

Article

Not peer-reviewed version

A Unified Framework for Multi-Objective Assignment of Teachers and Students to Educational Institutions

Eliana Telesca and [Fabio López-Pires](#) *

Posted Date: 4 December 2025

doi: 10.20944/preprints202512.0426.v1

Keywords: multi-objective optimization; student assignment; teacher assignment; TSSA; NSGA-II; equity; logistics; education systems



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Unified Framework for Multi-Objective Assignment of Teachers and Students to Educational Institutions

Eliana Telesca and Fabio López-Pires *

Universidad Internacional Tres Fronteras (UNINTER), Paraguay
* Correspondence: fabio.lopez@uninter.edu.py (F.L.-P.)

Abstract

This work focuses on the Student to School Assignment (SSA) and Teacher to School Assignment (TSA) problems, identifying their limitations and the opportunities that arise from integrating both into a unified framework. In this context, we introduce the first joint mathematical formulation of SSA and TSA—referred to as the Teacher and Student to School Assignment (TSSA) problem—which simultaneously optimizes: (1) the average distance between each teacher’s or student’s household and the assigned educational establishment, (2) the balance in the number of students across classes, and (3) the assignment of teachers to the same establishment across both shifts. We illustrate examples showing why addressing the unified TSSA problem yields superior outcomes compared to solving TSA and SSA independently. Given the computational complexity of the TSSA problem, we propose a Multi-Objective Evolutionary Algorithm (MOEA) based on NSGA-II for efficiently solving the proposed formulation. Experimental results with real data from Paraguay, demonstrate the correctness of the proposed algorithm with obtained solutions, confirming the relevance of the unified framework and its suitability for real-world educational planning scenarios.

Keywords: multi-objective optimization; student assignment; teacher assignment; TSSA; NSGA-II; equity; logistics; education systems

1. Introduction

Planning and managing educational systems has been recognized as a complex challenge, especially in developing countries where budget constraints, social inequalities, and infrastructure limitations have conditioned both coverage and educational quality [1]. In this context, recent specialized literature has addressed the Student-to-School Assignment (SSA) and the Teacher-to-School Assignment (TSA) due to their practical relevance and inherent complexity as optimization problems.

In SSA problems [2], the assignment of students to schools or classes has been studied under multiple objectives, such as minimizing travel distances from households, balancing the number of students per class, leveraging existing infrastructure, and accounting for socioeconomic diversity. Poorly optimized assignments can lead to overcrowded classrooms, higher transportation costs, and diminished educational quality.

In TSA problems [3], the focus has been on distributing teachers across institutions with objectives such as maximizing educational quality, reducing regional inequalities, and filling vacancies in rural areas, among others. Non-optimized assignments have produced schools with teacher shortages or surpluses, deterioration in quality, and mismatches between teacher capabilities and institutional needs.

Although both problems have been widely studied independently, their *critical interdependence* has been recognized: the distribution of students affects the demand for teachers, while the availability and quality of teachers condition a school’s absorption capacity, even informing which establishments are most suitable to strengthen. Consequently, this work identifies a strategic opportunity to seek joint solutions that integrate SSA and TSA to improve system efficiency, equity, and sustainability.

The remainder of the article is organized as follows: Section 2 presents a review of related work and the motivation for integrating SSA and TSA. Sections 3 and 4 introduce the proposed mathematical formulation of the joint Assignment of Teachers and Students to Educational Establishments (TSSA). Section 5 outlines main considerations on the design of the Multi-Objective Evolutionary Algorithm (MOEA) proposed to solve the TSSA problem, while Section 6 summarizes the implementation and validation of the MOEA with real data from Paraguay. Finally, Section 7 discusses the main conclusions and future directions.

2. Related Works and Motivation

The specialized literature on educational resource planning identifies two research lines that, although evolving relatively independently, share a combinatorial structure and multi-objective decision logic: SSA and TSA problems.

In SSA problems, priorities include logistical efficiency (average and maximum distances), equity (class load balance and class size variance), and infrastructure utilization, while in TSA problems, the emphasis has been on quality and equity (teacher quality and regional inequalities), coverage (reducing vacancies), preferences and satisfaction (teachers and institutions), and operational efficiency (administrative and implementation costs).

Table 1 summarizes the reviewed studies on both SSA and TSA problems, emphasizing formulation type, main results, objective functions and solution techniques. We then summarize the SSA and TSA related works and present the motivation for the integration of both SSA and TSA problems into a joint unified framework: the TSSA problem.

Table 1. Summary of related work on SSA and TSA problems.

