

[193006] Simheurística para el problema de VRP con demandas y tiempos estocásticos

David Felipe Gómez Rivero ^{a,c}, Sergio Andrés Forero Piraquive ^{a,c}, Santiago

Ordoñez Triana ^{a,c}, Juan Camilo Vanegas Rojas ^{a,c}

Eliana María González Neira, PhD. ^{b,c}

^aEstudiante de Ingeniería Industrial

^bProfesor, Director del Proyecto de Grado, Departamento de Ingeniería Industrial

^cPontificia Universidad Javeriana, Bogotá, Colombia

There are different types of vehicle routing problems (VRP) which consider, usually independently, different characteristics as: costs, quantity of product transported, customer demand, arrival time and attention time. Moreover, most of the academic literature consider that the parameters are all deterministic. This project considers a bi-objective VRP with stochastic demands and time windows. The objectives are total cost and tardiness. The demand of all clients must be satisfied by just one time visit.

In order to obtain a solution that gather all the specifications mentioned above, a simheuristic procedure was proposed. The simheuristic designed hybridizes the Non-Dominated Sorting Genetic Algorithm (NSGA-II) with a Montecarlo simulation. The NSGA-II allows obtaining the Pareto frontier of cost and tardiness objectives.

A deterministic mixed integer linear programming (MILP) model of the problem was formulated in a first stage and its results therefore simulated through a Montecarlo procedure in order to evaluate the performance of the proposed simheuristic. The MILP model was executed for small, medium, and large instances. For small instances, the model was executed until obtaining the optimal solution. In the case of medium and large instances, the MILP model was executed for a time limit of two hours. The simulation of the best obtained solutions for each objective function were performed with two probability distributions (uniform and lognormal) and four coefficients of variation (0, 0.10, 0.15 and 0.20).

The NSGA-II simheuristic was parametrized and then executed for all instances with the same two probability distributions and four coefficients of variation used in the simulation of MILP model solutions. The simheuristic results then were compared with the simulated results of the MILP model making possible to evidence the quality of the solution obtained by the simheuristic, and the time of execution until obtaining them. For small instances, in the deterministic case (coefficient of variation 0), the relative percentage deviation (RPD) was 0.26% for the costs objective and the absolute difference was 0.004 for the tardiness objective when compared with the solutions found by the MILP model. For medium instances in the deterministic case, the RPD for the cost objective function was 1.45% where MILP model had a better performance, but the simheuristic found better solutions for the tardiness objective function with an RPD -77.38%. In the case of big instances, the simheuristic was able to find a set of non-dominated solutions in a few minutes, while the MILP model could not find any feasible solution after 2 hours of execution. Finally, an experimental design was performed to identify the influence of probability distribution and coefficient of variation of the stochastic parameter in the results.

1. Justificación y planteamiento del problema

El problema de ruteo de vehículos (VRP) busca encontrar una combinación de rutas factible para un determinado vehículo, el cual hace entrega de bienes y servicios a un conjunto de clientes definidos (Baradaran, Shafaei, & Hosseinian, 2019). Este problema existe en diversas industrias relacionadas con: la entrega de materias primas a restaurantes o tiendas, el abastecimiento dinámico de combustible, la reparación y entrega de electrodomésticos en los domicilios, el transporte de prisioneros entre cárceles a tribunales, el transporte de muestras médicas de sangre desde los consultorios a los laboratorios y la recolección de basura y residuos de la ciudad hasta el lugar residual de basuras, entre muchos otros.

Las características que tradicionalmente consideran los estudios en VRP son: costos de ventas perdidas, costos de operación del camión y los costos relacionados con el transporte de productos, teniendo en consideración la cantidad de producto transportado. Por otro lado, también son considerados los tiempos de llegada para cada

cliente, las cantidades demandadas para cada cliente, su tiempo de servicio dependiendo de cada camión y el tiempo de viaje, para este caso desde un almacén a los minoristas y también su tiempo entre minoristas. Estas últimas variables en la vida real son variables aleatorias del sistema.

El VRP es un problema difícil de resolver, pues está clasificado como NP-duro (Lenstra & Kan, 1981). Por tanto, instancias medianas y grandes del problema VRP original y sus variantes tienen dificultad para ser resueltas por métodos exactos en tiempos computacionales razonables (Braekers, Hartl, Parragh, & Tricoire, 2016). Es por ello que se han utilizado métodos heurísticos y metaheurísticos para su solución.

