

A Unified Modified Vogel Framework for Transportation, Transshipment, ATSP and Capacitated Flows: Duality Certificates, Algorithmic Extensions and a Scalability Study

Oumar Koné¹, Pawoumodom Matthias Takouda², Youssou Gningue³

¹Basic and Applied Sciences—Mathematics and Computer Science Laboratory, Nangui Abrogoua University, Abidjan, Côte d'Ivoire

²Research Group in Operations, Analytics and Decision Sciences, School of Business Administration, Laurentian University, Sudbury, Canada

³Full and Emeritus Professor of Mathematics, Laurentian University, Sudbury, Canada
Email: koneoumar.sfa@univ-na.ci, mtakouda@laurentian.ca, ygningue@cs.laurentian.ca

How to cite this paper: Koné, O., Takouda, P.M. and Gningue, Y. (2026) A Unified Modified Vogel Framework for Transportation, Transshipment, ATSP and Capacitated Flows: Duality Certificates, Algorithmic Extensions and a Scalability Study. *Open Journal of Applied Sciences*, 16, 3841-3860.
<https://doi.org/10.4236/ojapps.2026.1610212>

Received: July 31, 2026

Accepted: October 12, 2026

Published: October 15, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

We develop a unified algorithmic and theoretical framework centred on the *Modified Vogel Method* (MVM) [1]. The framework covers five closely related models arising in logistics and combinatorial optimization: 1) balanced linear transportation problems (LTPs); 2) unbalanced LTPs with dummy-aware reduction; 3) transshipment networks compressed by Floyd minimum-cost paths into small equivalent transportation problems; 4) asymmetric travelling salesman problems (ATSPs) via an MVM-improved nearest-neighbour heuristic (MVM-NNA); and 5) capacitated (bounded) transportation problems through a VAM-B allocation rule with box complementary-slackness certificates. Building on a duality-based reframing, we interpret row/column reduction as the construction of dual-feasible potentials—the same mechanism that underlies the first phase of the Hungarian algorithm—and we state optimality certificates for both the unbounded and capacitated settings. We then present a carefully engineered, reproducible computational study comprising 1401 scalability runs plus a companion small-scale battery designed to reconnect with [1]. Reading the evidence by *scale* recovers that chapter's qualitative message at sizes $n \in \{3, \dots, 15\}$ (MVM matches or beats VAM on 92.6% of a chapter-like schedule, with mean LP gaps at 3.96% versus 7.35%), while the large schedule (n up to 1000) shows the natural evolution and nuance: pathwise dominance softens to 63.6% overall, yet mean gaps remain favourable (20.01%

versus 28.63% on LP-checked sizes). Floyd compression shrinks equivalent TPs by a mean factor of 6.4; MVM-NNA improves tours by 15.71% on average against fixed-start NNA; capacitated VAM-B recovers the technical note's 2×2 optimum and attains a mean LP gap of 21.37%. We discuss certificates, open structural questions, and a MODI-style correction paired with VAM-B.

Keywords

Transportation Problem, Modified Vogel Method, VAM, Transshipment, Floyd-Warshall, Asymmetric TSP, Nearest Neighbour, Capacitated Flows, Linear Programming Duality, Complementary Slackness, Scalability, Computational Study

1. Introduction

1.1. Motivation and Industrial Context

The linear transportation problem (LTP) asks how to ship a homogeneous commodity from a set of sources to a set of destinations so as to meet supplies and demands at minimum total unit-shipping cost. Despite its classical status [2]-[4], the LTP remains a workhorse model in distribution planning, warehouse-to-retail allocation, and as a building block inside richer network designs [5]. In many operational settings, the planner faces variants that are only superficially different from the balanced bipartite LTP: supplies and demands that do not match (unbalanced problems); intermediate cross-docks (transshipment); routing decisions that visit each customer once (travelling salesman); and contractual or physical capacities on individual origin-destination cells. A methodological approach that remains coherent across these variants is therefore of both theoretical and practical interest.

1.2. From VAM to MVM

Among constructive heuristics that produce basic feasible solutions for the LTP, Vogel's Approximation Method (VAM) [2] is repeatedly reported among the strongest classical options [6]. VAM assigns flow iteratively to a cell selected by line *penalties*—the difference between the two cheapest costs on a row or column—thereby capturing opportunity costs rather than myopic least-cost choices.

[1], extending the earlier reduced-matrix construction of Diagne and Gningue [7], proposed the Modified Vogel Method (MVM): first apply admissible row and column reductions of the cost matrix [8], then run VAM on the reduced matrix, with local re-reduction after each assignment and an internal optimality test based on the number of reduction rounds and the uniqueness of largest penalties. On a small random battery (sizes roughly 3 - 15), the chapter reported that MVM always improved upon—or at least matched—VAM, with improvement rates up to about 21%, and proved optimality in several structurally identifiable cases. The present paper is written as a *direct prolongation* of that programme: we first reproduce a

chapter-scale regime where the same qualitative conclusion reappears, then let instance size grow so that the evolution of pathwise and mean-gap behaviour becomes visible, and finally address the chapter's open call for comparisons against exact optima (via LP oracles) together with the later unbalanced, transshipment, ATSP and capacitated extensions.

1.3. Related Extensions in the Literature Strand

Three further developments motivate the present unification.

- [9] treated unbalanced LTPs by dummy rows/columns while ordering reduction so that big- M dummy costs leave the active reduced matrix, avoiding the pathologies discussed by [10]–[13].
- [14] reduced transshipment networks to *small* transportation problems by computing Floyd minimum-cost paths between sources and destinations, contrasting the classical buffer/all-node reformulations that inflate dimensions [15] [16].
- [17] improved the asymmetric nearest-neighbour heuristic by selecting the start city from the first MVM penalty iteration and refining tours via *missing zeros*.

1.4. Contributions

This paper consolidates the above strand into a single journal-level treatment with the following contributions.

- 1) **Unified modelling.** We present MVM as a family covering balanced and unbalanced LTPs, Floyd-compressed transshipment, MVM-NNA for ATSP, and capacitated VAM-B, with a common reduction-and-penalty engine.
- 2) **Duality certificates.** We give self-contained statements and proofs of complementary-slackness optimality for unbounded LTPs, and of box-constrained certificates for capacitated LTPs [18], clarifying the link to Hungarian potentials.
- 3) **Worked illustrations.** We reproduce [1]'s 4×5 example (optimal cost 339), the note's 2×2 capacitated counterexample (optimal cost 32), and the logic of Floyd compression.
- 4) **Reproducible scalability study.** Over 1401 runs with strict separation of models, instances, parameters and results, we report quality, runtime, certificates and dimension-reduction metrics, with LP baselines on moderate sizes.
- 5) **Open problems.** We formulate Hall-type support questions and worst-case re-reduction bounds as concrete open problems for the VAM/MVM literature, which remains largely empirical [6].

1.5. Organisation

Section 2 recalls notation and the classical LTP. Section 3 develops the duality reframing. Section 4 presents balanced and unbalanced MVM algorithms. Section 5 treats Floyd-compressed transshipment. Section 6 develops MVM-NNA. Section 7 presents capacitated VAM-B. Section 8 details the experimental protocol

and results. Section 9 discusses limitations and open questions. Section 10 concludes.