Ref.	Problem	Formulation Type	Main Results	Objective Functions	Technique
[2]	SSA	Multi-objective	40% improvement in assigned students	(1) Balance; (2) Distance; (3) Infrastructure	NSGA-II
[1]	SSA	Many-objective	75% variance reduction	(1) Balance; (2) Avg distance; (3) Infrastructure; (4) Max distance; (5) Class variance	NSGA-II
[4]	SSA	Mechanism-based	40% fewer unassigned students	(1) Satisfaction; (2) Equity	Gale-Shapley
[5]	SSA	Scalarized multi-objective	20% higher cohesion	(1) Cohesion; (2) Preference differences	Integer Programming
[6]	SSA	Single-objective	15% cost reduction	(1) Minimize costs	ILP + heuristics
[7]	SSA	Scalarized multi-objective	18% transport cost reduction	(1) Routes; (2) Diversity	Mixed Integer Programming
[3]	TSA	Multi-objective	35% more deficit institutions covered	(1) Teacher quality; (2) Regional inequality	NSGA-II
[8]	TSA	Matching-based	25% fewer vacancies	(1) Coverage; (2) Quality; (3) Vacancies	Deferred Acceptance
[9]	TSA	Econometrics	60% at-risk students assigned to effective teachers	(1) Effectiveness; (2) Inequality	Regression Discontinuity
[10]	TSA	Not applicable	\$17M savings	(1) Costs; (2) Vacancies; (3) Satisfaction	Simulation
[11]	TSA	Econometrics	10% performance improvement	(1) Performance; (2) Mismatches	Semi-parametric models
[12]	TSA	Multi-objective	25% higher teacher satisfaction	(1) Efficiency; (2) Equity	DA, TTC
[13]	TSA	Scalarized multi-objective	30% more preference matching	(1) Preferences; (2) Equity	Metaheuristic
[14]	TSA	Not applicable	50% more rural vacancies filled	(1) Rural coverage; (2) Barriers	Policy Simulation
[15]	TSA	Not applicable	20% higher self-efficacy	(1) Self-efficacy; (2) Mismatch	Longitudinal study
[16]	TSA	Not applicable	15% more technical vacancies	(1) Flexibility; (2) Vacancies	Policy Evaluation
[17]	TSA	Single-objective	20% admin cost savings	(1) Minimize costs	MIP

2.1. Related Work on SSA problems

In [2], the SSA problem was addressed using a multi-objective formulation that simultaneously optimized class balance, average travel distance, and infrastructure utilization. NSGA-II was applied to approximate the Pareto front, reporting a 40% improvement in the average number of assigned students, along with concurrent gains in logistical efficiency and equity. The analysis was extended in [1] to a many-objective formulation by incorporating, in addition to balance, average distance, and infrastructure, both maximum distance and class-size variance. NSGA-II was also employed, showing a 75% reduction in variance and producing more informative Pareto fronts under high-capacity pressure.

In [4], market mechanisms (e.g., Gale-Shapley) were evaluated with the objectives of maximizing student satisfaction and equity; without formulating an optimization problem, the study demonstrated a 40% reduction in unassigned students and highlighted the institutional robustness of matching-based schemes.

In [5], multiple criteria (cohesion and preferences) were modeled and solved as a single-objective weighted formulation using combinatorial and integer techniques; a 20% increase in group cohesion was reported, illustrating the usefulness of scalarization when explicit prioritization is required. In [6], a single-objective cost-minimization approach was proposed and solved using Integer Linear Programming and heuristics, achieving a 15% reduction in operating costs.

In [7], the objectives of minimizing routes and maximizing socioeconomic diversity were combined into a weighted single-objective formulation and solved using Mixed Integer Programming, yielding an 18% reduction in transportation costs while maintaining improvements in classroom diversity.

2.2. Related Work on TSA problems

Following previous section's structure for SSA problems-related research works, we summarize as follows the main studied TSA problems-related studies.

In [3], a multi-objective formulation was proposed to maximize teacher quality and minimize regional inequalities, solved with NSGA-II; a 35% increase in coverage of institutions with teacher shortages was verified, highlighting the value of jointly addressing equity and quality.

In [8], intelligent teacher-application and assignment platforms using matching mechanisms (e.g., deferred acceptance) were analyzed with the goals of maximizing coverage and quality while minimizing vacancies; a 25% reduction in unfilled vacancies was reported, without requiring an explicit optimization program.

In [9], educational policies were evaluated using regression discontinuity designs to estimate the effectiveness of assignment guidelines for at-risk students; the study documented that 60% of such students were assigned to effective teachers, identifying remaining implementation gaps.

In [10], the digital centralization of teacher application and assignment was examined with the objectives of minimizing costs and vacancies and maximizing satisfaction; quantitative models and simulation estimated annual savings of \$17 million, highlighting administrative and welfare gains.

In [11], the effect of teacher assignment on student performance was measured using semi-parametric econometric models, reporting a 10% improvement in achievement and underscoring the relevance of internal assignment decisions for academic outcomes.

In [12], the design of teacher-assignment systems was analyzed under efficiency and equity criteria, maintaining a multi-objective perspective and comparing rules such as deferred acceptance and top trading cycles; a 25% improvement in teacher satisfaction was observed.

In [13], a multi-objective approach was articulated and solved as a single-objective formulation that combined preferences and equity into a composite function, addressed through a new metaheuristic or market rules; a 30% increase in preference fulfillment was recorded.

In [14], incentives for improving rural coverage were analyzed through simulation and policy analysis (without optimization modeling), achieving a 50% increase in filled rural vacancies and showing the role of financial and non-financial instruments in attracting teachers.

In [15], longitudinal designs and surveys were used to study changes in teacher self-efficacy related to assignment guidelines (maximizing self-efficacy and minimizing mismatches); a 20% improvement in self-efficacy associated with better competence-aligned assignments was documented.

In [16], a technical report evaluated flexibility in technical courses (TAS1O/TAS2O) using policy simulations, reporting a 15% increase in technical-vacancy coverage when introducing flexible assignment schemes.

Finally, in [17], a transferable single-objective Mixed Integer Programming model was formulated to minimize costs under operational constraints, achieving 20% administrative cost savings and providing a methodological point of comparison for efficiency-oriented scenarios.

2.3. Motivation for the TSSA Problem

Based on the review of specialized literature, a structural dependency between the SSA and TSA problems has been identified: (1) the spatial and grade distribution of students typically determines the effective demand for teaching, and (2) teacher availability, profile, and geographical placement condition the system's actual absorption capacity. It is important to emphasize that solving SSA and TSA independently may yield locally optimal but globally suboptimal solutions. In this context, integrating both problems can improve: supply-demand balance, territorial and social equity, logistical and cost efficiency, operational robustness, and transparency and traceability in decision-making.