De acuerdo con la revisión realizada por (Braekers et al., 2016) la literatura en VRP de los últimos años ha analizado múltiples variantes del problema que incluyen características reales. No obstante, de acuerdo con los hallazgos de los autores, esas características reales han sido frecuentemente consideradas de manera individual o en otros casos de manera simultánea con un número limitado de otras características. Por ello, los autores recomiendan enriquecer el análisis de problemas VRP considerando simultáneamente varias características reales y desarrollar métodos eficientes de solución para esos problemas complejos. Entre las características reales que deberían ser analizadas se encuentran la estocasticidad de la demanda y la estocasticidad de los tiempos de viaje. Al respecto se encontraron las variaciones del problema VRP, siendo la más estudiada el *Vehicle Routing Problem with Stochastic Demands* (VRPSD) (A. A. Juan, Faulin, Jorba, Caceres, & Marquès, 2013). También se encontraron trabajos sobre *Stochastic Vehicle Routing Problem with Time Windows* (SVRPTW) (Rajabi-Bahaabadi, Shariat-Mohaymany, Babaei, & Vigo, 2019), *VRP with Stochastic Demands and Dynamic Request* (VRPSDDR) (Sarasola, Doerner, Schmid, & Alba, 2016), *VRP with Stochastic Demands and Customers* (VRPSDC) (Balaprakash, Birattari, Stützle, & Dorigo, 2015), *VRP with Split deliveries and stochastic demands* (VRPSDSD) (Lei, Laporte, & Guo, 2012) y el *VRP with Stochastic Demands and Duration Constraints* (VRPSD – DC) (Erera, Morales, & Savelsbergh, 2010) (Goodson, Thomas, & Ohlmann, 2016).

De los trabajos mencionados es posible concluir que se ha estudiado el VRP estocástico con las siguientes características de manera independiente (separada): demandas estocásticas, tiempos de desplazamiento estocásticos, tiempos de servicio, ventanas de tiempo, flota de camiones homogéneas, flota de camiones heterogéneas, entregas divididas, multidepósito, reaprovisionamiento y recogidas y entregas. Adicionalmente, estos trabajos consideran sólo como función objetivo la minimización de costos, minimización de distancia, minimización de tiempos de entrega, maximización de utilidades, minimización de retrasos, minimización de fracción de la demanda no atendida, minimización de utilidades y minimización de desviaciones respecto al depósito. Sin embargo, ninguno de los trabajos analiza de manera simultánea las demandas y tiempos estocásticos con la minimización de dos objetivos: costos y tardanza.

El estudio de un VRP estocástico implica no sólo el uso de algún método metaheurístico para su solución sino la implementación simultánea de alguna metodología que permita tratar la estocasticidad del problema. Una de las metodologías en auge para la solución de problemas de optimización combinatorio estocástico son las simheurísticas (A. A. Juan, Faulin, Grasman, Rabe, & Figueira, 2015). Estas son la combinación de las metaheurísticas con simulación de Montecarlo para tratar la parte estocástica del problema. Por otra parte, considerando que se resolverá un problema bi-objetivo se pretende obtener la frontera de Pareto de los objetivos de minimización de costos y de tiempos de entrega con el uso de la metodología *Non-Dominated Sorting Genetic Algorithm-II* (NSGA-II) propuesta por (Deb, Agrawal, Pratap, & Meyarivan, 2000a). De esta manera, surge la pregunta de investigación: ¿cómo debe ser diseñada e implementada una simheurística en combinación con la metodología NSGA-II para un VRP con demandas y tiempos estocásticos y ventanas de tiempo para minimizar costos y tardanza?

2. Antecedentes

Antes de llegar a lo que hoy se conoce como '*Vehicle Routing Problem*' (VRP) existieron diferentes artículos que dieron un punto de partida al estudio y desarrollo de este tema. Por ejemplo, (Dantzig, Fulkerson, Johnson, Chvátal, & Cook, 1954) estudiaron y propusieron un método de solución al problema del agente viajero o *Travelling Salesman Problem* (TSP) que consiste en minimizar los costos o distancias recorridas por un único agente (vehículo) con capacidad infinita. Sin embargo, (Clarke & Wright, 1964) fueron los primeros que incorporaron más de un vehículo en la formulación del problema de transporte. Por esta razón este estudio puede ser considerado como el primero en la literatura del VRP como se conoce actualmente. En la década de

los 70's las publicaciones sobre VRP incrementaron, esta vez enfocadas a ambientes más realistas; dentro de las publicaciones que se pueden destacar están el ruteo de vehículos de servicio público (Marks & Striker, 1971), el ruteo de vehículos para la recolección de residuos sólidos (Liebman, 1970) y sistemas de Telebus (Wilson & Sussman, 1971). En 1983 Solomon añadió unas restricciones relacionadas con ventanas de tiempo a los problemas clásicos de VRP, generando diferentes configuraciones estáticas a este tipo de problemas. Para la década de los 90's, se incrementó el número de investigaciones y nuevos desarrollos de problemas relacionados con el VRP, esto como consecuencia del surgimiento de nuevas herramientas tecnológicas que permitieron introducir algoritmos de búsqueda más complejos y utilizar las meta-heurísticas para darle solución a estos problemas (Eksioglu, Vural, & Reisman, 2009).

A lo largo de estos años, desde su creación, se han estudiado diversas variaciones del VRP tales como '*Capacited vehicle routing problem*' (CVRP), '*Vehicle routing problem with time windows*' (VRPTW), '*Vehicle routing problem with stochastic demands*' (VRPSD), entre otros.

En el CVRP se tiene un número determinado de nodos para atender con un depósito inicial, donde se encuentran vehículos con una capacidad finita. El objetivo final es determinar una solución óptima de rutas en la cual cada vehículo empieza y termina en el depósito, cada cliente es visitado una única vez y, además, la capacidad del camión no es excedida (Sorensen & Marc, 2009).