2. The Classical Linear Transportation Problem

2.1. Balanced Linear Transportation Problem

Definition 1 (Balanced LTP). Given m sources with supplies $a \in \mathbb{R}_+^m$, n destinations with demands $b \in \mathbb{R}_+^n$ satisfying $\sum_i a_i = \sum_j b_j$, and a cost matrix $C = (c_{ij}) \in \mathbb{R}_+^{m \times n}$, the balanced LTP is

$$\begin{aligned} \min_X \quad & \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\ \text{s.t.} \quad & \sum_{j=1}^n x_{ij} = a_i, \quad i = 1, \dots, m, \\ & \sum_{i=1}^m x_{ij} = b_j, \quad j = 1, \dots, n, \\ & x_{ij} \geq 0, \quad \forall i, j, \end{aligned} \quad (1)$$

where x_{ij} represents the number of units of flow shipped from origin i to destination j , for all i, j .

The dual of (1) uses free variables $(u, v) \in \mathbb{R}^m \times \mathbb{R}^n$:

$$\max \sum_i a_i u_i + \sum_j b_j v_j \quad \text{s.t.} \quad u_i + v_j \leq c_{ij} \quad \forall i, j. \quad (2)$$

The quantities $r_{ij} := c_{ij} - u_i - v_j$ are the reduced costs. Dual feasibility is achieved when $r \geq 0$.

2.2. Vogel Penalties

Definition 2 (Line penalty). For a (possibly reduced) cost matrix R and an active row i , the Vogel penalty is $p_i = r_{(2)} - r_{(1)}$, where $r_{(1)} \leq r_{(2)}$ are the two smallest finite costs on that row among active columns. The Vogel penalty q_j , for an active column j , is defined analogously.

On a fully reduced matrix, one typically has $r_{(1)} = 0$, so penalties coincide with second-smallest entries [1].

Classical VAM repeatedly selects the line of the largest penalty, assigns as much flow as possible to its cheapest active cell, updates residual supply/demand, and deletes saturated lines until a basic feasible solution is obtained.

3. Duality Reframing of MVM

3.1. Reduction as Dual-Feasible Potential Construction

Definition 3 (Row-column reduction). Starting from $C \geq 0$, set $u_i := \min_j c_{ij}$ and replace c_{ij} by $c_{ij} - u_i$ creating the matrix C' . Then, set $v_j := \min_i c'_{ij}$ on the intermediate matrix C' and subtract. The output R has a zero in every row and column, and $r_{ij} = c_{ij} - u_i - v_j \geq 0$.

Thus, reduction is precisely the construction of a dual-feasible potential pair. This is the same algebraic move as the first phase of the Hungarian algorithm for

the assignment problem (the special case $m = n$, $a_i = b_j = 1$).

Theorem 4 (Complementary slackness for LTP). Let (u, v) be dual feasible for (2) and let x be primal feasible for (1).

If

$$x_{ij} > 0 \Rightarrow c_{ij} = u_i + v_j$$

(i.e., x is supported on the zero set of the reduced matrix), then x is optimal for (1) and (u, v) is optimal for (2).

Proof. For any primal-feasible x' , non-negativity of reduced costs yields

$$\sum_{ij} c_{ij} x'_{ij} \geq \sum_{ij} (u_i + v_j) x'_{ij} = \sum_i u_i a_i + \sum_j v_j b_j =: D.$$

Evaluating the objective at the particular x that is supported on zeros gives $\sum c_{ij} x_{ij} = D$. Hence, x meets the dual lower bound and is optimal; D certifies dual optimality of (u, v) by weak-duality tightness. $\square$

Corollary 1 (MVM zero-support certificate). If VAM, run on the reduced cost matrix, produces a basic feasible solution using only zero-reduced-cost cells, that solution is optimal for the original LTP.

3.2. Local Re-Reduction and Hungarian-Style Repair

When an assignment removes the unique zero of a surviving line, MVM subtracts the new line minimum from that line. Algebraically, this process modifies the potentials to re-establish dual feasibility and introduce a new zero element. While conceptually similar to the Hungarian dual update, this local variant is driven by VAM's greedy penalty mechanism rather than an alternating-path search.

Remark 5. [1] prove that if no further reduction occurs after the initial pass ($N_{\text{red}} = 1$), or if every successive line removal is associated with a unique largest penalty, then MVM's solution is optimal. These combinatorial tests are claimed as sufficient in the paper. In our implementation, however, the UniqueLpen flag allows false positives when $N_{\text{red}} > 1$. Consequently, we restrict these criteria $N_{\text{red}}/\text{UniqueLpen}$ to diagnostics and instead verify optimality using a dual complementary-slackness certificate (ensuring valid potentials exist for the support of X), for problem sizes where $mn \leq 2500$.

Operationally, the "Cert. MVM" flag does not rely on potentials accumulated during greedy local re-reduction. After X is obtained, we recover witnessing duals by solving the linear feasibility system $u_i + v_j = c_{ij}$ on the support of X and $u_i + v_j \leq c_{ij}$ everywhere (HiGHS; $mn \leq 2500$). Feasibility of this system is exactly the complementary-slackness certificate of Theorem 4, and it remains well-defined for degenerate or disconnected supports.

3.3. Support Feasibility as a Hall-Type Condition

Let G_0 be the bipartite graph whose edges are the zero cells of the reduced matrix. One reduction pass suffices for a fully certified optimum if and only if there exists a feasible transportation flow supported on G_0 . By standard Gale-Hoffman theory, this holds if and only if, for every set S of sources,

$$\sum_{i \in S} a_i \leq \sum_{j \in N_{G_0}(S)} b_j,$$

and symmetrically for destination sets. This diagnostic is independent of VAM's greedy choices and can be checked by a single max-flow computation. We state as an open conjecture that the number of additional local re-reduction rounds is at most $\min(m, n) - 1$ in the worst case, analogous to the Hungarian algorithm's bound of n augmenting phases for $n \times n$ assignment.

4. Algorithms for Balanced and Unbalanced MVM

4.1. Algorithm for the Balanced Case

Algorithm 1 summarizes MVM for a balanced instance. The implementation tracks N_{red} , ties for the largest penalty and the internal optimality flags of [1].

Example 6. ([1] 4×5 instance). [1] illustrates MVM on

$$C = \begin{pmatrix} 1 & 3 & 8 & 7 & 6 \\ 8 & 10 & 7 & 9 & 6 \\ 2 & 11 & 6 & 8 & 9 \\ 5 & 9 & 8 & 7 & 7 \end{pmatrix}, \quad a = (15, 25, 10, 16), \quad b = (10, 15, 12, 14, 15).$$

After reduction and four assignment iterations (including a three-way tie broken by least cost), MVM returns a solution of total cost $TC = 339$ that coincides with the transportation-simplex optimum. Our implementation reproduces this value exactly and confirms optimality with an LP solver.

Algorithm 1. Modified Vogel method (balanced LTP).

