless of which aircraft type served them:

$$QOS = \sqrt{\frac{1}{N}\sum_{r}\frac{(D_r - S_r)^2}{C_{\max}}} \tag{8}$$

In both expressions, N is the number of routes with positive demand, Dr is the original demand on route r in tons, and Sr is the cargo actually served. Lower values are better. Equation (8) was adopted specifically to compare against Maul's [7] multi fleet results, where the same denominator convention applies, ensuring the comparison is like-for-like.

3.7 Algorithm Summary and Objective Configurations

The scheduler operates in two phases. Phase 1 assigns the first flight to each aircraft from its starting base using GRA grades adjusted by CDD and selects deterministically, with no randomization. Phase 2 is the GRASP expansion loop, which iterates over all aircraft sorted by earliest available time and continues assigning flights until no aircraft can accept any further assignment with a positive combined score. The complete procedure is illustrated in Figure 1- Phase 1 assigns the first flight from each aircraft's starting base. Phase 2 is the GRASP expansion loop, which continues until no aircraft records progress in a full iteration. (S_i in the figure corresponds to S_r in equation (6).

https://doi.org/10.71945/icas2026_0816

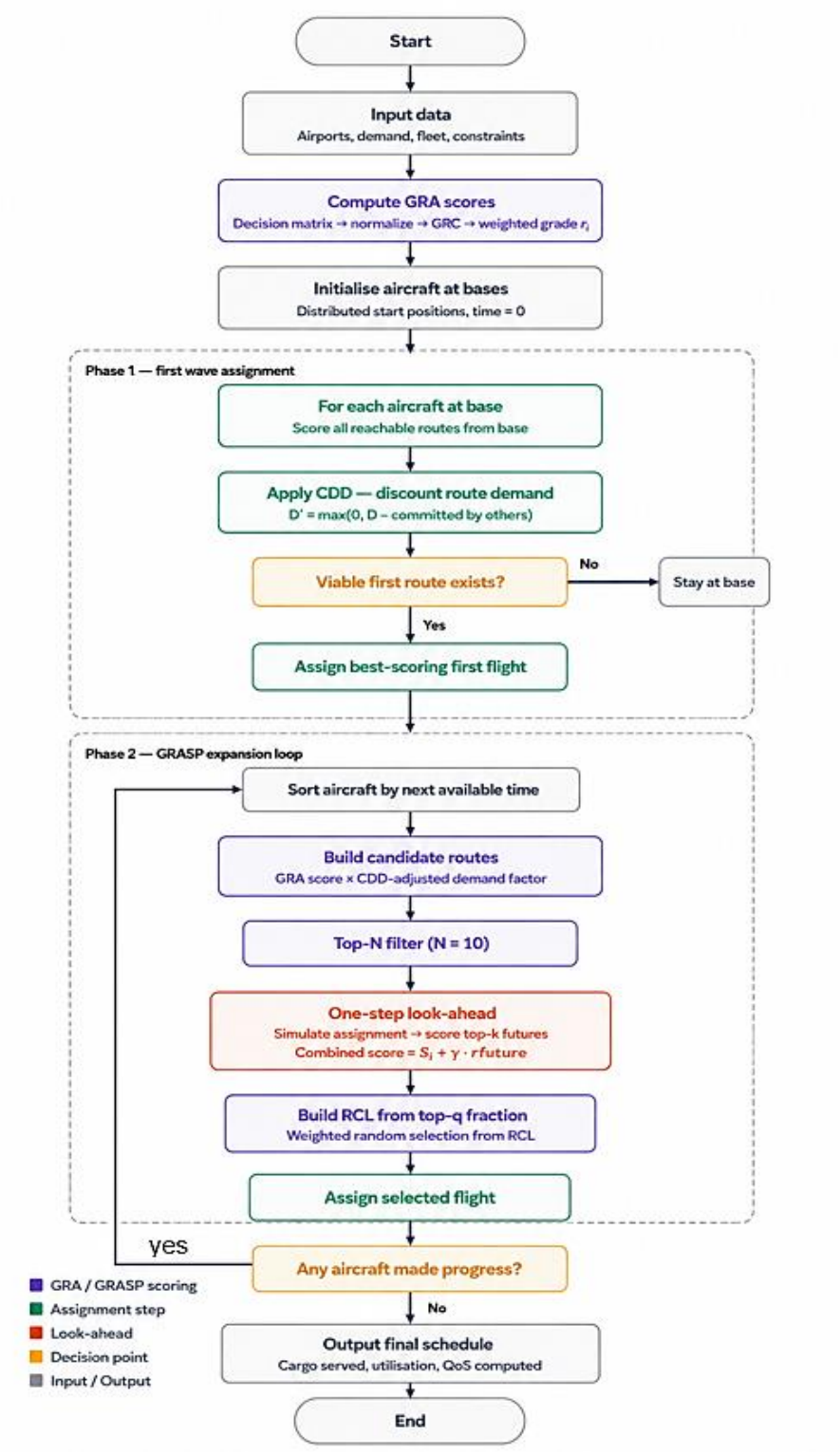


Figure 1: Limited horizon grey GRASP look-ahead scheduler algorithm

Two objective configurations are used throughout the experiments, corresponding to Maul's [7] Time and Utilization objectives. Table 1 lists the weight assigned to each index under each mode.

https://doi.org/10.71945/icas2026_0816

Objective Mode	BT	DI	Risk	RawD	LF[1]
Max Utilization	0.40	0.25	0.05	0.15	0.15
Max Time	0.08	0.27	0.10	0.40	0.15

Table 1: Objective weight configurations.

BT = block time; DI = demand intensity; Risk = composite risk index; RawD = raw route demand; LF = load factor index. [1]LF is included only in multi-fleet experiments. In single-fleet runs the four-index weight set is used with weights renormalized accordingly (Max Utilization: BT 0.50, DI 0.30, Risk 0.05, RawD 0.15; Max Time: BT 0.10, DI 0.35, Risk 0.10, RawD 0.45).

The Max Utilization configuration places the largest weight on block time, which drives the scheduler toward longer routes that keep aircraft airborne for a greater fraction of the available window. The Max Time configuration shifts weight heavily onto raw demand, directing aircraft toward the highest-volume routes regardless of distance and thereby maximizing total cargo throughput at some cost to fleet utilization. Neither weight set was tuned by a genetic algorithm or any external optimizer; all values are assigned manually and held constant across all 50 runs per objective mode. This constitutes a deliberate pre-optimization baseline, and GA-based weight tuning is identified as the primary direction for further work in Section 8.