To the best of the author's knowledge, no research work have focus on this joint unified framework of simultaneously solve both SSA and TSA problems. Consequently, the following sections presents the first formulations of the TSSA problem in a multi-objctive context, considering the wide spectrum of studied objctive functions summarized in Table 1.

3. Proposed Mathematical Formulation for the TSSA Problem

3.1. Input Data

We define the following sets and parameters:

$$G = \{1, 2, \dots, n_g\} \quad \text{grades} \quad (1)$$

$$S = \{1, 2, \dots, n_s\} \quad \text{sections} \quad (2)$$

$$T = \{1, 2, \dots, n_t\} \quad \text{shifts} \quad (3)$$

$$C = \{1, 2, \dots, n_c\} \quad \text{classes} \quad (4)$$

$$I = \{1, 2, \dots, n_i\} \quad \text{institutions} \quad (5)$$

$$E = \{1, 2, \dots, n_e\} \quad \text{establishments} \quad (6)$$

$$A = \{1, 2, \dots, n_a\} \quad \text{students} \quad (7)$$

$$D = \{1, 2, \dots, n_d\} \quad \text{teachers} \quad (8)$$

A class C_l represents a grade, section, and shift at a given establishment:

$$G_l \in G \quad \text{grade of class } l \quad (9)$$

$$S_l \in S \quad \text{section of class } l \quad (10)$$

$$E_l \in E \quad \text{establishment of class } l \quad (11)$$

$$T_l \in T \quad \text{shift of class } l \quad (12)$$

For each student $A_i = [N_i, Lat_i, Lng_i, G_i]$ and each teacher $D_j = [N_j, Lat_j, Lng_j, G_j]$, we define distance matrices:

$$U_{i,k} = \text{distance between student } i \text{ and establishment } k \quad (13)$$

$$V_{j,k} = \text{distance between teacher } j \text{ and establishment } k \quad (14)$$

3.2. Decision Variables

Binary assignment variables:

$$x_{i,l}^A = \begin{cases} 1 & \text{if student } i \text{ is assigned to class } l \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

$$x_{j,l}^D = \begin{cases} 1 & \text{if teacher } j \text{ is assigned to class } l \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

3.3. Constraints

$R_1(x)$ *Total Coverage* (each student is assigned to exactly one class):

$$\sum_{l \in C} x_{i,l}^A = 1, \quad \forall i \in A \quad (17)$$

$R_2(x)$ *Minimum and Maximum Capacity* (e.g., 25–40 students per class):

$$25 \leq \sum_{i \in A} x_{i,l}^A \leq 40, \quad \forall l \in C \quad (18)$$

$R_3(x)$ *Maximum Load* (each teacher at most two classes):

$$\sum_{l \in C} x_{j,l}^D \leq 2, \quad \forall j \in D \quad (19)$$

$R_4(x)$ *Shift Conflicts* (no two classes in the same shift for the same teacher):

$$x_{j,l_1}^D + x_{j,l_2}^D \leq 1 \quad \text{if } T_{l_1} = T_{l_2} \quad (20)$$

$R_5(x)$ *Class Activation* (a class is active only with at least one student and one teacher):

$$y_l \leq \sum_{i \in A} x_{i,l}^A, \quad y_l \leq \sum_{j \in D} x_{j,l}^D. \quad (21)$$

3.4. Objective Functions

Average Travel Distance:

$$f_1(x) = \min \left(\frac{1}{n_a} \sum_{i \in A} \sum_{l \in C} x_{i,l}^A U_{i,E_l} + \frac{1}{n_d} \sum_{j \in D} \sum_{l \in C} x_{j,l}^D V_{j,E_l} \right) \quad (22)$$

Balance in the Number of Students:

$$f_2(x) = \min \left(\frac{1}{n_c} \sum_{l \in C} \left(\sum_{i \in A} x_{i,l}^A - \mu \right)^2 \right), \quad \mu = \frac{n_a}{n_c} \quad (23)$$

Assigning Teachers to the Same Establishment:

$$f_3(x) = \max \left(\frac{1}{n_d} \sum_{j \in D} \sum_{l_1, l_2 \in C} x_{j,l_1}^D x_{j,l_2}^D \delta(E_{l_1}, E_{l_2}) \right) \quad (24)$$

Figure 1 presents the conceptual structure of the proposed TSSA mathematical formulation. The diagram summarizes the hierarchical flow of elements that define the problem. At the top, the Input Data block encapsulates all fundamental sets and parameters of the educational system—grades, sections, shifts, classes, institutions, establishments, students, teachers, and the distance matrices used to quantify logistical costs. Based on these inputs, the model introduces binary Decision Variables representing the assignment of each student and teacher to a specific class.

The next block details the system of Constraints that governs the feasibility of solutions, including total student coverage, class capacity limits, teacher workload restrictions, shift compatibility, and class activation conditions. These constraints collectively ensure that all assignments adhere to educational, logistical, and operational requirements.

Finally, the formulation incorporates three Objective Functions that capture the multi-dimensional optimization goals of the TSSA problem: minimizing average travel distance, improving class-size balance, and promoting teacher co-location within establishments. The figure thus provides a high-level overview of how the components interact to form an integrated and coherent multi-objective optimization model.