Require: Cost matrix C , supply a , demand b with $\sum a = \sum b$

Ensure: Shipment X , cost, certificate flags

- 1: Reduce C to R (rows then columns); set $N_{\text{red}} \leftarrow 1$, $UniqueLpen \leftarrow \text{true}$
- 2: **while** more than one active line remains **do**
- 3: Compute penalties on active rows/columns of R
- 4: Let L be a line attaining the largest penalty; if tied, mark tie
- 5: Assign $x_{kr} \leftarrow \min(a_k, b_r)$ to a zero reduced-cost cell on L
- 6: Update residuals; delete saturated line(s)
- 7: Locally re-reduce remaining matrix if needed; update N_{red}
- 8: If a new reduction occurred after a tie, set $UniqueLpen \leftarrow \text{false}$
- 9: **end while**
- 10: Fill the last line; return X
- 11: **Certificate (implementation):** dual potentials exist for the support of X (complementary slackness / MODI). Flags N_{red} and $UniqueLpen$ are retained as diagnostics only— $UniqueLpen$ alone is not used as a certificate (empirically unsafe when $N_{\text{red}} > 1$).

4.2. Algorithms for Unbalanced Problems

If $\sum a \neq \sum b$, a dummy destination (excess supply) or dummy source (excess demand) of residual volume $|\sum a - \sum b|$ and cost $M \gg 0$ restores balance [9]. The crucial algorithmic point is the *order* of reduction:

- **Dummy column:** Reduce rows first, then columns—zeros appear off the dummy,

and column reduction clears M from the active reduced matrix;

- *Dummy row*: Reduce columns first, then rows, symmetrically.

Thereafter MVM proceeds as in the balanced case; the reported shipment is restricted to original cells, and dummy flow is recorded separately. This strategy is related to, but distinct from, the classical modifications of [10]–[13], which either ignore dummy costs in penalties or replace M by the largest genuine cost.

5. Transshipment via Floyd Compression

5.1. Problem and Classical Reformulations

A transshipment problem ships from sources to destinations through a directed network that may contain pure transshipment (junction) nodes. Flow-conservation constraints hold at junctions; supplies and demands appear at terminals [15]. Two classical LP reformulations build large equivalent transportation problems in which junctions become both sources and destinations, with buffer values and big- M arcs for forbidden trips [14] [16] [19]. Their transportation matrices are of order $(|V| \times |V|)$ and become cumbersome even for hand-worked textbook instances [15].

5.2. Floyd Reduction to a Small TP

Proposition 7 (Floyd compression). Consider a transshipment network with *uncapacitated* arcs, that is, with finite arc costs and node supplies and demands but no arc bounds, and assume every required source-destination pair is reachable. Then replacing each such pair by its minimum-cost path length yields an equivalent transportation problem whose optimal cost equals that of the original transshipment model, and expanding the transportation shipment along the corresponding shortest paths recovers an optimal transshipment flow of equal cost.

The resulting TP has size $|S| \times |D|$ only. After MVM (or any TP solver) returns shipments y_{ij} , each positive y_{ij} is expanded by backtracking the Floyd next-hop matrix to primal arc flows [14]. Dimension reduction versus the classical all-node model is quadratic in $|V|$ versus $|S||D|$.

6. MVM-NNA for the Asymmetric TSP

6.1. Nearest Neighbour and Its Sensitivity to the Start

Given an asymmetric distance matrix $D \in \mathbb{R}_+^{n \times n}$ (zero diagonal), the nearest-neighbour algorithm (NNA) builds a tour by always appending the nearest unvisited city, then returning to the start. NNA is fast but highly sensitive to the initial city; a poor start can force expensive late insertions [20].

6.2. Using One MVM Iteration as a Start Selector

MVM-NNA proceeds as follows.

- 1) Reduce D (forbidding self-loops) to obtain potentials and reduced costs.
- 2) Compute row penalties; take as start the row of largest penalty.

3) Run NNA from that start.

4) Optionally improve: Identify tour arcs with positive reduced cost (*missing zeros*); restart or reconnect from those cities using cheaper successors on the reduced row.

Only the first MVM iteration is used for start selection; full MVM is not run on the TSP matrix. Missing zeros also indicate promising neighbourhoods for local search or seeds for evolutionary populations—remarks already made by [17] and consistent with TSP local-search practice [20].

7. Capacitated Extension (VAM-B)

7.1. Bounded Transportation Problem

Definition 8 (Capacitated LTP). We reconsider the classical LTP (1). If we add to the problem the constraints $0 \leq x_{ij} \leq u_{ij}$ (with $u_{ij} = +\infty$ allowed), we obtain a bounded transportation problem. Two-sided bounds reduce to this form by using the standard shift $x = \ell + y$ [18].

7.2. Why Naïve Zero-Support Certificates Fail

Let's consider the following example.

Example 9. Take $a = b = (10, 10)$, $C = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$ and bound $x_{11} \leq 6$. Parametrizing $t = x_{11} \in [0, 6]$ yields total cost $50 - 3t$, minimised at $t = 6$ with cost 32 and $X = \begin{pmatrix} 6 & 4 \\ 4 & 6 \end{pmatrix}$. Under potentials from ordinary reduction, cell $(1, 2)$ may carry positive reduced cost while being basic—so a naïve port of Corollary 1 incorrectly refuses to certify an already optimal solution. With corrected potentials $(u, v) = (3, 1; 1, 0)$ one obtains $r_{11} = -3 \leq 0$ at capacity and zeros on interior cells, matching the box certificate below. Our solver and LP oracle both recover cost 32.

7.3. Allocation Rule and Box Certificate

VAM-B uses the same penalties on the reduced matrix, but allocates

$$qty = \min \{a_i^{\text{rem}}, b_j^{\text{rem}}, u_{ij}^{\text{rem}}\}.$$

If capacity binds while supply and demand remain positive, the cell is *blocked* (removed from further consideration) without deleting its row or column; local re-reduction applies as in unbounded MVM.

Theorem 10 (Box complementary slackness). Let x be feasible for the capacitated LTP. If there exist potentials (u, v) such that, writing $r_{ij} = c_{ij} - u_i - v_j$,

$$\begin{aligned} r_{ij} &\geq 0 && \text{whenever } x_{ij} = 0, \\ r_{ij} &= 0 && \text{whenever } 0 < x_{ij} < u_{ij}, \\ r_{ij} &\leq 0 && \text{whenever } x_{ij} = u_{ij}, \end{aligned}$$

then x is optimal [18].

Economically, a negative reduced cost on a capacity-tight cell means the plan-

ner would ship more through that cheap arc if the capacity allowed—exactly the signal that binding capacities should produce. Automatic recovery of such potentials after a greedy pass is harder than in the unbounded case. Our experiments report honest certification rates alongside LP gaps.

8. Computational Scalability Study

8.1. Design Principles

The experimental pipeline enforces a strict separation of concerns:

- **Models:** Algorithms only (VAM, MVM, Floyd, MVM-NNA, VAM-B, LP);
- **Instances:** Seeded generators and literature fixtures;
- **Parameters:** YAML configuration (sizes, replicates, LP cell budget);
- **Results:** CSV metrics, figures, and LaTeX macros.