4. Relationship to Prior Work and Scope of Comparison

The results presented in Section 5 compare the Grey-GRASP-Lookahead scheduler directly against the schedules reported by Maul [7] on the same six-airport Indian cargo benchmark network. This comparison is informative but not a controlled single-variable experiment. Several architectural differences exist between the two methods that cannot be resolved without fundamentally altering one or the other. These differences are stated here explicitly so that the results in Section 5 can be interpreted correctly.

1. **Grey Theory formulation**. Maul [7] applies Grey Clustering with triangular whitening weight functions, partitioning candidate routes into decision categories by membership grade. The present method applies Grey Relational Analysis with Grey Relational Coefficients, measuring each route's proximity to an ideal reference sequence in normalized index space. Both are established Grey system theory applications [2] that produce different numerical scores from identical input data. The comparison is therefore between two Grey-based schedulers, not a single-variable experiment.
2. **Weight optimization.** Maul's [7] Grey index weights are tuned by a Genetic Algorithm over 15 or more generations on the benchmark dataset, with the weight vector passed to GADO [6] as the design variable. The weights in Table 1 of the present work are manually configured and held constant across all runs. The results in Section 5 therefore represent what the scheduling logic achieves without any parameter tuning. Weight optimization is identified as the primary direction for further work in Section 8.
3. **Objective function components.** Maul's [7] objectives incorporate route-level operating cost data and slot waiting times, neither of which was available for replication. Their absence affects the cargo result under the Utilization objective, where Maul's [7] cost-aware Grey scorer produces more aggressive trunk-route concentration than the present configuration can replicate.
4. **Operational constraints.** Two constraints in Maul's [7] formulation are absent from the validation configuration. First, Maul's [7] Break-Even Load Factor gate prevents assignment to any route where expected load factor falls below the aircraft type's break-even threshold. Without a cost matrix this constraint has no economic justification in the present model and was excluded. Second, Maul [7] models ten-minute slot windows at each airport, with waiting times included in his operating time metric. The present model measures operating time as pure block time only, making operating time figures structurally incomparable between the two methods.

https://doi.org/10.71945/icas2026_0816

5. **Validation configuration versus full implementation.** The scheduler was configured for validation to match Maul's [7] architectural assumptions as closely as possible. Features specific to the full implementation were excluded to avoid penalizing the present method for constraints Maul's [7] scheduler does not impose. The comparison in Section 5 is between two schedulers operating under the same free-end positioning and coverage-as-output assumptions.

5. Validation Against Maul's [7] Benchmark

The scheduler is validated on the six-airport Indian cargo network used by Maul [7], comprising the airports Chennai (MAA), Bangalore (BLR), Bombay (BOM), Delhi (DEL), Kolkata (CCU), and Hyderabad (HYD). The network parameters match Maul's [7] specification exactly: a 615-minute operating window, a 45-minute turnaround time at all airports, and pairwise demand data taken directly from his published dataset. Three experimental configurations are evaluated: a single homogeneous fleet under original demand, the same fleet under ten times the original demand, and a heterogeneous two-type fleet under original demand. In all cases, 50 independent runs are completed per objective mode, with the GRASP alpha parameter sampled freshly for each run. Results are compared against the three individual schedule outcomes Maul [7] reports per objective for each configuration.

5.1 Single Fleet, Original Demand

The single-fleet configuration uses five aircraft each with a payload capacity of 15 tonnes, matching Maul's [7] specification. Total network demand is 159.5 tonnes. Table 2 presents mean results across 50 runs per objective mode alongside Maul's [7] best reported schedule per objective.

Metric	Maul Best (Max Util)	Ours Mean ± SD (Max Util)	Maul Best (Max Time)	Ours Mean ± SD (Max Time)
Cargo Served (t)	123.0	112.07 ± 3.71	120.0	122.17 ± 3.60
Operating Time (min)	2430	1978.2 ± 88.4	2525	2036.4 ± 48.7
QOS	0.308	0.23 ± 0.02	0.296	0.17 ± 0.01
Fleet Utilization	0.571	0.643 ± 0.029	0.555	0.662 ± 0.016
Cargo per Hour (t/hr)	3.04	3.40 ± 0.17	2.85	3.60 ± 0.15

Table 2: Single fleet results under original demand - Grey-GRASP-Lookahead vs Maul [7]

Under the Max Time objective the scheduler serves 122.17 tonnes on average against Maul's [7] best of 120.0 tonnes, a gain of 1.8%, while simultaneously achieving higher fleet utilisation (66.2% versus 55.5%) and substantially better QOS (0.17 versus 0.296). Cargo throughput efficiency reaches 3.60 tonnes per hour against Maul's [7] 2.85, a difference of 26.3%. These advantages are attributable to two mechanisms working together. The distributed base initialization places aircraft at the four highest outbound-demand airports from the start of the window, eliminating the positioning flights that a hub-based architecture requires before revenue service begins. The GRASP expansion loop then continues assigning flights until the window is exhausted, whereas Maul's [7] scheduler terminates an aircraft when no route above the BELF threshold remains reachable.

Under the Max Utilisation objective, the scheduler serves 112.07 tonnes on average against Maul's [7] best of 123.0 tonnes, a shortfall of 8.9%. As established in Section 4, this gap is attributable to two compounding differences rather than a scheduling logic failure. Maul's [7] GA-tuned weights under the Utilisation objective steer his Grey scorer more aggressively toward high-volume trunk routes than the manually configured weights in Table 1 can replicate without optimisation. Additionally, Grey Clustering with whitening functions creates sharper distinctions between high and low demand route categories than Grey Relational Coefficients, which evaluate all routes against the same ideal reference uniformly. The scheduler nonetheless outperforms Maul [7] on four of five

https://doi.org/10.71945/icas2026_0816

metrics under this objective: QOS is 25% better, fleet utilisation is 7.2 percentage points higher, and cargo per hour is 11.8% greater.

The operating time gap in both objectives reflects the metric definition difference stated in Section 4. Maul's [7] operating time accumulates block time, turnaround time, and slot waiting times at every airport. The present metric sums block minutes only. The true flying time in Maul's [7] schedules is therefore lower than his reported figures, and the gap is not interpretable as a difference in scheduling intensity.

5.2 Single Fleet, 10× Demand

The 10× demand scenario scales all origin-destination demand values by a factor of ten, producing a total network demand of 1,595 tonnes against a fleet capacity of 75 tonnes per day. Neither scheduler can serve more than approximately 19% of demand under these conditions. The comparison shifts from coverage breadth to scheduling efficiency under saturation: which method extracts the most cargo from a severely supply-constrained fleet. Table 3 presents results.