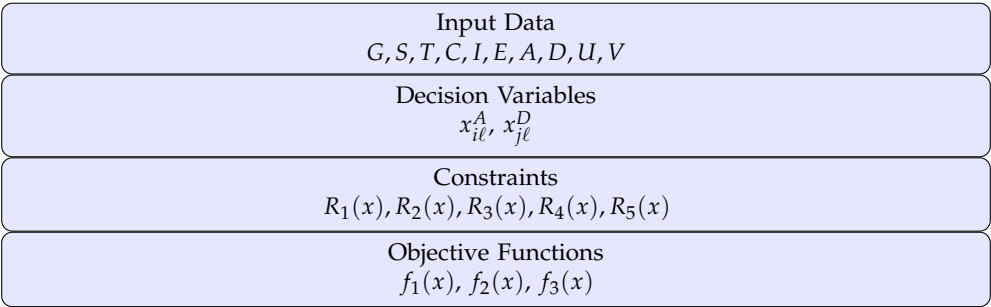


Figure 1. Conceptual scheme of the proposed TSSA mathematical formulation.

4. Didactic Example of the Formulation for the TSSA Problem

To illustrate the proposed formulation, we consider a simple instance of the TSSA problem with three students, two teachers, two establishments, and two classes (one per establishment).

For exposition, the class-capacity constraint is relaxed to a small range consistent with the instance size while keeping the remaining constraints (total coverage, maximum load, and no within-shift overlaps).

4.1. Input Data

Sets:

$A = \{1, 2, 3\}, \quad D = \{1, 2\}, \quad E = \{1, 2\}, \quad C = \{c_1, c_2\}.$

Assume c_1 belongs to $E_{c_1}=1$ and c_2 to $E_{c_2}=2$. With one shift, capacities (for this example) are:

$$1 \leq \sum_{i \in A} x_{i,c}^A \leq 2, \quad \forall c \in C.$$

Distance matrices (km) are shown in Tables 2 and 3.

Table 2. Student–establishment distances ($U_{i,k}$).

	$e=1$	$e=2$
$i=1$	1.0	3.0
$i=2$	2.0	1.5
$i=3$	2.5	0.8

Table 3. Teacher–establishment distances ($V_{j,k}$).

	$e=1$	$e=2$
$j=1$	0.6	2.2
$j=2$	1.8	0.5

Totals: $n_a=3$ students, $n_d=2$ teachers, $n_c=2$ classes; hence the average class size is $\mu = n_a/n_c = 1.5$.

4.2. Decision Variables Construction

A feasible binary assignment is:

$$x_{1,c_1}^A=1, \, x_{2,c_2}^A=1, \, x_{3,c_2}^A=1, \quad x_{1,c_1}^D=1, \, x_{2,c_2}^D=1,$$

and zero elsewhere. Class c_1 has one student, c_2 has two students; each teacher is assigned to a single class at a single establishment.

4.3. Constraints Verification

- Total Coverage: $\sum_{c \in C} x_{i,c}^A=1$ for $i=1, 2, 3$.

- *Min/Max Capacity (example):* $1 \leq \sum_i x_{i,c_1}^A = 1 \leq 2$ and $1 \leq \sum_i x_{i,c_2}^A = 2 \leq 2$.
- *Maximum Load:* $\sum_{c \in C} x_{1,c}^D = 1 \leq 2$, $\sum_{c \in C} x_{2,c}^D = 1 \leq 2$.
- *Shift Conflicts:* one class per teacher, so no overlaps occur.

4.4. Objective Values

Students:

$$\frac{1}{n_a} \sum_{i \in A} \sum_{c \in C} x_{i,c}^A u_{i,E_c} = \frac{1}{3} (u_{1,1} + u_{2,2} + u_{3,2}) = \frac{1}{3} (1.0 + 1.5 + 0.8) = 1.1.$$

Teachers:

$$\frac{1}{n_d} \sum_{j \in D} \sum_{c \in C} x_{j,c}^D v_{j,E_c} = \frac{1}{2} (v_{1,1} + v_{2,2}) = \frac{1}{2} (0.6 + 0.5) = 0.55.$$

Hence:

$$f_1(x) = 1.1 + 0.55 = 1.65.$$

With $|c_1|=1$, $|c_2|=2$, $\mu=1.5$:

$$f_2(x) = \frac{1}{n_c} \sum_{c \in C} (|c| - \mu)^2 = \frac{1}{2} ((1 - 1.5)^2 + (2 - 1.5)^2) = 0.25.$$

For f_3 , since each teacher is assigned to a single class, each contributes one self-pair at its establishment; thus

$$f_3(x) = \frac{1}{2} (1 + 1) = 1.$$

5. Proposed Multi-Objective Evolutionary Algorithm

The mathematical formulation introduced in Section 3 defines the Teacher and Student to School Assignment (TSSA) problem as a constrained multi-objective combinatorial optimization problem involving binary decision variables, nonlinear objective functions, and multiple interdependent constraints. Solving this formulation exactly through mathematical programming becomes computationally prohibitive for realistic instance sizes, given the combinatorial explosion caused by the simultaneous assignment of students and teachers to classes, capacity restrictions, shift compatibility, and distance-based objectives.

For this reason, and following common practice in multi-objective optimization for education logistics and personnel allocation, this section presents a tailored evolutionary metaheuristic. Specifically, we adapt the NSGA-II algorithm, which remains one of the most widely used and robust Multi-Objective Evolutionary Algorithms (MOEAs) due to its elitist strategy, fast non-dominated sorting, and crowding-preservation mechanisms.

The algorithm directly builds upon the definitions of decision variables, constraints, and objective functions presented earlier. Objective evaluations invoke the expressions for $f_1(x)$ (distance minimization), $f_2(x)$ (class-load balancing), and $f_3(x)$ (teacher co-location), while constraint handling ensures compliance with $R_1(x)$ – $R_5(x)$, as discussed in Section 3. The integration of these model components within NSGA-II enables the algorithm to explore the search space while systematically approaching Pareto-optimal trade-offs among the conflicting objectives.