This architecture prevents silent coupling between experiment control and algorithmic code and supports exact regeneration of tables and figures.

8.2. Instance Families and Metrics

Unless otherwise stated, we use seed base 42 and thirty-five independent instances per size. Sizes follow a doubling schedule from 10 to 1000:

$$n \in \{10, 20, 40, 80, 160, 320, 640, 1000\}.$$

1) **Balanced LTP:** Square instances $m = n$. Balanced supplies are drawn i.i.d. as $a_i \sim \text{Unif}\{10, \dots, 50\}$ and demands form a random positive integer partition of $S = \sum_i a_i$ into n parts, obtained from $n-1$ ordered uniform cuts on $\{1, \dots, S-1\}$; costs satisfy $c_{ij} \sim \text{Unif}\{1, \dots, 100\}$. Metrics: VAM/MVM cost, runtime, MVM certificate, LP gap when $mn \leq 2500$ (i.e., $n \leq 40$).

2) **Unbalanced LTP:** Unbalanced instances reuse this construction and then rescale supplies by the stated ratio in $\{0.7, 1.3\}$, rounded and floored at 1, leaving costs and demands unchanged.

3) **Transshipment:** $|S| = |D| = n$ with a capped transit layer; metrics include classical-to-reduced cell-count ratio.

4) **ATSP:** ATSP matrices are i.i.d. asymmetric and *non-metric*: $D_{ij} \sim \text{Unif}\{1, \dots, 99\}$ for $i \neq j$ with $D_{ii} = 0$, and no Euclidean perturbation is applied; this matters because metric structure would materially change nearest-neighbour behaviour. Compare NNA (start 0), MVM-NNA and multi-start NNA (up to 10 starts).

5) **Capacitated:** For capacitated instances, about 30% of the cells receive a finite upper bound drawn uniformly from $[0.5\bar{f}, 2.5\bar{f}]$, where $\bar{f} = (\sum_i a_i)/(mn)$ is the mean cell flow of a uniform shipment, so capacities are calibrated to the flow scale induced by (a, b) rather than fixed a priori. If a row or column then carries total capacity below its supply or demand, all finite caps on that line are lifted to $+\infty$, which restores primal feasibility by construction. Metrics: LP gaps and box-certificate flags.

We use the instances from Example 6 (cost 339) and Example 9 (cost 32) as golden

targets for our unit-test oracles. In addition, a *chapter-scale companion battery* (sizes 3 - 15, including a 27-instance schedule mirroring [1] and a dense 455-instance design) is analyzed in Section 8.3 to reconnect small-scale conclusions with the large-scale schedule above.

8.3. Balanced Transportation

We read the balanced evidence in three nested regimes—small ([1]’s scale), medium (start of the full schedule), and large—so that the 2015 conclusions remain recoverable before nuances appear; **Table 1** condenses the ladder.

Small scale ([1] regime). We first rebuild a battery in the spirit of [1]: 27 balanced LTPs with square sizes 5, 10, 15 (four replicates each) plus fifteen mixed shapes with $m, n \in \{3, 4, 5\}$, all solved by the same VAM/MVM engines and an LP oracle. On this schedule, MVM matches or beats VAM on 92.6% of instances (strict improvement on 51.9%; MVM worse on only 7.4%), mean LP gaps are 3.96% versus 7.35%, exact optimality of the MVM shipment occurs on 55.6% of instances, and the largest relative cost improvement over VAM reaches 19.7%—the same order of magnitude as the chapter’s reported maximum of 20.92%. Absolute “always” pathwise dominance is therefore almost recovered on a chapter-sized sample; the few reversals are consistent with finite-sample fluctuation and deterministic tie-breaking once several lines share the largest penalty.

Table 1. Scale ladder for balanced MVM versus VAM (same generators and solvers). Pathwise rate = share of instances with MVM cost $\leq$ VAM cost, with Wilson 95% CI. Gaps are mean relative gaps to LP optima (%; unavailable when $n \geq 80$). Mean paired relative improvement $RI = (c^{VAM} - c^{MVM}) / c^{VAM}$ with Student 95% CI.

Regime	N	MVM $\leq$ VAM (%)	Wilson 95% CI	Gap MVM	Gap VAM	Mean RI (%)	RI 95% CI
Chapter-like schedule ($n \in \{3, \dots, 15\}$)	27	92.6	[76.6, 97.9]	3.96	7.35	2.77	[-0.22, 5.76]
Dense small ($n = 3, \dots, 15, 35$ reps)	455	70.8	[66.4, 74.8]	8.46	11.13	1.80	[0.91, 2.69]
of which $n \leq 5$	105	82.9	[74.5, 88.9]	1.69	2.15	0.26	[-0.76, 1.28]
Full battery, $n = 10$	35	77.1	[61.0, 87.9]	6.67	12.54	4.52	[0.71, 8.34]
Full battery, all $n \leq 1000$	280	63.6	[57.8, 69.0]	20.01	28.63	2.80	[1.48, 4.12]
Full battery, $n \geq 80$ (no LP)	175	60.6	[53.2, 67.5]	—	—	1.00	[-0.60, 2.60]

Alongside pathwise rates, **Table 1** reports the mean paired relative improvement $RI = (c^{VAM} - c^{MVM}) / c^{VAM}$ with a Student 95% confidence interval, and Wilson 95% intervals on the MVM $\leq$ VAM frequency. On the chapter-like schedule, the mean RI is only 2.77% (CI [-0.22, 5.76]), yet the pathwise rate remains high (92.6%, Wilson [76.6, 97.9]), showing that dominance is frequent even when average cost gains are modest. On the full battery, the mean RI is 2.80% (CI [1.48, 4.12]) with pathwise rate 63.6% (Wilson [57.8, 69.0]); for $n \geq 80$ the mean RI CI [-0.60, 2.60] covers zero, consistent with the non-monotonic pathwise pattern visible by size.

A denser companion design (455 square instances, $n = 3, \dots, 15, 35$ replicates

per n) confirms that the [1]'s message is a genuine *small-size phenomenon* rather than an artefact of 27 lucky draws: overall $MVM \leq VAM$ holds on 70.8% of instances (82.9% when $n \leq 5$), while mean gaps remain favourable (8.46% versus 11.13%). Thus, at the scale studied in 2015, our prolongation yields the same qualitative conclusion: MVM is the better constructive start.

Evolution along the full schedule. Table 2 and Figure 1, Figure 2 summarise the double-up schedule $n \in \{10, \dots, 1000\}$ (280 runs). At the first rung ($n = 10$)—still close to the [1]'s window— $MVM \leq VAM$ holds on 77.1% of draws and mean gaps are 6.67% versus 12.54%. As n grows within the LP-checked window, mean-gap advantages widen (21.6% vs 29.5% at $n = 20$; 31.8% vs 43.9% at $n = 40$), in line with the [1]'s claim that reduced-cost penalties are qualitatively sharper. Aggregating all full-battery sizes, MVM beats or ties VAM on 63.6% of instances; mean LP gaps are 20.01% versus 28.63%; exact LP matches occur for 2.9% (MVM) versus 1.0% (VAM); dual complementary-slackness certificates (for $mn \leq 2500$) fire on 1.1% of all runs.