El VRPSD, es un problema en el cual un grupo de clientes tienen demandas aleatorias que deben ser atendidas por una flota homogénea de vehículos saliendo desde un depósito, con todos los recursos disponibles (A. A. Juan et al., 2011). El objetivo es determinar una solución óptima de rutas que minimice los costos de acuerdo con restricciones como que las rutas deben empezar y terminar en el depósito, cada vehículo tiene una capacidad, la demanda estocástica de los clientes debe ser satisfecha, cada vehículo debe surtir a un solo cliente y un vehículo no debe parar dos veces en un cliente sin incurrir en una penalidad (A. A. Juan et al., 2013). Como parte de su investigación (Balaprakash et al., 2015) hacen uso de metaheurísticas como '*Random restart local search*', '*Iterated local search*', '*Memetic algorithm*' y '*Ant colony optimization*' para resolver el problema. Otra forma de abordar el VRPSD es estableciendo políticas de reaprovisionamiento para evitar que existan casos donde las demandas de los clientes no sean atendidas en su totalidad. Para este último caso, (Salavati-Khoshghalb, Gendreau, Jabali, & Rei, 2019) proponen la implementación de una adaptación de un "*Integer L-shaped algorithm*" y (Biesinger, Hu, & Raidl, 2018) proponen un algoritmo genético. Actualmente existen metaheurísticas muy potentes que pueden ser aplicadas para resolver este problema y así lograr un mejor acercamiento a las situaciones de industrias reales. (A. Juan et al., 2011) logran resolver un problema de VRPSD transformando el problema en pequeños problemas de CVRP usando una simulación de Montecarlo.

El VRPTW es una variante del VRP en la que los vehículos no pueden atender a los clientes antes de una hora determinada, y el servicio debe acabar antes de una hora máxima. A esto se le conoce como una ventana de tiempo. Estas se pueden clasificar en '*Soft Time Windows*' y '*Hard Time Windows*'. (Errico, Desaulniers, Gendreau, Rei, & Rousseau, 2018) explican que la principal diferencia entre estos dos tipos de ventanas de tiempo es que los servicios fuera de las ventanas '*Soft Time Windows*' incurrir en una penalización, mientras que los servicios fuera de las ventanas '*Hard Time Windows*' están prohibidos. Al igual que las demás variantes del VRP, este tipo de problema tiene en cuenta restricciones de capacidad de los camiones.

Un aspecto importante que se suele tener en cuenta en los problemas VRPTW son los tiempos de servicio. Estos son los tiempos que demora un camión en atender al cliente y completar el servicio antes de continuar con la ruta hasta la siguiente parada. Cuando el problema contempla ambas variables (tiempo de traslado y tiempo de servicio) como variables estocásticas, se le conoce como '*Stochastic Vehicle Routing Problem*' (SVRP) (Errico et al., 2018). En la literatura se encuentran diferentes aproximaciones a este problema utilizando modelos estadísticos e implementando "*Multi-Population Memetic Algorithm*" (Gutierrez, Dieulle, Labadie, & Velasco, 2018) o algoritmos un poco más complejos como "*Multi-Objective Memetic Algorithm*" y "*Multi-Objective Iterated Local Search*" (Miranda, Branke, & Conceição, 2018) obteniendo resultados factibles con el fin de reducir los costos totales.

(Baradaran et al., 2019) analizan una variante VRPTW el '*vehicle routing problem with multiple prioritized time windows*' (VRPMPTW) que consiste en una flota heterogénea con restricciones de intervalos de tiempo inciertos, es decir, que el intervalo de tiempo de entrega a cada cliente puede ser estricto o moderado. El nivel

de priorización define estricto como una única ventana de tiempo aceptada para la entrega y moderado como una ventana que puede ser cumplida en otro horario. Los autores solucionan el problema a partir de tres diferentes perspectivas. La primera hace uso de tiempos y costos determinísticos. En la segunda, se tiene una función bi-objetivo que minimiza los costos totales de transporte y maximiza la satisfacción del cliente. Finalmente, en la tercera perspectiva, plantean un modelo estocástico multi-objetivo donde los tiempos de viaje y los costos son variables aleatorias.

Existen otras variaciones del VRP que aumentan la complejidad del problema, pero que son coherentes con algunas situaciones reales que ocurren en la industria como el '*Vehicle routing problem with simultaneous pick and delivery*' (VRPSPD) y el '*VRPSDSD*'. El VRPSPD es un problema que no sólo considera la entrega de producto a cada cliente, sino que adicionalmente considera la recogida de producto que fue entregado al cliente anteriormente y que se considera obsoleto o que puede ser reutilizado (Lagos et al., 2018). Esta variación del problema aumenta la complejidad puesto que implica varias consideraciones a tener en cuenta sobre la capacidad de los vehículos al momento de realizar el ruteo. Para resolver este problema, autores como (Lagos et al., 2018) implementaron un 'Particle Swarm Optimization Algorithm' (PSO), mientras que (Montero, José Miranda-Bront, & Méndez-Díaz, 2017) resolvieron una versión simplificada del problema a partir de un '*Variable neighborhood search*' (VNS). Por otro lado, el VRPSDSD es una variante del VRP en la cual los clientes pueden ser atendidos por más de un vehículo, por lo que las entregas se pueden repartir entre diferentes vehículos que tengan un cliente en común dentro de su ruta. (Lei et al., 2012) proponen la heurística '*Neighbourhood Search*' para resolver este problema encontrando que el *Split Delivery* resulta ser beneficioso en varias ocasiones, reduciendo los costos de operación.