Metric	Maul Best (Max Util)	Ours Mean ± SD (Max Util)	Maul Best (Max Time)	Ours Mean ± SD (Max Time)
Cargo Served (t)	240.0	291.81 ± 8.52	235.7	302.16 ± 7.65
Operating Time (min)	2470	2073.0 ± 46.6	2665	2227.5 ± 61.6
QOS	0.074	—	0.220	—
Fleet Utilization	0.581	0.674 ± 0.015	0.558	0.724 ± 0.020
Cargo per Hour (t/hr)	5.83	8.45 ± 0.19	5.31	8.14 ± 0.25

Table 3: Single fleet results under 10× demand – Grey-GRASP-Lookahead vs Maul [7]

The scheduler serves 291.8 tonnes under Max Utilisation and 302.2 tonnes under Max Time, exceeding Maul's [7] corresponding best results by 22% and 28% respectively. Fleet utilisation reaches 67.4% and 72.4% against Maul's [7] 58.1% and 55.8%. Cargo throughput efficiency reaches 8.45 and 8.14 tonnes per hour against Maul's [7] 5.83 and 5.31, differences of 45% and 53%.

These gains are driven by the interaction of distributed initialization and the look-ahead mechanism under saturation conditions. When demand is uniformly high across all routes, the CDD prevents aircraft from converging on the same corridors, and the look-ahead steers each aircraft away from destinations from which no productive onward routing is available within the remaining window. The result is that aircraft cycle repeatedly on short high-demand corridors such as BOM-DEL and BOM-BLR, which can be completed multiple times within 615 minutes, rather than committing to single long legs that exhaust the window after one or two flights.

QOS values are not compared in this scenario. At 10× demand, required service frequencies per route reach three to thirteen flights per day while the five-aircraft fleet can physically complete one to four flights per route. The resulting frequency gap is determined by fleet capacity rather than scheduling quality, making QOS a measure of supply constraint rather than schedule design. Both methods face the same structural limitation and the metric does not discriminate between them at this demand level.

5.3 Multi – fleet application

The multi-fleet configuration uses three Boeing 737-200 aircraft at 16.78 tonnes payload and two Boeing 757-200 aircraft at 39.0 tonnes payload, matching Maul's [7] multi fleet specification. Total fleet capacity per day is 128.3 tonnes against a network demand of 159.5 tonnes. Rather than reporting mean results across 50 runs, Table 4 presents a direct best-run comparison, selecting the single best schedule from the 50 runs by the relevant objective metric and comparing it against Maul's [7] single reported schedule per objective. This places both methods on equal footing as individual schedule outcomes.

https://doi.org/10.71945/icas2026_0816

Metric	Maul's (Max Util)	Our Best (Max Util)	Δ	Maul's (Max Time)	Our Best (Max Time)	Δ
Cargo Served (t)	127.2	133.97	+5.3%	140.5	142.71	+1.6%
Operating Time (min)	2545	2360	−7.3%	2085	2160	+3.6%
QOS	0.25	0.06	−76.7%	0.25	0.03	−86.1%
Fleet Utilization	0.600	0.767	+27.9%	0.490	0.702	+43.4%
Cargo per Hour (t/hr)	3.00	3.41	+13.5%	4.04	3.96	−1.9%

Table 4: Multi-fleet best-run comparison- Grey-GRASP-Lookahead vs Maul [7]

Under the Max Utilisation objective, the best run serves 133.97 tonnes against Maul's [7] 127.2 tonnes, an improvement of 5.3%, while achieving fleet utilisation of 76.7% against his 60.0% and QOS of 0.06 against his 0.25. Under the Max Time objective the best run serves 142.71 tonnes against Maul's [7] 140.5 tonnes, a gain of 1.6%, with fleet utilisation of 70.2% against his 49.0% and QOS of 0.03 against his 0.25. Cargo per hour under Max Time is 3.96 tonnes per hour against Maul's [7] 4.04, a difference of 1.9% that falls within normal run-to-run variation and is treated as a statistical tie. The flight sequence for the best Max Time run is shown in Figure 2.

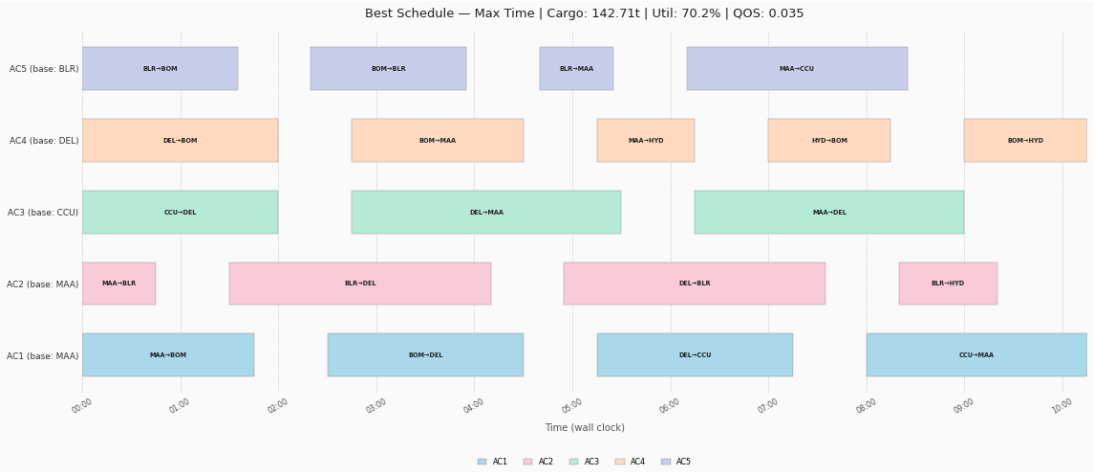


Figure 2: Best schedule produced under the Max Time objective.