Table 2. Balanced LTP scalability (35 instances per size; LP gaps for $n \leq 40$).

n	Gap MVM	Gap VAM	$MVM \leq VAM$ Cert.	MVM	t_{MVM} (s)	t_{VAM} (s)
10	6.67%	12.54%	77.1%	8.6%	0.004	0.002
20	21.57%	29.45%	62.9%	0%	0.015	0.011
40	31.80%	43.91%	65.7%	0%	0.055	0.020
80	—	—	42.9%	0%	0.18	0.068
160	—	—	71.4%	0%	0.79	0.31
320	—	—	65.7%	0%	3.69	1.54
640	—	—	57.1%	0%	18.6	8.58
1000	—	—	65.7%	0%	56.6	26.9

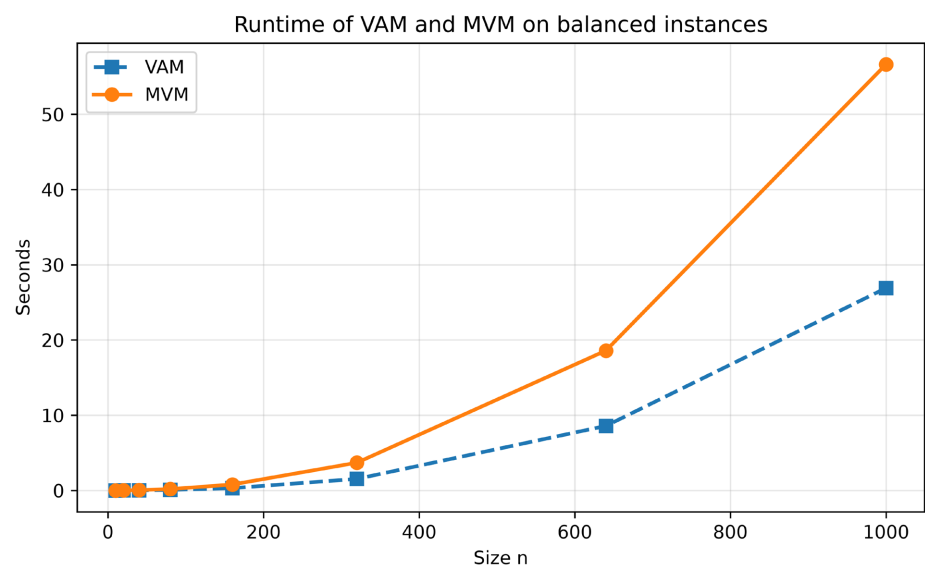


Figure 1. Mean runtime of VAM and MVM versus size on balanced random LTPs ($n \in \{10, \dots, 1000\}$, 35 instances per size).

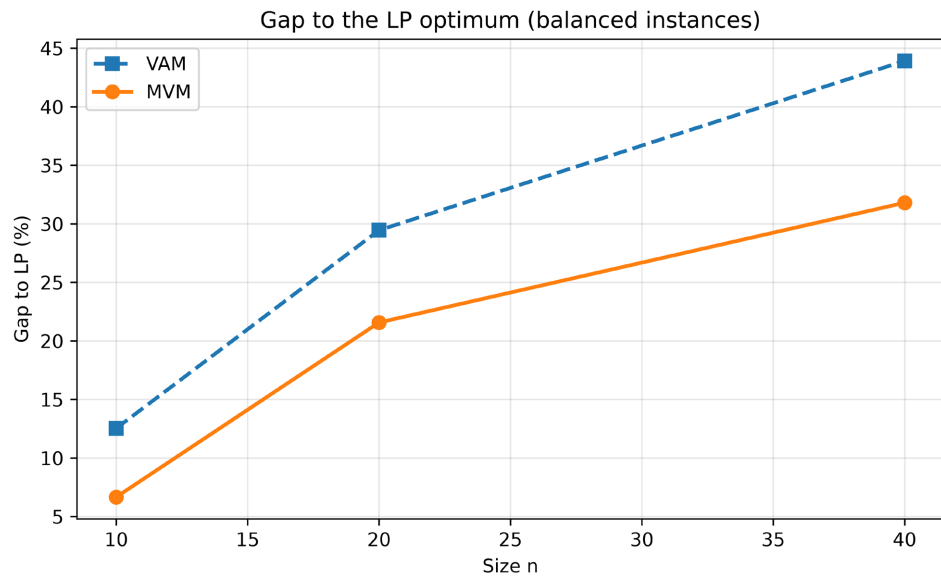


Figure 2. Mean relative gap to LP (%) versus size (LP skipped for $n \geq 80$).

Table 2 makes the size-wise dispersion explicit: Wilson intervals on $MVM \leq VAM$ are wide at $N = 35$ per size (e.g., $[28.0, 59.1]$ at $n = 80$ versus $[61.0, 87.9]$ at $n = 10$), and the mean RI CI excludes zero at $n \in \{10, 20, 40\}$ but covers zero for several larger sizes, including a negative point estimate at $n = 80$. These paired summaries use only the already stored per-instance costs (c^{VAM}, c^{MVM}) ; no new runs were required.

Nuances at large sizes. Pathwise dominance is incomplete once leaving the micro-regime: the fraction $MVM \leq VAM$ falls to 60.6% for $n \geq 80$ and can dip lower at individual sizes (e.g., 42.9% at $n = 80$). That does *not* reverse the mean-quality message where LP checks are available, but it shows that “MVM always beats VAM” is a small-scale statement, not a universal law. Reporting distributions—and not only averages—therefore matters [6]. Runtimes scale as expected for an $O(n^3)$ -style greedy loop: MVM is roughly twice as slow as VAM in our Python implementation, yet remains practical up to $n = 1000$ (about 57 s versus 27 s per instance on average for VAM). Absolute timings differ from the chapter’s Java codes; the quality narrative, not the stopwatch, is the object of the prolongation.

Deterministic tie-breaking. When several active lines attain the same largest Vogel penalty, we break ties deterministically: rows are preferred over columns and, within that class, the line of smallest index is selected. On the chosen line, the flow is assigned to a zero reduced-cost cell if one exists, and otherwise to the cheapest active cell (ties broken by the smallest secondary index). This rule is not a bookkeeping detail. Replaying the published seeds—which reproduces the reported MVM costs exactly—shows that at least one largest-penalty tie arises on 92.5% of the 482 small-scale instances (Wilson 95% CI $[89.8, 94.6]$) and on 100% of the 245 full-battery instances with $10 \leq n \leq 640$ (CI $[98.5, 100.0]$), that is, on essentially every instance once $n \geq 10$; the only tie-free draws occur at $n \in \{3, 4, 5\}$. Two

mechanisms explain this saturation: the diagnostic is triggered by a single tie among the $m + n - 1$ iterations, and on a reduced integer cost matrix many line penalties equal zero, so ties are the rule rather than the exception. Consequently, the tie-breaking order—not only the penalty criterion—is part of the specification of MVM, and any alternative order defines a different constructive heuristic. **Table 3** reports the rates by regime.