La gran mayoría de la literatura revisada sobre el VRP incluye una sección de 'Benchmarking', en la cual los métodos propuestos para la generación de soluciones son puestos a prueba con una gran cantidad de instancias diferentes para medir su eficiencia, y los resultados son comparados con información real de diferentes sectores de la industria. Esta parte de la investigación es de vital importancia, pues la investigación en algoritmos no exactos sin instancias de Benchmark o que no usan información de la industria en sus pruebas carecen de precisión y puede ser perjudicial al momento de la implementación en la industria (Rincon-Garcia, Waterson, & Cherrett, 2018).

3. Objetivos

Diseñar e implementar una simheurística para solucionar el problema de VRP con ventanas de tiempo, demandas estocásticas y tiempos de entrega estocásticos para la minimización de los costos y la tardanza.

Objetivos Específicos:

- Plantear un modelo matemático de programación lineal que permita resolver el problema VRP determinístico que incluye las características de ventanas de tiempo
- Diseñar una simheurística para resolver el problema VRP que incluye las características de ventanas de tiempo, demandas y tiempos estocásticos.
- Evaluar la efectividad de la simheurística contra la simulación de la solución óptima dada por el modelo matemático (para el caso determinístico) en instancias pequeñas.
- Evaluar la efectividad de la simheurística contra la simulación de la mejor solución encontrada después de dos horas de ejecución del modelo matemático (para el caso determinístico) en instancias grandes.

4. Modelo matemático del VRP (determinístico) con ventanas de tiempo

A continuación, se presenta el modelo matemático determinístico del VRP con ventanas de tiempo:

Conjuntos:

- I: Vehículos: {1...N}
- J: Nodos: {0...N}

Parámetros:

- cf : Costo fijo [\$] de usar un vehículo
- d_j : Demanda [unidades] del nodo $j \in J$
- t_{jk} : Tiempo de desplazamiento [minutos] para llegar del nodo $j \in J$ al nodo $k \in J$
- $wMin_j$: Hora mínima de llegada [minutos] al nodo $j \in J$
- $wMax_j$: Hora máxima de llegada [minutos] al nodo $j \in J$
- P : Penalización [\$] por entrega fuera de la ventana de tiempo
- k : Capacidad [unidades] de los vehículos
- M : Número muy grande.
- cv : Costo variable por tiempo recorrido por el camión [\$/min]

Variables de decisión:

- X_{ij} : 1 Si el vehículo $i \in I$ atiende al nodo $j \in J$, 0: d.l.c.
- Y_{ijk} : 1 Si el vehículo $i \in I$ se desplaza desde el nodo $j \in J$ hasta el nodo $k \in J$, 0 d.l.c.
- H_{ij} : Hora en la que el vehículo $i \in I$ atiende al nodo $j \in J$.
- A_{ij} : 1 Si el vehículo $i \in I$ atiende el nodo $j \in J$ dentro de la ventana de tiempo, 0 d.l.c
- F_{ij} : Tardanza total del vehículo $i \in I$ en el nodo $j \in J$.

Funciones Objetivo:

$$\begin{aligned} \text{Min Costo: } & \sum_{i \in I} (cf * X_{i0}) + \sum_{i \in I} \sum_{j \in J} (P * A_{ij}) \\ & + \sum_{i \in I} \sum_{j \in J} \sum_{k \in J} (cv_{ij} * Y_{ijk} * t_{jk}) \end{aligned} \quad (1)$$

$$\text{Min Tardanza: } \sum_{i \in I} \sum_{j \in J} F_{ij} \quad (2)$$

Restricciones:

$$\sum_{i \in I}^N X_{ij} = 1 \quad \forall j \in J/j \neq 0 \quad (3)$$

$$\sum_{j \in J}^N X_{ij} * d_j \leq K \quad \forall i \in I \quad (4)$$

$$A_{ij} \leq X_{ij} \quad \forall i \in I, \forall j \in J \quad (5)$$

$$H_{ik} \geq H_{ij} + t_{jk} - M(1 - Y_{ijk}) \quad \forall i \in I, \forall j \in J, \forall k \in J/k \neq 0 \quad (6)$$

$$\sum_{k \in J}^N Y_{iok} \leq 1 \quad \forall i \in I \quad (7)$$

$$\sum_{j \in J}^N Y_{ij0} \leq 1 \quad \forall i \in I \quad (8)$$

$$A_{ij} * wMin_j \leq H_{ij} \quad \forall i \in I, \forall j \in J/j \neq 0 \quad (9)$$

$$A_{ij} * M \geq (H_{ij} + 0.01) - wMax_j \quad \forall i \in I, \forall j \in J \quad (10)$$

$$H_{ij} \geq \sum_{k \in J}^N Y_{ijk} * wMin_j - M * (1 - X_{ij}) \quad \forall i \in I, \forall j \in J \quad (11)$$

$$\sum_{k \in J}^N Y_{ijk} = X_{ij} \quad \forall i \in I, \forall j \in J \quad (12)$$