The most significant finding in the multi-fleet comparison is the QOS advantage. Maul's [7] QOS of 0.25 is identical across both objectives because his BELF gate systematically prevents the Boeing 757-200 from serving any route where expected load factor falls below break-even. Routes such as HYD-MAA (0.405 tonnes), HYD-CCU (0.500 tonnes), and CCU-HYD (0.325 tonnes) fall below the 757's BELF threshold and receive zero service in Maul's [7] schedule. These routes contribute their maximum possible frequency gap to the QOS calculation, inflating the metric regardless of how well the remaining routes are served. Without a BELF gate, the present scheduler assigns the 757 to these thin routes when no higher-demand assignment is available, reducing the frequency gap on every served route and producing consistently lower QOS across all runs. This reveals a fundamental tension in cargo scheduling between payload concentration and network coverage: BELF enforcement maximises per-flight efficiency but abandons thin routes entirely, while unconstrained scheduling sacrifices marginal per-flight efficiency in exchange for broader and more uniform network service. Which strategy better serves an airline's operational objectives depends on its network model and commercial priorities. The route networks for the two best-run schedules are shown in Figures 3 and 4; multi-fleet configuration (Data 3). Blue edges: Boeing 737-200 (16.78 t); orange edges: Boeing 757-200 (39.0 t). Line thickness proportional to cargo carried. 18 flights, 18 origin destination pairs served, illustrating the allocation of 737-200 and 757-200 capacity across the

https://doi.org/10.71945/icas2026_0816

six-airport network.

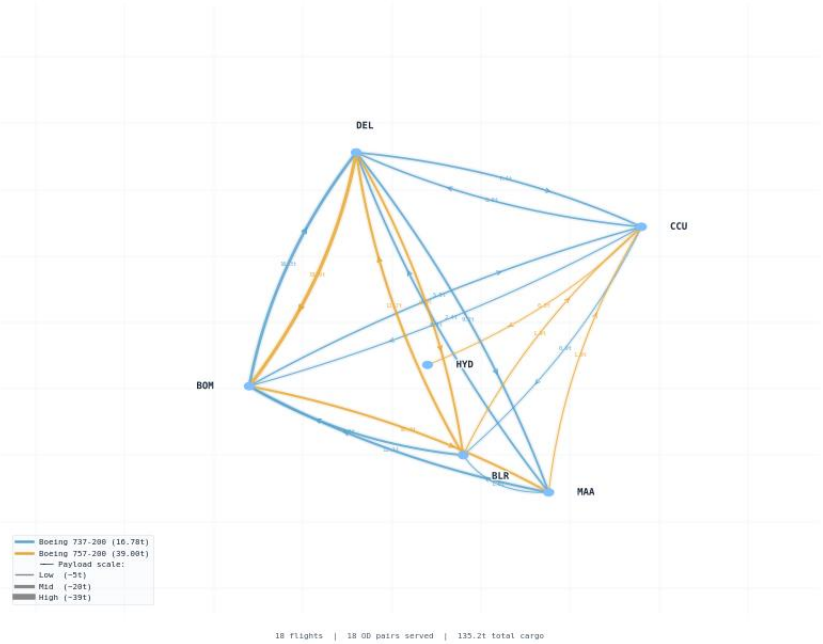


Figure 3: Route network for the best-run schedule under the Max Utilization objective.

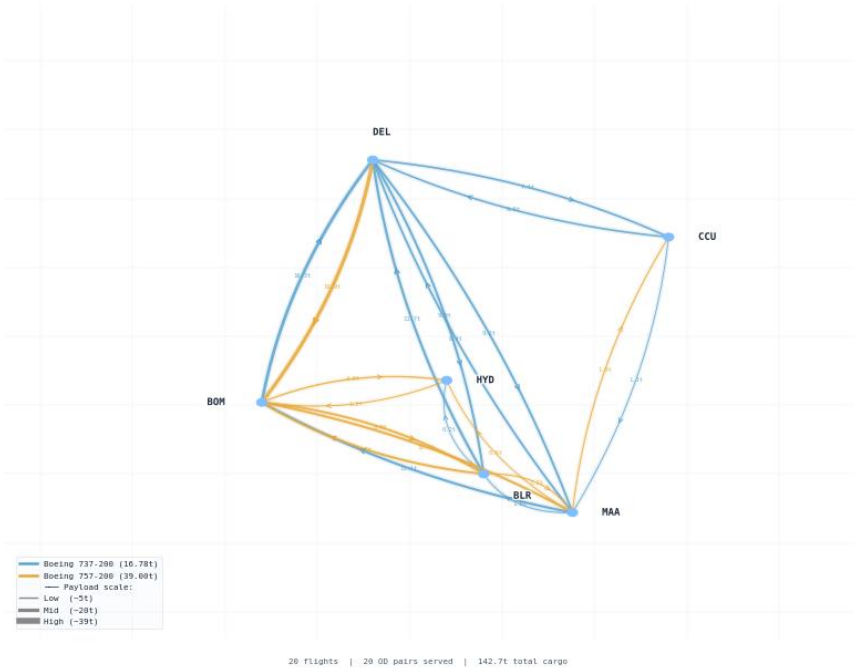


Figure 4: Route network for the best-run schedule under the Max Time objective.

The 100% feasibility rate across all 100 runs in the multi-fleet experiments, with zero infeasible schedules rejected, is noted as a robustness finding. Maul [7] presents one schedule per objective with no information on rejection rate or schedule diversity. The standard deviations reported in the mean-result summary (cargo SD 5.16 tonnes under Max Utilisation, 3.44 tonnes under Max Time, representing coefficients of variation of 3.9% and 2.5% respectively) confirm that the best-run results in Table 4 are not statistical outliers but rather the upper end of a consistently high-quality distribution.

6. Application Cases

https://doi.org/10.71945/icas2026_0816

The validation experiments in Section 5 establish that the Grey-GRASP-Lookahead scheduler produces competitive results on a known benchmark under controlled conditions. The two application cases presented here serve a different purpose. They trace the development of the methodology from its conceptual origin to a substantially more complex operational setting, and they identify the conditions under which the scheduler generalises beyond the six-airport cargo context of the benchmark.

6.1 Proof of Concept — Four Airport Cargo Network

The look-ahead mechanism was first tested on a minimal synthetic network of four airports, labelled A through D, with fixed pairwise demand and deterministic flight times. A single aircraft operated under a total available window of 180 minutes. The scheduling objective was to maximise cumulative reward, defined as a weighted combination of cargo transported and flight time consumed. This configuration was chosen specifically to isolate the contribution of the look-ahead mechanism, since with one aircraft and four airports the CDD and GRASP mechanisms have no meaningful role. The only variable is whether the aircraft evaluates downstream consequences before committing to each leg.