Table 3. Frequency of instances in which at least one largest-penalty tie occurs during MVM, with Wilson 95% confidence intervals. Obtained by replaying the published seeds, so the reported MVM costs are reproduced exactly.

Regime	N	≥ 1 tie (%)	Wilson 95% CI
Chapter-like schedule ($n \in \{3, \dots, 15\}$)	27	88.9	[71.9, 96.1]
Dense small ($n = 3, \dots, 15, 35$ reps)	455	92.7	[90.0, 94.8]
All small-scale instances	482	92.5	[89.8, 94.6]
Full battery, $n = 10$	35	100.0	[90.1, 100.0]
Full battery, $n = 20$	35	100.0	[90.1, 100.0]
Full battery, $n = 40$	35	100.0	[90.1, 100.0]
Full battery, $n = 80$	35	100.0	[90.1, 100.0]
Full battery, $n = 160$	35	100.0	[90.1, 100.0]
Full battery, $n = 320$	35	100.0	[90.1, 100.0]
Full battery, $n = 640$	35	100.0	[90.1, 100.0]
Pooled, $10 \leq n \leq 640$	245	100.0	[98.5, 100.0]

8.4. Unbalanced Transportation

On 280 runs with ratios 0.7 and 1.3, dummy-aware MVM always returns a feasible shipment whose total matches $\min(\sum a, \sum b)$. Internal certification occurs on 1.1% of instances—the same order of magnitude as in the balanced battery—suggesting that dummy elimination does not by itself make certificates rarer. A fuller comparison against [12] and [11] on published unbalanced batteries is left for a follow-up dedicated note; our focus here is scalability of the unified engine.

8.5. Transshipment

All 280 generated networks were solved successfully (**Table 4**, **Figure 3**). The mean classical-to-Floyd cell-count ratio is 6.4, confirming the dimensionality argument of [14]. Even the smallest tested shape already yields a ninefold cell-count reduction; the largest reaches sixteenfold. Wall-clock times remain modest because Floyd is cubic in $|V|$ with small constants for the sizes considered, and the subsequent TP is tiny.

Table 4. Transshipment: Floyd/Dijkstra compression versus classical all-node TP size (35 instances per size).

$ S $	$ T $	$ D $	Classical dim.	Reduced TP	Cell-count ratio	Mean time (s)
10	10	10	30×30	10×10	9.00	0.007

Continued

20	20	20	60 × 60	20 × 20	9.00	0.027
40	40	40	120 × 120	40 × 40	9.00	0.11
80	40	80	200 × 200	80 × 80	6.25	0.35
160	40	160	360 × 360	160 × 160	5.06	1.39
320	40	320	680 × 680	320 × 320	4.52	6.25
640	40	640	1320 × 1320	640 × 640	4.25	32.9
1000	40	1000	2040 × 2040	1000 × 1000	4.16	113.1

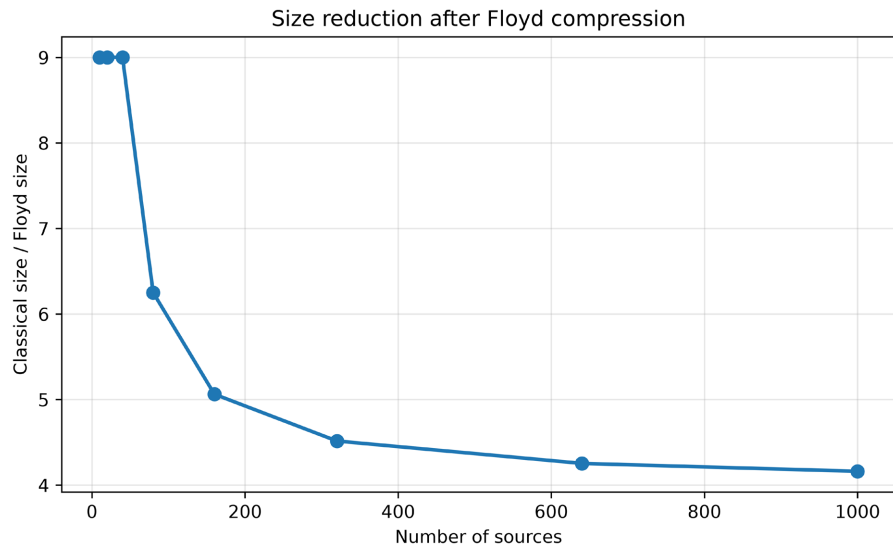


Figure 3. Mean classical-to-Floyd transportation cell-count ratio versus number of sources.

Independent validation of the compression. Reporting solution success and the cell-count ratio alone does not establish that the expand/backtrack step is implemented correctly, since a bug in path reconstruction would still return a well-formed shipment. We therefore add three independent checks on every transshipment instance within the LP cell budget. The first isolates the *reformulation* from the heuristic: because MVM solves the compressed problem only approximately, we compare the two *optima* rather than the MVM cost itself, requiring

$$|c_{\text{cmp}}^* - c_{\text{LP}}^*| \leq \varepsilon \max(1, c_{\text{LP}}^*), \quad \varepsilon = 10^{-6}, \quad (3)$$

where c_{cmp}^* is the LP optimum of the Floyd-compressed transportation problem on the shortest-path cost matrix and c_{LP}^* is the optimum of the classical all-node transshipment LP solved directly on the original network (a min-cost flow with uncapacitated arcs and node balances $a_k - b_k$). The tolerance is required because both quantities are computed in floating point. The second check validates the expansion itself: re-evaluating the reconstructed arc flows f on the original arc costs must reproduce the cost reported for the compressed solution,

$$\left| \sum_{(u,v) \in A} c_{uv} f_{uv} - c^{\text{cmp}} \right| \leq \varepsilon \max(1, c^{\text{cmp}}), \quad (4)$$

which fails as soon as a path is expanded along arcs other than those realising the shortest-path length. The third check verifies the *feasibility* of the reconstructed flow through the conservation residual at every node k ,

$$\rho_k = \left| \sum_{(k,v) \in A} f_{kv} - \sum_{(u,k) \in A} f_{uk} - (a_k - b_k) \right|, \quad (5)$$

where a_k and b_k are the supply and demand at k (both zero at transit nodes). Together, the three metrics separate a modelling error ((3) would fail), a path-reconstruction error ((4) would fail) and a bookkeeping error ((5) would fail).

Table 5 reports the outcome on the 105 transshipment instances with $|S||D| \leq 2500$. Both cost identities hold on every instance (100.0% and 100.0%, Wilson 95% CI [96.5, 100.0] for the first), and all conservation residuals are exactly zero rather than merely below tolerance, since supplies, demands and arc costs are integral and the expansion only adds integers. This supports Proposition 7 empirically and confirms that the $(|S| + |T| + |D|)^2 \rightarrow |S||D|$ reduction is obtained without loss of optimality.