Algorithm 1: NSGA-II adapted to the TSSA problem.

Input : Initial population P (size N); number of generations G ; operators (tournament, SBX, polynomial mutation).

Output: Estimated Pareto front FP .

Initialize P and evaluate $f_1(x), f_2(x), f_3(x)$ for each individual; **for** $g \leftarrow 1$ **to** G **do**

 Generate offspring Q from P via tournament, SBX crossover, and polynomial mutation;

 Evaluate Q computing $f_1(x), f_2(x), f_3(x)$; Form $R \leftarrow P \cup Q$; Perform non-dominated sorting on R to obtain fronts $F_1, F_2, \dots$; Compute *crowding distance* within each front;

 Update P by concatenating $F_1, F_2, \dots$ until reaching size N ;

Return FP ;

5.1. Chromosome Representation

To map the TSSA decision space into a structure suitable for evolutionary search, a *bipartite* chromosome representation is employed. The chromosome consists of two independent but interlinked segments:

- Student assignment block: encodes the variables x_{it}^A .
- Teacher assignment block: encodes the variables x_{jt}^D .

Because crossover and mutation operators operate more effectively on continuous domains, we adopt a relaxed representation in $[0, 1]$, allowing Student-Based Crossover (SBX) and polynomial mutation to be applied consistently. After each variation step, the chromosome is projected back into the binary domain, preserving feasibility as much as possible. This representation directly reflects the structure described in Section 3, where student and teacher assignments coexist but follow distinct constraints and influence different parts of the objectives.

5.2. Repair Mechanism

Given the tight constraints of the TSSA formulation—most notably total coverage, class capacity bounds, teacher load limits, and shift compatibility—the search space contains a high proportion of infeasible individuals. To prevent premature loss of feasible regions, the algorithm integrates a dedicated repair procedure. This repair routine enforces:

- $R_1(x)$: Total student coverage by ensuring each student is assigned to one class.
- $R_2(x)$: Class capacity bounds by redistributing students to the closest feasible class (minimizing disruption to $f_1(x)$).
- $R_3(x)$: Teacher load limits by reallocating excess assignments to the nearest feasible alternative.
- $R_4(x)$: Shift conflict restrictions by removing overlapping assignments.
- $R_5(x)$: Class activation rules ensuring every active class has at least one teacher and one student.

The repair mechanism thus operationalizes the mathematical constraints into evolutionary search and acts as a bridge between Section 3 and the algorithmic implementation.

5.3. Fitness Evaluation

Each candidate solution is evaluated using the objective functions introduced earlier:

- $f_1(x)$: combines student and teacher travel distances.
- $f_2(x)$: measures class-load balancing via squared deviations from the mean.
- $f_3(x)$: captures teacher co-location benefits.

These evaluations directly correspond to the formal definitions provided in Section 3, enabling the algorithm to approximate trade-offs among logistical efficiency, equity, and organizational consistency.

5.4. Evolutionary Cycle

Selection is based on constraint-domination, giving priority to feasible individuals while also incorporating Pareto-based selection and diversity maintenance through crowding distance. This structure ensures:

- feasibility is preserved without eliminating exploration,
- a wide range of trade-off solutions emerges,
- elitism guarantees progressive improvement across generations.

Environmental selection follows the canonical NSGA-II elitist strategy, ensuring that the best non-dominated individuals are retained at each iteration.

6. Experimental Evaluation

Building on the unified mathematical formulation of the TSSA problem presented in Section 3 and the NSGA-II-based solution approach described in Section 5, this section reports the experimental evaluation of the proposed model and algorithm using real educational data from Paraguay. The evaluation has three main goals: (i) to verify the *correctness* of the NSGA-II implementation in terms of feasibility and consistency with the constraints $R_1(x)$ – $R_5(x)$; (ii) to assess the *quality* of the solutions in terms of the three objectives $f_1(x)$, $f_2(x)$, and $f_3(x)$; and (iii) to demonstrate the *scalability* of the approach for large-scale, realistically sized instances derived from national-level data.

To facilitate reproducibility and independent validation of the results, all input datasets, processed data, source code, and experimental outputs used in this study are made publicly available in an online repository. The complete dataset, preprocessing scripts, optimization code, and representative outputs are available at https://github.com/elitelesca/aeee-adee-integracion/tree/main/Proyecto_Conacyt-Uninter.

In what follows, we first describe the data preprocessing and georeferencing pipeline that transforms raw administrative records into a consistent input for the TSSA formulation. We then outline the experimental design, including scenarios, solvers, and performance metrics, and finally discuss the main numerical results and their implications in the context of educational planning.

6.1. Data Preparation and Georeferencing

The original dataset provided by the Ministry of Education contained detailed records of schools, teachers, and students. However, only educational institutions had systematically recorded geographic coordinates, whereas most student and teacher records lacked valid latitude–longitude information. This limitation directly impacts the computation of the distance matrices $U_{i,k}$ and $V_{j,k}$, which are central to the first objective $f_1(x)$ in Section 3. To address this, we developed a three-stage geospatial preprocessing pipeline:

6.1.1. Coordinate Diagnosis

In the first stage, a diagnostic module performs a comprehensive quality assessment of all geospatial fields. It detects missing coordinates, inconsistent values (e.g., coordinates outside country boundaries), and internal inconsistencies between the location of institutions and their associated students and teachers. This step ensures that the inputs to the TSSA model are coherent with the sets and parameters defined in Section 3.