Two strategies were compared. The greedy strategy selected A→B as the first leg on the basis of its high immediate reward. This choice constrained all subsequent routing: from B, the only reachable continuation within the remaining window was C, producing the sequence A→B→C with a cumulative reward of 13.50 and 90 minutes of operating time consumed. Figures 5 and 6 show the resulting Gantt charts; Aircraft 3 and Aircraft 4 terminate early in the greedy schedule due to the myopic first-leg choice. The look-ahead strategy evaluated the downstream state after each candidate first leg before committing. Although A→C offered a lower immediate reward than A→B, the one-step simulation revealed that C connected to D within the remaining window, producing the sequence A→C→D with a higher cumulative reward and 105 minutes of operating time consumed.

This result confirms the core motivation for the look-ahead mechanism in the simplest possible setting. A purely greedy rule selects the locally best option and closes off superior continuations. A one-step forward evaluation costs one additional scoring computation per candidate and recovers a globally better sequence. Figure 7 compares the three scheduler variants: the Grey Hybrid configuration achieves 92.9% demand satisfaction against 50.4% for Strict GRASP and 44.0% for GRASP alone, at fleet utilization of 70.2%. The four-airport case provided the verified conceptual foundation on which the full multi-aircraft scheduler was constructed.

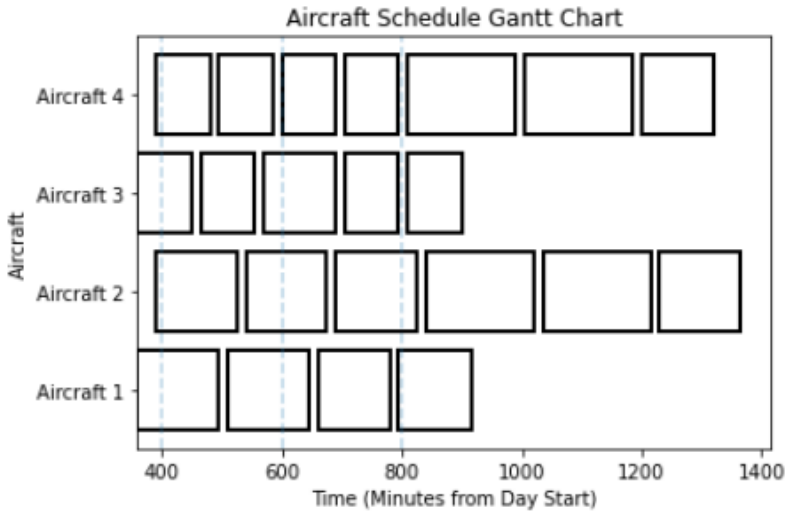


Figure 5: Greedy scheduler output for the four-airport proof-of-concept network without look-ahead

https://doi.org/10.71945/icas2026_0816

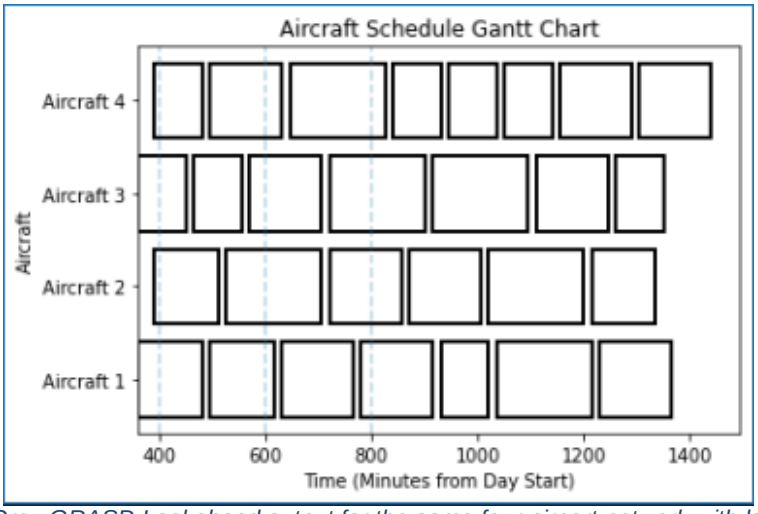


Figure 6: Grey-GRASP-Lookahead output for the same four-airport network with look- ahead

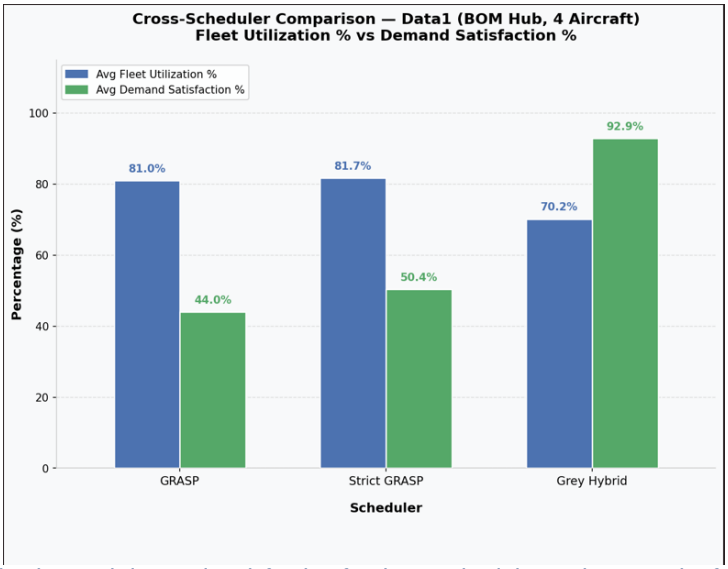


Figure 7: Fleet utilization and demand satisfaction for three scheduler variants on the four-airport network.

6.2 Extension to a Regional Passenger Network

This configuration differs from the validation setup in Section 5 in four respects: hub return is enforced as a hard constraint, a Break-Even Load Factor soft penalty discourages sub-break-even assignments in route scoring, aircraft type matching logic directs high-demand routes toward ERJ175 and low-demand routes toward ERJ145 through a two-threshold classification, and a mandatory origin-destination coverage phase ensures every feasible route receives at least one service attempt before the GRASP expansion loop begins. Their inclusion demonstrates that the Grey-GRASP-Lookahead architecture accommodates constraint additions as modular layers without requiring modification to the core scoring, CDD, or look-ahead logic.