$$\sum_{j \in J}^N Y_{ijk} = X_{ik} \quad \forall i \in I, \forall k \in J \quad (13)$$

$$F_{ij} \geq H_{ij} - wMax_j \quad \forall i \in I, \forall j \in J/j \neq 0 \quad (14)$$

$$Y_{ijj} = 0 \quad \forall i \in I, \forall j \in J \quad (15)$$

$$H_{ij} \geq 0 \quad \forall i \in I, \forall j \in J \quad (16)$$

La ecuación (1) es la función objetivo de costos que incluye los costos fijos y variables de transporte, así como el costo de penalización por violar la ventana de tiempo. La ecuación (2) minimiza el tiempo de llegada tarde a un cliente. El conjunto de restricciones (3) asegura que cada nodo debe ser atendido una única vez. La ecuación (4) permite atender la demanda de todos los nodos que cada vehículo visita. El conjunto de restricciones (5) tiene como objetivo establecer el cumplimiento de las ventanas de tiempo así el vehículo no atienda el nodo para evitar el costo de penalización. La ecuación (6) permite fijar la hora de atención de un cliente contemplando que la hora de atención del cliente debe ser mayor a la hora de atención de un nodo atendido anteriormente. La ecuación (7) asegura que todos los vehículos deben salir del nodo inicial. La ecuación (8) asegura que todos los vehículos deben volver al nodo inicial. El conjunto de restricciones (9) obliga a la variable A_{ij} al cero cuando no se cumple el límite inferior de la ventana de tiempo. Hace que se penalice cuando no se cumplen las ventanas de tiempo. El conjunto de restricciones (10) obliga a la variable A_{ij} al cero cuando no se cumple el límite superior de la ventana de tiempo. Hace que se penalice cuando no se cumplen las ventanas de tiempo. La restricción (11) asegura que la hora de atención al nodo debe ser mayor o igual al límite inferior de la ventana de tiempo del nodo. La restricción (12) asegura que si un vehículo incluye el arco (j, k) en la ruta, X_{ij} debe ser 1. Al igual que la anterior restricción, la restricción (13) asegura que si un vehículo incluye el arco (j, k) en la ruta, X_{ik} debe ser 1. El conjunto de restricciones (14) permite almacenar los valores de tardanza obtenidos en la solución. La restricción (15) evita que se pueda atender el mismo dos veces. Finalmente, el conjunto de restricción (16) evita la negatividad de la variable H_{ij} .

5. Simheurística para resolver el VRP estocástico con ventanas de tiempo

Para resolver el VRP con ventanas de tiempo, demandas estocásticas y tiempos de recorrido estocásticos se diseñó una simheurística que hibridiza la metaheurística NSGA-II con simulación de Montecarlo. El NSGA-II es un algoritmo genético creado específicamente para problemas de optimización multiobjetivo (Deb, Agrawal, Pratap, & Meyarivan, 2000b). La simheurística propuesta comienza con la generación de soluciones iniciales de manera aleatoria para la primera generación. Cada solución es una matriz que cruza los vehículos a enrutar, contra los clientes a atender. Las filas representan los vehículos y las columnas representan los clientes. El valor que se asigna a cada casilla de la matriz corresponde al orden en el que se atenderá el cliente dentro de la ruta de cada camión. Si el valor de la casilla es cero significa que ese camión no atiende al cliente que corresponde a la columna de la casilla. Dentro de la matriz no se contempla el nodo de origen (nodo cero) ya que se sobreentiende que todos los vehículos deben empezar y finalizar la ruta en este nodo. Sin embargo, el tiempo de desplazamiento hasta el nodo de origen y el regreso al mismo, es tenido en cuenta a la hora de calcular el

valor de las funciones objetivo. En la Figura 1, se muestra un ejemplo de una solución para una instancia del problema planteado en el que existen seis vehículos para atender seis clientes.

		Nodo o Cliente					
		1	2	3	4	5	6
Vehículo	1	0	2	0	3	1	0
	2	0	0	1	0	0	0
	3	0	0	0	0	0	1
	4	0	0	0	0	0	0
	5	1	0	0	0	0	0
	6	0	0	0	0	0	0

Figura 1. Ejemplo de solución.

Según lo explicado anteriormente, para el caso del primer vehículo, la ruta sería: 0-5-2-4-0, teniendo en cuenta que el nodo cero es el nodo de origen. En este mismo ejemplo, se puede ver que los vehículos 4 y 6 no tienen ningún nodo asignado en la ruta, indicando que estos vehículos no se usarían en esta solución.

Cada solución es evaluada con una simulación de Montecarlo, para la cual se generan números aleatorios que siguen una distribución de probabilidad determinada y sus parámetros son definidos para establecer las variables de tiempos de desplazamiento entre nodos y las demandas de cada nodo para cada una de las réplicas. En cada réplica de la simulación se calcula el valor de ambas funciones objetivo para cada solución. El número de réplicas de cada simulación se determina siguiendo la metodología propuesta por (Framinan & Perez-Gonzalez, 2015) en la que se realizan el número de réplicas necesarias hasta lograr un porcentaje de error determinado con un nivel de confianza específico. Para este caso en específico se determinó un porcentaje de error del 5% alrededor del valor esperado y un nivel de confianza del 95%. Una vez realizada la simulación, se calcula el valor esperado de ambas funciones objetivo para cada solución. En la Tabla 1, se muestra un ejemplo de la simulación de Montecarlo con ocho réplicas para una solución.