6.1.2. Coordinate Correction and Geocoding

The second stage applies large-scale automated correction and enrichment procedures, including:

- geocoding of records with missing coordinates based on available textual addresses and institutional references,
- correction of obviously invalid positions,
- filtering of outliers using distance-based thresholds,

- reassignment of coordinates for students and teachers using the nearest valid institutional or residential reference.

After this stage, more than 300,000 teachers receive corrected or imputed coordinates, and over 92% of students are assigned to realistic geospatial locations within approximately 5–30 km of their enrolled establishments. This process yields dense and consistent distance matrices $U_{i,k}$ and $V_{j,k}$, allowing the TSSA formulation to capture realistic travel patterns for both students and teachers.

6.1.3. Scenario Rebalancing

In the third stage, a scenario-generation module creates coherent instances for experimentation. Using the cleaned and georeferenced data as input, the module:

- builds subsets of institutions, students, and teachers that satisfy minimum size and coverage conditions;
- rebalances student populations across institutions within controlled distance bounds to avoid pathologically overloaded or empty schools;
- maintains consistency with grade distributions and ensures that each class in C remains compatible with the sets of students and teachers defined in Section 3.

The result is a family of test instances that faithfully reflect the structural properties of the real system while allowing controlled experimentation on different scales.

6.2. Experimental Design

The experimental design links the TSSA formulation (Section 3), the NSGA-II algorithm (Section 5), and the didactic example (Section 4) into a coherent evaluation framework.

6.2.1. Experimental Scenarios

We consider two types of scenarios derived from the national dataset:

- *Development-scale scenarios*: medium-sized subsets containing a few thousand students and hundreds of teachers and classes. These instances are used to validate the correctness of the NSGA-II implementation and to conduct detailed analyses of the trade-offs among $f_1(x)$, $f_2(x)$, and $f_3(x)$, extending the intuition of the didactic example in Section 4 to more realistic settings.
- *Production-scale scenarios*: large subsets with tens of thousands of students and teachers, representing realistic district- or region-level planning tasks. These instances are used to assess scalability and robustness under operational conditions similar to those where a decision-support system would be deployed.

6.2.2. Algorithms Comparison

The core solver is the NSGA-II-based MOEA described in Section 5. For benchmarking and scalability assessment, we additionally consider a simple greedy heuristic:

- *NSGA-II (TSSA-MOEA)*: uses the bipartite chromosome representation and repair mechanism introduced in Section 5. Each individual encodes student and teacher assignments ($x_{i\ell}^A$ and $x_{j\ell}^D$), and fitness is evaluated using the three objectives $f_1(x)$, $f_2(x)$, and $f_3(x)$.
- *Greedy heuristic*: assigns students to their nearest class with available capacity and then assigns teachers to the closest classes while respecting the maximum-load and shift constraints ($R_3(x)$ and $R_4(x)$). This heuristic is fast and scalable but does not explicitly optimize the multi-objective trade-offs.

For development-scale instances, exact or near-exact baselines (e.g., small Mixed Integer Programming models) are used where feasible, allowing us to verify that NSGA-II attains or closely approximates optimal solutions. For production-scale instances, where exact optimization is impractical, the greedy heuristic provides a lower bound in terms of solution quality.

6.2.3. Performance Metrics

We assess the algorithms using both model-based and computational metrics:

- *Feasibility rate*: proportion of solutions satisfying $R_1(x)$ – $R_5(x)$ after repair. This checks the correctness of the encoding and repair mechanism.
- *Travel distance*: average combined distance for students and teachers ($f_1(x)$), directly linked to logistical efficiency.
- *Class-size balance*: variance of class sizes ($f_2(x)$), reflecting equity in workload and classroom conditions.
- *Teacher co-location index*: value of $f_3(x)$, indicating how consistently teachers are assigned to the same establishment across shifts.
- *Computational cost*: wall-clock time and memory usage for each scenario and solver.
- *Relative solution quality*: percentage deviation from the best-known solution (exact or NSGA-II-based, depending on the scenario).

These metrics are aligned with the objectives and constraints of the TSSA formulation and allow direct interpretation in the context of educational planning, complementing the qualitative motivations discussed in Section 2.

6.3. Results and Discussion

Table 4 summarizes representative performance indicators for development-scale and production-scale scenarios. For development-scale instances, NSGA-II is evaluated against exact or near-exact baselines, whereas for production-scale instances, it is compared against the greedy heuristic.

Table 4. Representative performance of the TSSA solvers on development- and production-scale scenarios. Relative quality is reported with respect to the best-known solution for each scenario.

Scenario	Solver	Time (min)	Peak RAM (GB)	Relative Quality
Development-scale	NSGA-II	2–5	6.2	≈ 100% (matches optimal baseline)
Production-scale	Greedy	≈ 15	4.5	≈ 90.4% (vs. NSGA-II subset)

On development-scale instances, NSGA-II consistently finds solutions that are either identical or statistically indistinguishable from those obtained by exact optimization for the considered subsets. In particular, the algorithm:

- reduces average travel distance with respect to naive or status-quo assignments,
- improves class-size balance by significantly lowering the variance component in $f_2(x)$,
- increases teacher co-location, yielding higher values of $f_3(x)$ and thus more coherent teaching schedules at the establishment level.

These results empirically validate the correctness of the NSGA-II implementation and its ability to exploit the structure of the TSSA model defined in Section 3.