The initial application to the full 182-pair demand matrix produced 37 percent demand satisfaction despite fleet utilisation of 70 to 75 percent, because the matrix encoded total multi-frequency daily demand rather than single-wave demand. Dividing demand by a frequency factor of five resolved this, demand satisfaction rose to 79.79 percent, fleet utilisation to 73.71 percent, and route coverage breadth to 131 of 182 origin-destination pairs (71.98 percent), with all 50 runs producing feasible schedules and zero rejections. All unserved pairs carried scaled demand below 20 passengers; no route above 60 passengers went unserved in any run. A representative schedule is shown in Figure 8.

https://doi.org/10.71945/icas2026_0816

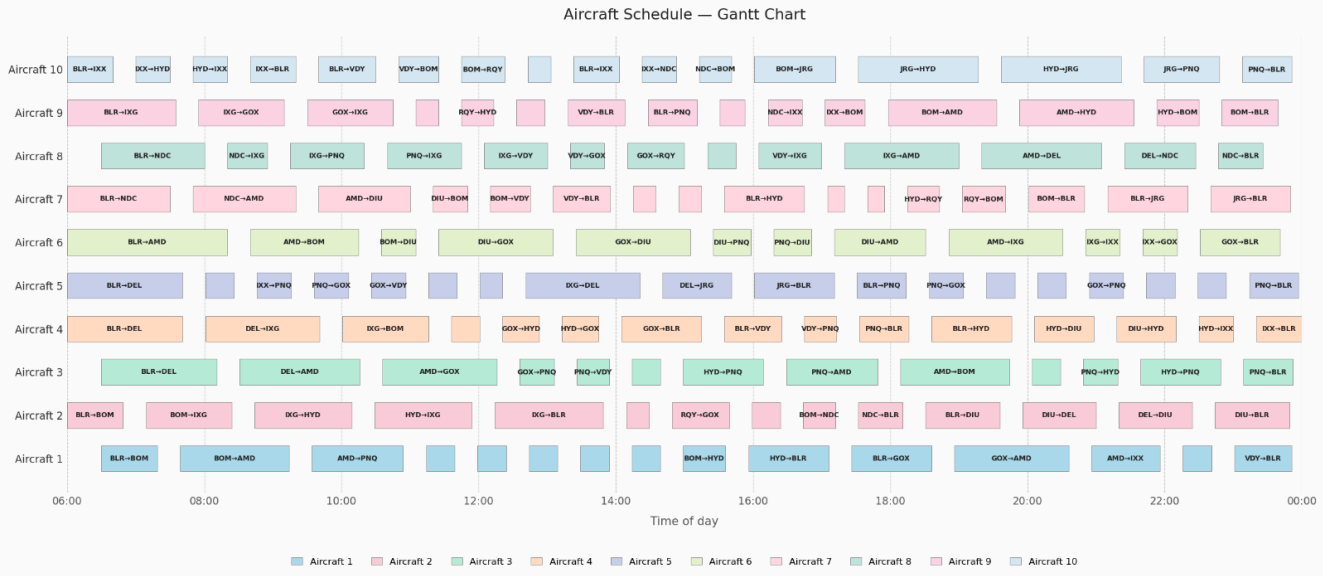


Figure 8: Sample schedule, 14-airport regional passenger network (Data 2).

Three behaviors emerged that were not observable at the benchmark scale, as illustrated in Figure 9, which shows the full network with hub BLR and line thickness proportional to passenger demand. Emergent fleet specialisation appeared without any hard assignment rule: larger aircraft self-selected onto trunk routes and smaller aircraft onto thinner connections, arising from the combined effect of Grey Relational Analysis scoring, Break-Even Load Factor penalties, and demand threshold classification.

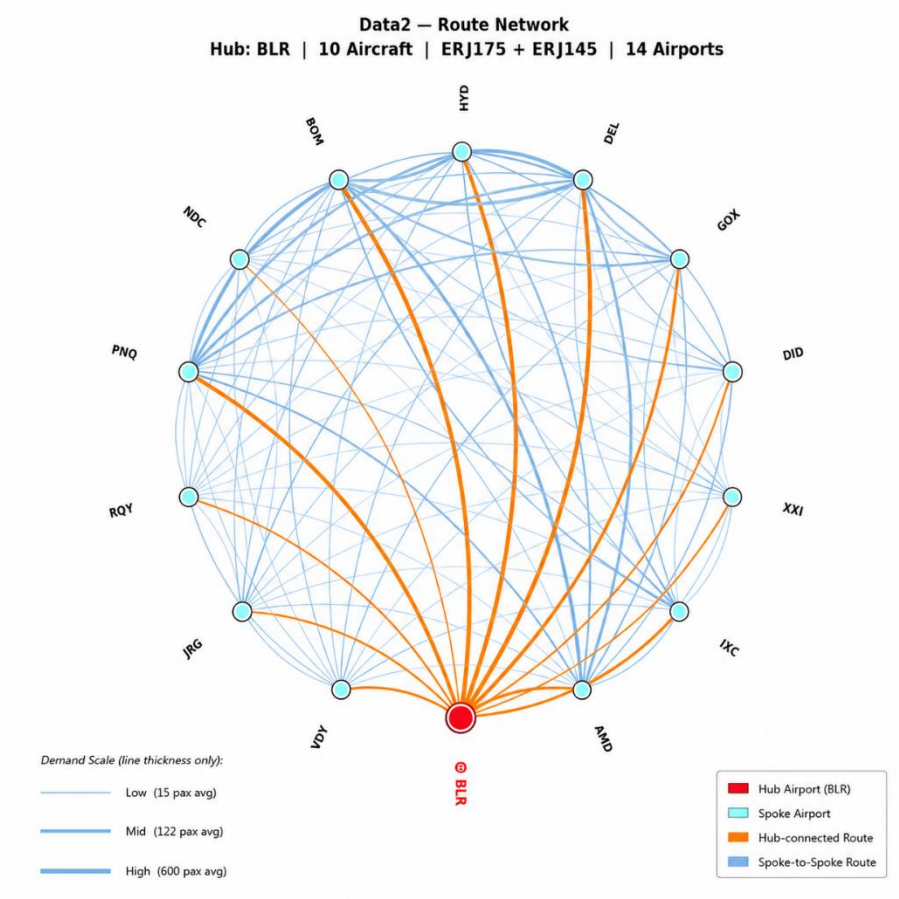


Figure 9: Route network, 14-airport regional passenger network (Data 2).

CDD importance scaled with network size: without demand discounting, multiple aircraft converged

https://doi.org/10.71945/icas2026_0816

on the same three or four highest-demand routes leaving the remainder unserved, whereas with CDD active demand distributed across 71.98 percent of the network. The one-step look-ahead handled the hub return feasibility requirement naturally, assigning zero combined score to any route from which hub return within the window was no longer feasible after the simulated assignment.

The core architecture transferred to the passenger context without structural modification, with operational extensions added as constraint layers, establishing that the methodology is not specific to cargo operations or small networks.