Table 5. Independent validation of the Floyd compression against a direct all-node transshipment LP, on the transshipment instances within the LP cell budget. “Equivalence” is the fraction of instances satisfying (3); “Expansion” is the fraction satisfying (4); residuals ρ_k are defined by (5) and are exactly zero because all supplies, demands and arc costs are integral.

$ S = D $	N	Equivalence (%)	Expansion (%)	Mean ρ_k	Max ρ_k
10	35	100.0	100.0	0	0
20	35	100.0	100.0	0	0
40	35	100.0	100.0	0	0
Pooled	105	100.0	100.0	0	0

8.6. Asymmetric TSP

Table 6 and **Figure 4** report ATSP results. Across 280 instances, MVM-NNA improves upon NNA started at city 0 by 15.71% on average and is competitive with—often better than—a multi-start NNA limited to ten starts. The relative improvement is largest at moderate sizes (about 27% at $n = 20$) and decreases gradually as n grows (about 6.6% at $n = 1000$), while remaining strictly positive on average throughout the schedule. This suggests that penalty-based start selection systematically avoids a large class of poor initial cities, with a shrinking but still useful edge once many cities dilute the first-move effect.

Table 6. ATSP: mean tour costs (35 instances per size). Improvement is relative to NNA from city 0.

n	NNA (start 0)	MVM-NNA	Multi-start NNA	Mean improv. vs NNA ₀
10	249.7	197.5	197.2	18.7%

Continued

20	326.1	232.6	247.4	26.6%
40	374.7	292.5	307.5	20.2%
80	482.0	390.5	394.3	17.1%
160	609.8	504.5	520.5	16.6%
320	779.4	682.5	690.7	11.8%
640	1102.9	1011.2	1013.1	8.1%
1000	1462.7	1364.1	1370.7	6.6%

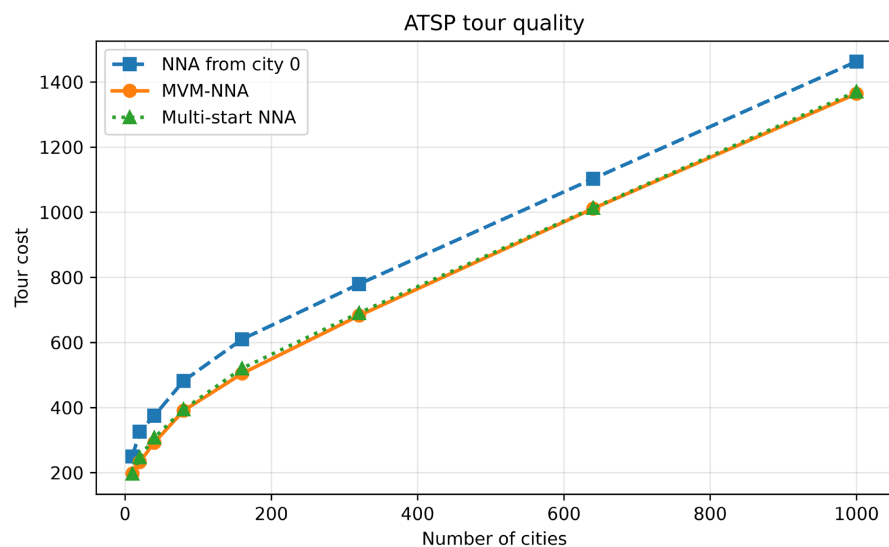


Figure 4. Mean tour cost of NNA (start 0), MVM-NNA and multi-start NNA versus n .

8.7. Capacitated Transportation

Table 7 summarises VAM-B. On 280 random bounded instances, the mean LP gap is 21.37%, with exact optimality on 18.1% of LP-checked cases (driven largely by the smallest sizes). Automatic box-potential certification after our current potential-repair heuristic is rare (0.0%)—an honest limitation: once capacities bind, the “all-zero” certificate essentially never fires, and constructing potentials that satisfy the three sign conditions is a genuine post-processing problem. The litmus Example 9 is recovered exactly (cost 32).

Table 7. Capacitated TP (VAM-B): mean LP gap (when available) and runtime (35 instances; $\approx 30\%$ cells bounded).

n	Mean LP gap	Mean $t_{\text{VAM-B}}$ (s)
10	3.83%	0.010
20	18.19%	0.029
40	42.09%	0.089
80	—	0.37
160	—	1.53

Continued

320	—	7.02
640	—	36.2
1000	—	109.1

8.8. Synthesis of Experimental Findings

1) **Small-scale concordance.** On [1]-sized LTPs, the 2015 qualitative conclusion is recovered: MVM is the better (or equal) constructive solution in the vast majority of draws, with mean LP gaps roughly half those of VAM on the chapter-like schedule.

2) **Evolution.** As n grows, mean-gap advantages persist and even widen on LP-checked sizes, while pathwise $\text{MVM} \leq \text{VAM}$ rates drift from the low-90% chapter-like level toward about 63.6% overall.

3) **Large-scale nuance.** Pathwise reversals become non-negligible for $n \geq 80$; internal certificates become rare; MVM remains a strong mean improver where LP evidence exists, but no longer a universal dominator.

4) **Extensions.** Floyd compression, MVM-NNA starts and VAM-B extend the same reduction-and-penalty idea beyond the balanced LTP treated in 2015.

9. Discussion, Limitations and Open Questions

9.1. Positioning within the VAM Literature

Experimental analyses of VAM variants [6] typically report solution quality on moderate batteries without internal optimality certificates tied to duality. MVM's contribution is not merely another tie-breaking rule on opportunity-cost matrices, but a reduction that coincides with dual potentials and a certificate theory that extends, carefully, to capacities. Our scalability evidence, while broader than [1], remains a laboratory study: costs are synthetic, networks are random, and industrial instance libraries would strengthen external validity.

9.2. Limitations of the Present Study

This paper presents the following limitations.

- **LP budget.** Instances with $mn > 2500$ are not LP-checked. Large-size quality is assessed relative to VAM (or multi-start NNA) rather than proven optima.
- **Implementation language.** Our pure Python implementation prioritizes clarity and testability over micro-optimization. Implementing a C++ or Julia port would reduce absolute execution times but would not alter the qualitative performance rankings at these scales.
- **Box potentials.** The post-processing mechanism used to certify Theorem 10 is heuristic. Consequently, a low automatic certification rate does not imply that VAM-B solutions are seldom optimal; rather, it indicates that our current approach cannot yet reliably construct the necessary witnessing potentials.
- **Literature batteries.** Dedicated re-runs of published unbalanced and VAM-var-

iant test sets [6] [12] would allow sharper head-to-head claims.

- **Capacitated Floyd compression.** The uncapacitated-arc hypothesis of Proposition 7 is essential rather than cosmetic: with arc capacities, several origin-destination pairs may route through a common arc, so shortest-path expansion can oversubscribe it and the compressed optimum is no longer attainable. Arc-capacitated transshipment therefore lies outside the present scope and is left for future work.