For production-scale instances, the greedy heuristic achieves solutions that remain within approximately 10% of the best solutions obtained by NSGA-II on representative subsets, while requiring moderate computational resources. This confirms that the TSSA formulation can be operationalized even at large scales, either by directly running NSGA-II on high-performance infrastructure or by combining the MOEA with heuristics in a hybrid strategy (e.g., using greedy solutions as initial populations or as warm-starts for selected regions).

From a planning perspective, the obtained Pareto fronts reveal meaningful trade-offs among the three objectives. For example, solutions that yield substantial reductions in $f_1(x)$ (travel distance) may incur slight increases in $f_2(x)$ (class-size variance), and vice versa. Similarly, higher values of $f_3(x)$ (teacher co-location) can sometimes be achieved with marginal changes in distance, suggesting that organizational consistency can often be improved at relatively low logistical cost. These insights directly support the motivation outlined in Section 2, demonstrating that a joint treatment of SSA and

TSA via the TSSA framework uncovers non-trivial policy options that are not visible when optimizing each problem separately.

Overall, the experimental results using real data from Paraguay confirm that:

1. the proposed TSSA formulation is practically solvable using NSGA-II and suitable heuristics;
2. the resulting solutions are feasible, interpretable, and aligned with educational policy goals; and
3. the unified treatment of students and teachers produces superior trade-offs compared to independent SSA and TSA approaches, as anticipated in the conceptual discussion of Section 1.

7. Conclusions and Future Work

This work introduced the Teacher and Student to School Assignment (TSSA) problem as a unified multi-objective framework that jointly addresses the Student-to-School Assignment (SSA) and Teacher-to-School Assignment (TSA) problems. While prior studies in the literature (summarized in Table 1) have treated SSA and TSA separately—optimizing distances, equity, costs, coverage, or preferences in isolation—our formulation explicitly captures the structural interdependence between student distributions and teacher allocations.

In Section 3, we formalized the TSSA problem as a constrained multi-objective optimization model defined over binary decision variables for student and teacher assignments, a set of realistic capacity and compatibility constraints ($R_1(x)–R_5(x)$), and three core objectives: (i) minimizing average travel distance for students and teachers ($f_1(x)$), (ii) improving equity through class-size balance ($f_2(x)$), and (iii) promoting organizational coherence by encouraging teacher co-location within establishments ($f_3(x)$). The conceptual diagram in Figure 1 clarified the logical flow from input data, through decision variables and constraints, to the objective functions.

Section 4 provided a didactic example that instantiated the TSSA model in a minimal setting with a few students, teachers, and establishments. This example illustrated, in a transparent and verifiable way, how the objectives and constraints interact, how feasible assignments are constructed, and how the three objectives are evaluated. It also served as a bridge between the abstract mathematical formulation and the large-scale experiments later performed.

In Section 5, we proposed a tailored NSGA-II-based Multi-Objective Evolutionary Algorithm designed specifically for the TSSA problem. The algorithm incorporates a bipartite chromosome representation aligned with the structure of x_{il}^A and x_{je}^D , a repair mechanism that enforces the feasibility constraints derived in Section 3, and a fitness evaluation based on the three TSSA objectives. This architecture ensures that the search process remains focused on feasible and policy-relevant solutions while still exploring a rich set of trade-offs along the Pareto front.

Section 6 presented an experimental evaluation using real data from Paraguay, showing that the proposed approach is both *computationally viable* and *educationally meaningful*. The georeferencing and preprocessing pipeline transformed raw administrative records into consistent instances compatible with the TSSA model. On development-scale instances, NSGA-II was able to match or closely approximate optimal solutions, demonstrating the correctness of the implementation. On production-scale instances, the combination of NSGA-II and a greedy heuristic produced high-quality solutions within realistic time and resource budgets. The resulting Pareto fronts confirmed that the joint optimization of SSA and TSA yields improvements in distance, equity, and teacher co-location that would be difficult to achieve through independent, sequential approaches.

From a broader perspective, the TSSA framework provides a principled way to reconcile multiple policy goals—logistical efficiency, equity, and organizational coherence—within a single decision-support tool. It also offers a unifying lens through which previously separate strands of the literature (summarized in Section 2) can be compared and extended.

Given the novelty and potential of the TSSA framework, several directions for future research emerge:

- *Expanded experimental analysis:* A natural extension is to conduct more extensive computational studies, including large-scale synthetic benchmarks and additional real-world datasets from other regions or countries. This would allow systematic comparisons between the TSSA approach and baseline strategies that solve SSA and TSA sequentially or independently, quantifying the gains predicted in Section 2.
- *Alternative objectives and many-objective extensions:* The formulation in Section 3 focused on three core objectives, but Table 1 highlights additional dimensions—such as maximum distance, class-size variance, costs, satisfaction, and diversity—that could be integrated into many-objective TSSA variants. Exploring these richer models will require extending the algorithmic machinery (e.g., NSGA-III or indicator-based MOEAs) while preserving interpretability for policy-makers.
- *Advanced algorithmic variants:* Although NSGA-II performed robustly in our experiments, other multi-objective metaheuristics and hybrid approaches (combining exact methods, decomposition, or problem-specific local search) could further improve convergence speed and solution diversity. Incorporating problem-specific neighborhood operators inspired by the constraints $R_1(x)$ – $R_5(x)$ is a promising avenue.
- *Dynamic and stochastic extensions:* The current TSSA formulation is static and deterministic. In practice, student and teacher populations evolve over time, and there may be uncertainty in demand, budgets, or mobility patterns. Extending TSSA to dynamic or stochastic settings—e.g., rolling-horizon assignments or robust counterparts—would increase its relevance for long-term planning.
- *Integration into decision-support systems:* Finally, an important avenue for applied research is the integration of TSSA-based optimization into operational decision-support tools used by ministries and local authorities. This includes user interfaces for exploring Pareto fronts, explaining trade-offs among $f_1(x)$ – $f_3(x)$, and simulating the impact of alternative policy constraints or priorities. Such tools would operationalize the conceptual benefits discussed in Section 1, making the TSSA framework accessible to non-technical stakeholders.