7. Discussion

The three case studies presented in this paper collectively demonstrate the scheduling behaviour of the Grey-GRASP-Lookahead method across increasing levels of network complexity, fleet heterogeneity, and demand scale. This section draws out the principal observations from each case, traces each to the specific architectural mechanisms responsible, and places them in the context of the limitations identified in prior work.

7.1 Four-Airport Case: Confirming the Look-Ahead Effect

As demonstrated in Section 6.1, the one-step look-ahead scheduler produced a higher cumulative reward than the greedy scheduler on an identical network by selecting a lower-immediate-reward first leg whose downstream continuation was more productive. This directly addresses the absence of anticipatory routing identified by Singh and Pant as a limitation of their methodology [4]. The improvement was achieved at the cost of one additional scoring computation per candidate per step, confirming that limited-horizon foresight is a low-cost intervention with measurable schedule quality benefit.

7.2 Maul's [7] Benchmark Case: Fleet Coordination and Scoring Under Constraint

The benchmark comparison against Maul [7] produced four observations that each trace to a specific architectural decision.

Fleet utilization consistently exceeded Maul's [7] results across all objectives and both demand levels, by margins ranging from 7.2 percentage points under original demand to 16.6 percentage points under 10× demand. The mechanism responsible is the interaction of distributed base initialisation and the GRASP expansion loop. Aircraft start at the four highest outbound-demand airports rather than at a single hub, meaning revenue-generating flights begin from the first minute of the window without any positioning requirement. The algorithm’s function then continues assigning flights to each aircraft until no feasible assignment with positive combined score remains, iterating across the full fleet until the window is exhausted. Maul's [7] scheduler terminates an aircraft when no route above the BELF threshold is reachable, causing early stoppage as high-demand trunk routes become saturated. The distributed initialisation and unconstrained expansion together ensure that the present scheduler extracts more block time from every aircraft in the fleet.

QOS was substantially better than Maul's [7] results across all configurations, with the largest advantage in the multi-fleet case where Maul's [7] QOS of 0.25 compared to values of 0.03 to 0.06 in the present results. As established in Section 5.3, this advantage is attributable to the absence of the BELF hard gate.

The cargo gap under the Utilisation objective, where the scheduler served 112.07 tonnes on average against Maul's [7] best of 123.0 tonnes, is structural rather than algorithmic. As established in Sections 4 and 5.1, it reflects Maul's [7] GA weight optimization advantage and the sharper route discrimination produced by Grey Clustering with whitening functions relative to Grey Relational Coefficients.

All 50 runs per objective mode produced feasible schedules with zero rejections across both demand levels and fleet configurations, with cargo standard deviations representing coefficients of variation between 2.5 and 2.9 percent. This level of consistency across runs where the GRASP alpha parameter is resampled from a range of 0.25 to 0.60 confirms that result quality is a property of the algorithm rather than of specific randomisation outcomes. Maul [7] presents three individual

https://doi.org/10.71945/icas2026_0816

schedule results per objective with no information on rejection rate or schedule diversity [7]. Singh similarly reported results from individual runs without statistical characterization [4, 5]. The 50-run statistical evaluation is therefore a methodological contribution in its own right, demonstrating that the scheduler reliably produces near-equivalent schedule quality across a wide range of stochastic parameter values, a prerequisite for operational confidence that neither prior work established.

7.3 Regional Passenger Network Case: Scalability and Emergent Behavior

The passenger network application produced three observations not visible in the benchmark case.

The first is emergent fleet specialisation. Without any hard routing rule assigning aircraft types to route categories, larger aircraft consistently self-selected onto trunk routes connecting the hub to high-demand metropolitan airports, while smaller aircraft covered thinner regional connections. The mechanism is the interaction of three scoring components: the Grey Relational Analysis score, which gives higher grades to routes with high raw demand and high demand intensity; the BELF soft penalty, which discounts the score of any assignment where expected load factor falls below break-even for the aircraft type; and the demand threshold classification, which signals the preferred aircraft type for each route based on remaining demand. The Grey scorer aggregates these signals into a combined score that consistently favours the larger aircraft on high-demand routes without explicit assignment. Singh's methodology could not produce this behaviour because it modelled only homogeneous fleets [4]. Maul [7] introduced heterogeneous fleet handling but his BELF gate hard-blocked type mismatches rather than producing emergent specialization through scoring incentives. The present approach produces commercially sensible fleet deployment as an output of the scoring architecture rather than as an imposed constraint.

The second observation is the demand scaling resolution. The initial application to the full demand matrix produced 37 percent demand satisfaction despite high fleet utilisation, because the matrix encoded total multi-frequency daily demand against a fleet sized for one scheduling wave. Dividing demand by a frequency factor of five raised demand satisfaction to 79.79 percent. The scheduler is a single-wave scheduler: it decides which aircraft goes where in one coordinated departure cycle, and the full daily demand is served across multiple such cycles. This distinction is operationally correct for hub-and-spoke regional operations and is a modelling insight rather than a limitation of the scheduling logic. Neither Singh nor Maul encountered this issue because their datasets were sized to single-cycle fleet capacity from the outset [4,7].

The third observation, that high-demand routes were consistently served before thin connections and all unserved pairs carried demand below 20 passengers, is discussed in full in Section 6.2. It confirms that Grey scoring combined with CDD produces the correct commercial priority ordering without any explicit route priority input.

8. Further Improvements

The results presented in this paper represent a pre-optimisation baseline configuration of the Grey-GRASP-Lookahead scheduler. Several extensions are identified that would materially improve both the quality and the operational scope of the methodology.

8.1 Grey Index Weight Optimisation

The Grey index weights used throughout this study were manually configured and held constant across all runs. Integrating an optimizer or an equivalent evolutionary algorithm as an outer loop over the GRASP scheduler is the single improvement most likely to close the cargo gap observed under the Utilisation objective in Table 2. In this formulation the weight vector would serve as the design variable, the mean cargo served across a fixed number of GRASP runs would serve as the objective function value returned to the optimizer, and the scheduler itself would remain structurally unchanged. This was the original architecture of Singh and Pant [4, 5] and its integration into the present framework is a natural next step.

https://doi.org/10.71945/icas2026_0816

8.2 Expanded Grey Index Set with Economic Layer

The four-index decision matrix used in the present work captures block time, demand intensity, route risk, and raw demand. Two additional indices are proposed for future incorporation. A Revenue Index would measure expected revenue per block hour on each candidate route, computed as payload carried multiplied by a yield rate per unit distance multiplied by route distance, divided by block time. A Cost Index analogous to Singh's original formulation [4] would measure direct operating cost per block hour, derived from aircraft type, stage length, and fuel burn parameters. Together these indices would allow the scheduler to distinguish between routes that are physically productive and routes that are economically productive.