Réplica	Función Objetivo 1 (\$)	Función Objetivo 2 (min)
1	228.731	22.88
2	229.838	21.97
3	238.120	46.56
4	225.690	16.68
5	234.527	21.66
6	228.198	13.72
7	233.842	31.66
8	227.780	9.39
Valor Esperado	\$230.840,75	23,065 min

Tabla 1. Ejemplo de simulación Montecarlo.

Teniendo en cuenta el valor esperado de ambas funciones objetivo, se ordena todo el conjunto de soluciones vigentes y se clasifican en diferentes fronteras o grupos. Esto se conoce como el *Non-Dominated Sorting*. Las soluciones que hacen parte del primer grupo (primera frontera) son las soluciones no dominadas. Las soluciones no dominadas son aquellas que, al momento de compararlas con todas las demás soluciones vigentes, en ningún caso fueron dominadas por otra solución. Para el caso del problema planteado en específico, se dice que una solución S_1 domina a una solución S_2 cuando $Tardanza(S_1) \leq Tardanza(S_2)$ y $Costo(S_1) < Costo(S_2)$ o también cuando $Tardanza(S_1) < Tardanza(S_2)$ y $Costo(S_1) \leq Costo(S_2)$. En cualquier otro caso se puede decir que ninguna solución domina a la otra. Para ilustrar mejor el concepto de dominancia entre soluciones, se

muestra un ejemplo a continuación. Es importante tener en cuenta que en el ejemplo se busca minimizar las funciones objetivo.

Solución	Función Objetivo 1 (\$)	Función Objetivo 2 (min)
1	228.731	22,88
2	229.838	23,97
3	228.731	46,56
4	225.69	30,68

Tabla 2. Ejemplo de dominancia.

En el ejemplo de la Tabla 2, la solución 1 domina la solución 2 porque el valor de ambas funciones objetivo es menor en la solución 1. La solución 1 también domina la solución 3 porque a pesar de tener el mismo valor en la primera función objetivo, la solución 1 tiene un valor menor en la segunda función objetivo. Al comparar la solución 1 y la solución 4, se puede observar que ninguna solución domina a la otra, pues la solución 4 tiene menor valor en la primera función objetivo, pero la solución 1 tiene menor valor en la segunda función objetivo. Al realizar todas las comparaciones entre las soluciones, se puede decir que las soluciones 1 y 4 son soluciones no dominadas.

Las soluciones que hacen parte del segundo grupo son las soluciones no dominadas cuando no se tienen en cuenta las soluciones del primer grupo. Este mismo criterio se usa iterativamente para asignar todas las soluciones vigentes a un grupo. Cuando todas las soluciones han sido ordenadas, se descartan las últimas soluciones del ranking que excedan el número de soluciones deseadas o el tamaño de la población. Es decir, si actualmente se cuenta con 100 soluciones, y el tamaño de la población es 80, se descartan las últimas 20 soluciones del ranking. A continuación, se muestra un ejemplo de un *Non-Dominated Sorting* para un conjunto de diez soluciones vigentes (tamaño de población 10).

POS.	ID	GRUPO	FOC	FOT
0	52	1	239107	39.4191
1	73	1	186273	75.5554
2	75	1	240996	38.7118
3	82	1	186527	47.2512
4	58	2	187206	65.7839
5	80	2	240046	41.6576
6	86	2	187998	49.0359
7	12	3	243615	45.6184
8	30	3	188525	102.778
9	59	3	189217	86.9017

Figura 2. Ejemplo de Non Dominated Sorting.

En la Figura 2 se muestran los resultados del *Non-Dominated Sorting*. La primera columna corresponde a la posición de la solución dentro del conjunto de soluciones. La segunda columna corresponde al número ID para identificar la solución evaluada. La tercera columna corresponde al grupo al que fue asignada la solución. La cuarta columna muestra el valor de la función objetivo de minimización de costos y la quinta columna indica el valor de la función objetivo de minimización de tardanza. En este caso se puede observar que, si no se tienen en cuenta las soluciones del grupo 1, las soluciones no dominadas son las pertenecientes al grupo 2. Lo mismo aplicaría para el grupo 3 en caso de tener más grupos sin tener en cuenta las soluciones de los grupos 1 y 2. En este punto del proceso se evalúa si se han generado el número de generaciones deseadas, o el número de

iteraciones que debe realizar la metaheurística. Si el criterio de parada se cumple, el proceso termina aquí. De lo contrario, se procede con los procesos que se explican a continuación.