9.3. Open Mathematical Questions

This work leaves several open mathematical questions for future research:

- 1) Resolve the theoretical validity of the Hall-type characterization for one-pass sufficiency (Section 3) and its capacitated counterpart.
- 2) Bound the maximum number of local re-reduction rounds for MVM and VAM-B to verify if the worst-case behavior is bounded by $\min(m, n) - 1$.
- 3) Pair VAM-B with a bounded-variable stepping-stone or MODI correction to guarantee finite optimality convergence while leveraging dual-based warm starts.
- 4) Model MVM penalty start-city selection in random asymmetric ATSP instances to mathematically explain why the observed 16% average improvement peaks at moderate values of n .

10. Conclusion

This work extends the Modified Vogel Method (MVM) of [1], confirming at small scales that MVM outperforms classical VAM in nearly all instances ($92.6\% \text{ MVM} \leq \text{VAM}$) while providing a sharper, more reliable constructive start for transportation solvers. As problem size increases, the absolute “always better” claim is refined to a scale-aware statement, with mean gaps remaining favorable (20.01% vs. 28.63%) and overall pathwise dominance tapering to 63.6% (60.6% for $n \geq 80$). Key contributions including duality certificates, dummy-aware reduction, and capacitated VAM-B extend this primal-dual greedy discipline, with future integration of a bounded MODI phase aimed at completing the theoretical framework.

Acknowledgements

We thank colleagues involved in earlier works on unbalanced transportation and Floyd-compressed transshipment. All computational artifacts (models, instances, parameters, and results) are organized to support verification and extension.

Author Contributions

The original idea for this study was conceived by Y. Gningue. O. Koné and P.M. Takouda developed the numerical extensions of this idea and designed the computational framework. O. Koné carried out the implementation and performed all numerical computations. All authors contributed to the analysis and interpretation of the results, and to the writing and revision of the manuscript. All authors read and approved the final version.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Almaatani, D., Diagne, S.G., Gningue, Y. and Takouda, P.M. (2015) Solving the Linear Transportation Problem by Modified Vogel Method. In: Cojocaru, M.G., Ko-tsireas, I.S., Makarov, R.N., Melnik, R.V.N. and Shodiev, H., Eds., *Interdisciplinary Topics in Applied Mathematics, Modeling and Computational Science*, Springer International Publishing, 13-19. https://doi.org/10.1007/978-3-319-12307-3_3
- [2] Reinfeld, N.V. and Vogel, W.R. (1958) *Mathematical Programming*. Prentice-Hall, viii+274 p.
- [3] Charnes, A. and Cooper, W.W. (1954) The Stepping Stone Method of Explaining Linear Programming Calculations in Transportation Problems. *Management Science*, **1**, 49-69. <https://doi.org/10.1287/mnsc.1.1.49>
- [4] Dantzig, G.B. (1963) *Linear Programming and Extensions*. Princeton University Press.
- [5] Chopra, S. and Meindl, P. (2006) *Supply Chain Management: Strategy, Planning, and Operation*. 3rd Edition, Pearson Prentice Hall.
<https://www.amazon.com/Supply-Chain-Management-Sunil-Chopra/dp/0131730428>
- [6] Mathirajan, M. and Meenakshi, B. (2004) Experimental Analysis of Some Variants of Vogel's Approximation Method. *Asia-Pacific Journal of Operational Research*, **21**, 447-462. <https://doi.org/10.1142/s0217595904000333>
- [7] Diagne, S.G. and Gningue, Y. (2011) Méthode de Vogel modifiée pour la résolution des problèmes de transport simple. *Applied Mathematical Sciences*, **5**, 2373-2388.
<https://www.m-hikari.com/ams/ams-2011/ams-45-48-2011/sanghareAMS45-48-2011.pdf>
- [8] Burkard, R.E. (2007) Admissible Transformations and Assignment Problems. *Vietnam Journal of Mathematics*, **35**, 373-386.
- [9] Almaatani, D., Diagne, S., Gningue, Y. and Takouda, P.L. (2014) Solving the Unbalanced Transportation Problem Using a Modified Vogel Method. *Proceedings of the Operational Research Division of the Annual Conference of the Administrative Sciences Association of Canada*, Vol. 35, 85-95. <https://tinyurl.com/fmjt7hrc>
- [10] Shimshak, D., Alan Kaslik, J. and Barclay, T. (1981) A Modification of Vogel's Approximation Method through the Use of Heuristics. *INFOR: Information Systems and Operational Research*, **19**, 259-263. <https://doi.org/10.1080/03155986.1981.11731827>
- [11] Goyal, S.K. (1984) Improving VAM for Unbalanced Transportation Problems. *Journal of the Operational Research Society*, **35**, 1113-1114.
<https://doi.org/10.1057/jors.1984.217>
- [12] Balakrishnan, N. (1990) Modified Vogel's Approximation Method for the Unbalanced Transportation Problem. *Applied Mathematics Letters*, **3**, 9-11.
[https://doi.org/10.1016/0893-9659\(90\)90003-t](https://doi.org/10.1016/0893-9659(90)90003-t)
- [13] Ramakrishnan, C.S. (1988) An Improvement to Goyal's Modified VAM for the Unbalanced Transportation Problem. *Journal of the Operational Research Society*, **39**, 609-610. <https://doi.org/10.1057/jors.1988.101>
- [14] Almaatani, D., Diagne, S., Diop, E.M., Gningue, Y. and Takouda, P.L. (2014) Reduction of Transshipment Problems to Smaller Transportation Problems Using Minimum Cost Path. *Proceedings of the Operational Research Division of the Annual Confer-*

- ence of the Administrative Sciences Association of Canada, Vol. 35, 76-84.
<https://tinyurl.com/ws65xxk2>
- [15] Hillier, F.S. and Lieberman, G.J. (2001) Introduction to Operations Research. 7th Edition, McGraw-Hill.
<https://abdulkadir.blog.uma.ac.id/wp-content/uploads/sites/621/2020/01/Introduction-to-Operations-Research-Hillier-Lieberman.pdf>
- [16] Taha, H.A. (2007) Operations Research: An Introduction. 8th Edition, Pearson Prentice Hall, xx+813 p.
<https://www.abebooks.com/9780131889231/Operations-Research-Introduction-Taha-Hamdy-0131889230/plp>
- [17] Almaatani, D., Boudjellaba, H., Gningue, Y., Takouda, P.M. and Zhu, H. (2016) Improving the NNA for the Travelling Salesman Problem Using the MVM. *International Journal of Latest Research in Science and Technology*, **5**, 106-109.
<https://tinyurl.com/mw66ysre>
- [18] Bazaraa, M.S., Jarvis, J.J. and Sherali, H.D. (2010) Linear Programming and Network Flows. 4th Edition, John Wiley & Sons, xiv+748 p.
- [19] Winston, W.L. (2004) Operations Research: Applications and Algorithms. 4th Edition, Thomson Brooks/Cole.
<https://www.cengage.com/c/ebook-operations-research-applications-and-algorithms-4e-winston/9780357337769/>
- [20] Applegate, D.L., Bixby, R.E., Chvátal, V. and Cook, W.J. (2007) The Traveling Salesman Problem: A Computational Study. Princeton Series in Applied Mathematics. Princeton University Press, 608 p.