8.3 Multi-Step Look-Ahead

The present look-ahead mechanism evaluates one step forward, which is sufficient to prevent the most common myopic failure mode of committing to a route that forecloses productive continuations. In networks with strong structural asymmetry, such as hub-and-spoke topologies where an early hub return eliminates all subsequent revenue flights, a two or three step horizon would provide materially better routing guidance. This extension is tractable within the existing architecture, as the top-N filter already limits the candidate set before look-ahead and can be tightened further to control the increased simulation cost.

8.4 Crew Constraints and Slot Modelling

Both Singh [4] and Maul [7] identified the absence of crew availability constraints as a limitation of the methodology. Incorporating crew as a scheduling resource would add a second feasibility dimension alongside aircraft availability, ensuring that assignments respect duty hour limits and rest compliance requirements. Airport slot modelling, present in Maul's formulation as 10-minute slot windows at each airport [7], was excluded from the present validation configuration to enable direct metric comparison. Its reintroduction would bring the operating time metric into alignment with Maul's [7] definition, enabling direct comparison on that metric also.

9. Conclusions

This paper presented a Grey-GRASP-Lookahead scheduling methodology for demand-responsive cargo and passenger airline operations, developed as an extension of the Grey Theory based scheduling lineage established by Singh and Pant [4, 5] and advanced by Maul [7]. The work was motivated by three structural limitations identified in prior studies: serial scheduling with no inter-aircraft coordination, purely greedy route selection with no anticipation of downstream routing consequences, and deterministic weight configurations with no mechanism for exploring diverse feasible schedules across repeated runs.

Four integrated mechanisms were introduced to address these limitations. Grey Relational Analysis with relational coefficients replaced Grey Clustering as the route scoring method, providing a reference-based evaluation that generalizes across network configurations without requiring dataset-specific whitening thresholds. A Greedy Randomized Adaptive Search Procedure with a Restricted Candidate List replaced deterministic greedy selection, enabling controlled stochastic exploration of the schedule space. A Committed Demand Discount mechanism introduced fleet-wide coordination by discounting each route's effective remaining demand by cargo or passengers already committed by other aircraft in the current scheduling attempt, preventing the demand pile-up that serial architectures cannot avoid. A one-step look-ahead evaluated the fleet-competitive state following each candidate assignment before committing to it, reducing the myopic routing decisions that characterize purely greedy schedulers.

Validation against Maul's [7] six-airport Indian cargo benchmark demonstrated that the methodology, operating without any parameter tuning, matches or exceeds GA-optimized Grey Clustering on four of five metrics under the Time objective and three of five under the Utilization objective. Fleet utilization exceeded Maul's [7] results by 7.2 to 16.6 percentage points across demand levels and

https://doi.org/10.71945/icas2026_0816

fleet configurations. Quality of Service improved by 25 to 86 percent depending on objective and fleet type. Cargo throughput efficiency reached 3.60 tons per hour against Maul's [7] 2.85 under the Time objective at original demand, and 8.14 to 8.45 tons per hour against Maul's [7] 5.31 to 5.83 under ten times demand. The single metric where Maul's [7] method outperformed the present approach, total cargo under the Utilization objective, is attributable to his GA weight optimization advantage and is identified as the primary target for the next development stage.

All 100 single-fleet benchmark runs produced feasible schedules with zero rejections, with cargo standard deviations representing coefficients of variation below three percent across all objectives and demand levels. This statistical characterization, absent from both prior works, demonstrates that the scheduler reliably produces near-equivalent schedule quality across a wide range of stochastic parameter values rather than depending on favorable individual outcomes.

The extension to a 14-airport regional passenger network with a 10-aircraft heterogeneous fleet demonstrated the generalizability of the methodology beyond cargo operations and small networks, with three emergent behaviors observed: fleet specialization, demand-driven route prioritization, and natural resolution of hub return feasibility through the look-ahead mechanism.

GA weight optimization, expanded Grey indices with economic criteria, multi-step look-ahead, and crew and slot modelling are identified as the principal further development directions, each compatible with the existing architecture.

10. Contact authors and Email Address

Rajkumar S. Pant, Professor, Aerospace Engineering Department, Indian Institute of Technology Bombay, Mumbai, India. Telephone: +91-9820084502. Telefax: +91(0)22-25720257. E-mail: rpant@iitb.ac.in

Palak Singh, Research Intern, Aerospace Engineering Department, Indian Institute of Technology Bombay, Mumbai, India. Telephone: +91-8840588273. E-mail: ps.3030.work@gmail.com

11. Copyright statement

The authors confirm that they, and/or their company or organization, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third-party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS proceedings or as individual off-prints from the proceedings.

References

[1] Barnhart C, Cohn A, Johnson E, Klabjan D, Nemhauser G, Vance P. Airline crew scheduling. In: Hall RW, editor. Handbook of Transportation Science. Boston: Kluwer Academic Publishers; 2003. p. 517–560.

[2] Deng J. Introduction to grey system theory. Journal of Grey Systems. 1989;1(1):1–24.

[3] Hsu C, Wen Y. Application of grey theory and multi-objective programming towards airline network design. European Journal of Operational Research. 2000; 127:44–68.

[4] Singh A. Studies in airline scheduling and network optimization [master's dissertation]. Mumbai, India: Indian Institute of Technology Bombay, Aerospace Engineering Department; 2007.

[5] Singh A, Pant RS. Airline network optimization using grey theory. In: Proceedings of the AIAA Aviation Technology, Integration and Operations Conference; 2007 Sep 18–20; Belfast, Northern Ireland. Reston (VA): American Institute of Aeronautics and Astronautics; 2007. Paper No. AIAA 2007-7757.

[6] Rasheed KM. GADO: A genetic algorithm for continuous design optimization [PhD dissertation]. New Jersey: Rutgers University; 1998.

[7] Maul F. Enhanced flight scheduling algorithm. Semester Arbeit. Mumbai, India: Aerospace Engineering Department, Indian Institute of Technology Bombay; 2014.

[8] Feo TA, Resende MGC. Greedy randomized adaptive search procedures. Journal of Global Optimization. 1995;6(2) Vol. 6, No.2, pp. 109-133.

https://doi.org/10.71945/icas2026_0816