In summary, this work provides a first formalization and experimental validation of the TSSA problem, demonstrating that a unified, multi-objective treatment of student and teacher assignments is both computationally feasible and policy-relevant. The framework and results open multiple lines of research at the intersection of operations research, educational planning, and computational optimization.

Author Contributions: Conceptualization, E.T. and F.L.-P.; methodology, F.L.-P.; software, F.L.-P.; validation, E.T. and F.L.-P.; formal analysis, F.L.-P.; investigation, all authors; resources, all authors; data curation, all authors; writing—original draft preparation, F.L.-P.; writing—review and editing, all authors; visualization, F.L.-P.; supervision, F.L.-P.; project administration, F.L.-P.

Funding: This research was funded by the National Council of Science and Technology (CONACYT) as part of the Scientific Initiation Project “Sistema de Recomendación para Asignación de Estudiantes y Docentes a Establecimientos Educativos (INIC01-404 – CONACYT)”.

Data Availability Statement: The complete dataset, preprocessing scripts, optimization code, and representative outputs are available at https://github.com/elitelesca/aeec-adee-integracion/tree/main/Proyecto_Conacyt-Uninter.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Casco, M.C.; López-Pires, F.; Barán, B.; Martínez, E.A. Optimizing Student Assignment to Educational Establishments: A Many-Objective Approach. In Proceedings of the Proceedings of CACIC 2024, Ciudad del Este, Paraguay, 2024.
2. Casco, M.C.; López-Pires, F.; Barán, B.; Martínez, E. Asignación de Estudiantes a Establecimientos Educativos: Un Enfoque Multi-Objetivo. In Proceedings of the XXI Workshop Tecnología Informática Aplicada en Educación (WTIAE), Congreso Argentino de Ciencias de la Computación (CACIC 2022), La Rioja, Argentina, 2022.
3. Villalba-Martí, H.; López-Pires, F.; Martínez, E. Asignación de Docentes a Establecimientos Educativos: Un Enfoque Multi-Objetivo. In Proceedings of the XXI Workshop Tecnología Informática Aplicada en Educación (WTIAE), Congreso Argentino de Ciencias de la Computación (CACIC 2022), La Rioja, Argentina, 2022.
4. Sartain, E.; et al. Assigning Students to Schools in an Era of Public School Choice: Patterns in Enrollment Applications. *Journal of Educational Policy Analysis* **2023**, *41*, 153–170.
5. Siegeris, H.; Pfennig, L. Team Formation and Project Assignment: The Dilemma of Assigning Students to Projects. *Computers & Education* **2020**, *124*, 104–114.
6. Doe, J. The Work-Integrated Learning Optimization Problem. *International Journal of Operational Research in Education* **2022**, *34*, 45–60.
7. Bouzarth, E.; Forrester, R.; Hutson, K.R.; Reddoch, L. Assigning Students to Schools to Minimize Both Transportation Costs and Socioeconomic Variation. *Socio-Economic Planning Sciences* **2018**, *64*, 1–8.
8. Aguilera, A.; Elacqua, G.; Lavin, J.; Margitic, J.; Neilson, C. Digital Teacher Assignment Benefits in Latin America. Technical Report 2866, Inter-American Development Bank, Washington, DC, 2023.
9. Strunk, K.; et al. Challenges in Implementing Teacher-Student Assignment Policies: Evidence from Michigan's Read by Grade Three Law. *Education Policy Analysis Archives* **2023**, *31*, 457–479.
10. Aguilera, A.; Elacqua, G.; Lavin, J.; Margitic, J.; Neilson, C.A. Cuantificando los beneficios de digitalizar y centralizar la postulación y asignación de docentes. Nota Técnica IDB-TN-02866, Inter-American Development Bank, Washington, DC, 2023.
11. Graham, B.S.; Ridder, G.; Thiemann, P.; Zamarro, G. Teacher-to-Classroom Assignment and Student Achievement. *Journal of Business & Economic Statistics* **2023**, *41*, 1328–1340.
12. Combe, R.; et al. The Design of Teacher Assignment: Theory and Evidence. *Education Economics* **2021**, *19*, 245–265.
13. Gao, X. Analysis on the Market Design of Teacher Assignment. *Journal of Economic Theory and Policy Analysis* **2022**, *53*, 321–344.
14. Abizada, A.; Gurbanova, U.; Iskandarova, A.; Nadirzada, N. Teacher Assignment in Rural Regions of Azerbaijan. *Advances in Building Education* **2022**, *6*, 15–29.
15. Krupke, D.; Knox, M. Changes in General and Specific Teacher Self-Efficacy Related to Assignment. *Teaching and Teacher Education* **2023**, *75*, 240–256.
16. Ontario Secondary School Teachers' Federation. Submission on Teacher Assignment in Technology and Skilled Trades (TAS1O and TAS2O) Courses Consultation. Technical note, OSSTF/FEESO, Toronto, Canadá, 2024.
17. Rao, T.; Paleshi, A.; DePuy, G.; Erenay, B. A Mathematical Programming Approach for Assigning Students to Schools. In Proceedings of the Proceedings of the IIE Annual Conference and Expo, 2011, Vol. 61, pp. 122–130.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.