Una vez se han ordenado y clasificado las soluciones en los diferentes grupos, y se desean generar más generaciones de soluciones, se realiza una estimación de densidad o *density estimation*. Esta estimación consiste en calcular un valor que indica qué tan poblada se encuentra la región en la cual se ubica una solución en el plano cartesiano según las coordenadas dadas por los valores de sus funciones objetivo. Este valor calculado se conoce como el *crowding distance* y se calcula para cada solución vigente. Para este estudio se utilizó el método para calcular el *crowding distance* propuesto por (Fu & Wen, 2017). Para realizar este cálculo se deben organizar las soluciones según el valor de la primera función objetivo. Una vez ordenadas las soluciones, se debe calcular la diferencia entre el valor de las funciones objetivo de la solución inmediatamente anterior y la solución inmediatamente posterior a la solución a la que se le está calculando el *crowding distance*. Este valor será denotado como d_1 . También se debe calcular la diferencia entre la primera y la última solución del ranking siguiendo este mismo orden. Este valor será denotado como F_1 . Estas mismas dos operaciones se realizan nuevamente, pero esta vez ordenando las soluciones según el valor de la segunda función objetivo, encontrando así los valores d_2 y F_2 *respectivamente*. Una vez se han calculado estos valores, se aplica la siguiente fórmula (17):

$$crowding\ distance = \frac{d_1}{F_1} + \frac{d_2}{F_2} \quad (17)$$

A continuación, se muestra un ejemplo del cálculo del *crowding distance* de la solución ID:52 con respecto a las soluciones pertenecientes al grupo 1 tomando la función objetivo de costo como la primera función objetivo y la función objetivo de tardanza como la segunda función objetivo.

ID	Función Objetivo 1 (\$)	Función Objetivo 2 (min)
73	186,273	75.555
82	186,527	47.251
52	239,107	39.419
75	240,996	38.712

Tabla 3. Ejemplo de soluciones ordenadas según su Función Objetivo Costos.

ID	Función Objetivo 1 (\$)	Función Objetivo 2 (min)
75	240,996	38.712
52	239,107	39.419
82	186,527	47.251
73	186,273	75.555

Tabla 4. Ejemplo de soluciones ordenadas según su Función Objetivo Tardanza.

$$\begin{aligned}
 d_1 &= 240,996 - 186,527 \\
 F_1 &= 240,996 - 186,273 \\
 d_2 &= 47.251 - 38.712 \\
 F_2 &= 75.555 - 38.712 \\
 crowding\ distance_{ID:52} &= \frac{240,996 - 186,527}{240,996 - 186,273} + \frac{47.251 - 38.712}{75.555 - 38.712} \\
 crowding\ distance_{ID:52} &= 1.227
 \end{aligned}$$

En la Figura 3, se muestra un ejemplo gráfico de las diferencias utilizadas en el cálculo del *crowding distance*. Se evidencian, en la figura, las cuatro soluciones pertenecientes al grupo 1 ubicadas según el valor de sus funciones objetivo. Las líneas de color rojo representan el valor de las diferencias F_1 y F_2 , mientras que las líneas de color verde representan el valor de las diferencias d_1 y d_2 .

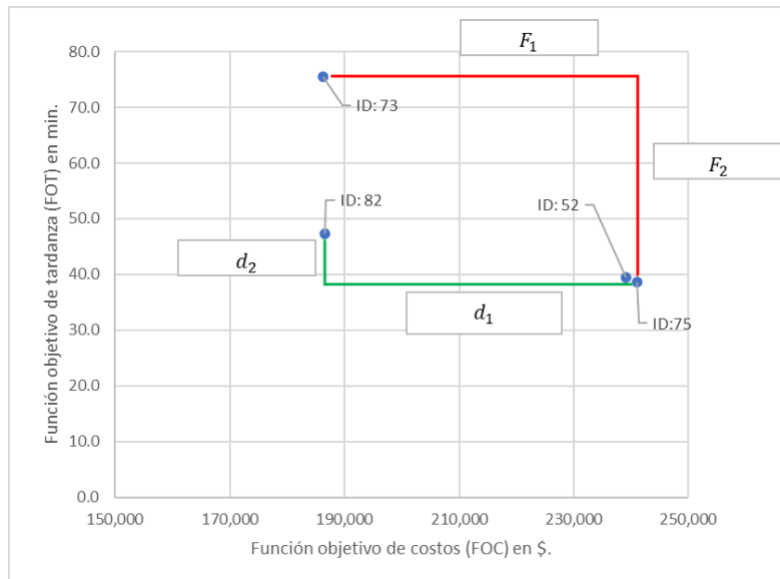


Figura 3. Ejemplo del cálculo del *crowding distance* de la solución ID:52

Luego de realizar esta estimación, se asigna una probabilidad de reproducción (de ser escogido como padre de una nueva solución) a cada solución basada en el grupo al que pertenece y al valor de su *crowding distance*. En este orden de ideas, a la solución con menor número de grupo, y mayor *crowding distance* se le asignará una probabilidad de reproducción mayor. Siguiendo con el ejemplo anterior, todas las soluciones del grupo 1 tendrán una probabilidad de reproducción mayor que las soluciones de los grupos 2 y 3. Adicionalmente, la solución ID: 82 tendrá mayor probabilidad de reproducción que la solución ID: 52 debido a que tiene mayor *crowding distance*, a pesar de pertenecer al mismo grupo. En el caso de las soluciones ID:73 e ID:75, al ser las soluciones con mayor FOT y FOC, respectivamente, se les asignará una probabilidad aún mayor que las otras dos soluciones ya que se desea tener una mayor probabilidad de que las soluciones de las siguientes generaciones tengan valores que amplíen más la frontera de soluciones. Una vez se ha hecho la estimación de densidad en todos los grupos, se eligen dos soluciones de manera aleatoria teniendo en cuenta su probabilidad de reproducción para ser las soluciones que se van a cruzar (ser soluciones padre), generando así una nueva solución. El cruce consiste en que la solución hijo hereda los valores exactamente iguales a los de la primera mitad de la matriz solución del padre 1, teniendo en cuenta todos los camiones (todas las filas). La segunda mitad de la matriz de la solución hijo se llena teniendo en cuenta los valores que hacen falta para completar una ruta continua, y a partir del orden en el que se asignaron los nodos de la segunda mitad del padre 2. Este procedimiento se repite para el número de nuevas soluciones (soluciones hijo) que se deseen generar. A continuación, se muestra un ejemplo de un cruce entre dos soluciones padre para generar una solución hijo.

En el ejemplo que se presenta en la Figura 4 se puede ver que el único valor que el hijo hereda en la ruta del primer camión del padre 1 es el número 2 en el segundo cliente. El único valor que debería heredar el hijo del padre 2 para la ruta del primer camión debería ser el número 3 en el cliente número 4. Sin embargo, como el número 1 no fue heredado del padre 1, se cambia el único valor heredado del padre 2 para garantizar una ruta continua y que no hayan “eslabones perdidos” en la ruta, dando como resultado la ruta: 0-4-2-0. En el caso del

segundo camión se hereda el número 1 en el tercer cliente del padre 1, y se debería heredar el número 1 en el quinto cliente del padre 2, sin embargo, heredar el mismo número dos veces genera una inconsistencia en la asignación de la ruta, por esta razón se cambia el valor heredado del padre 2 y se asigna el siguiente valor que no ha sido utilizado para continuar con la asignación de la ruta. Esto mismo ocurre en el caso del quinto camión.

		Padre 1						Padre 2						
		Nodo o Cliente						Nodo o Cliente						
Vehículo		1	2	3	4	5	6		1	2	3	4	5	6
	1	0	2	0	3	1	0	1	1	0	2	3	0	0
	2	0	0	1	0	0	0	2	0	0	0	0	1	0
	3	0	0	0	0	0	1	3	0	1	0	0	0	0
	4	0	0	0	0	0	0	4	0	0	0	0	0	0
	5	1	0	0	0	0	0	5	0	0	0	0	0	1
	6	0	0	0	0	0	0	6	0	0	0	0	0	0

		Hijo					
		Nodo o Cliente					
Vehículo		1	2	3	4	5	6
	1	0	2	0	1	0	0
	2	0	0	1	0	2	0
	3	0	0	0	0	0	0
	4	0	0	0	0	0	0
	5	1	0	0	0	0	2
	6	0	0	0	0	0	0

Figura 4. Ejemplo de cruce

Luego de realizar el cruce, algunas soluciones son seleccionadas para mutar siguiendo una probabilidad de mutación asignada a todas las soluciones. La mutación a la que son sometidas las soluciones hijo seleccionadas consiste en intercambiar un par de columnas de la matriz solución seleccionadas aleatoriamente para cambiar el resultado de las funciones objetivo de dicha solución y poder generar soluciones que diversifiquen la frontera de soluciones. En la Figura 5 se muestra el ejemplo de una mutación de una solución con la misma solución hijo que se mostró anteriormente.

		Hijo					
		Nodo o Cliente					
Vehículo		1	2	3	4	5	6
	1	0	2	0	1	0	0
	2	0	0	0	0	2	1
	3	0	0	0	0	0	0
	4	0	0	0	0	0	0
	5	1	0	2	0	0	0
	6	0	0	0	0	0	0

Figura 5. Ejemplo de mutación.

Una vez se han generado las nuevas soluciones, se les aplica la simulación de Montecarlo para encontrar el valor esperado de ambas funciones objetivo de cada solución y se repite todo el procedimiento hasta cumplir con un número de iteraciones establecido. En la Figura 6, se presenta un diagrama de flujo que resume el proceso desarrollado por la simheurística.

6. Evaluación de la simheurística con respecto al modelo matemático en instancias pequeñas

En esta sección se presentan los resultados del modelo matemático determinísticos (coeficiente de variación 0) y simulados teniendo en cuenta dos distribuciones de probabilidad (LogNormal y uniforme) y tres coeficientes de variación (0.1, 0.15 y 0.2) para instancias pequeñas y se evalúa la simheurística para esas mismas instancias con el fin de comparar los resultados de las funciones objetivo en ambos casos.

Se realizaron pruebas del modelo matemático con 18 instancias pequeñas de diferentes tamaños. La instancia más grande de las clasificadas como “pequeñas” para estas pruebas contemplaba 8 clientes y 3 vehículos para atenderlos. Se calculó la diferencia porcentual entre los resultados obtenidos por la simheurística con coeficiente de variación 0 (que es la simheurística para el caso determinístico) y los resultados arrojados por el modelo matemático en Gusek. Los resultados se muestran en la Tabla 5.

Para la evaluación del desempeño entre el modelo y la simheurística fue de suma importancia cuantificar los indicadores ‘relative percentage deviation’ (RPD), para la función objetivo de costos, y la diferencia absoluta (DA) para la función objetivo de tardanza. Estos indicadores se realizaron de la siguiente manera.

$$RPD \text{ determinístico} = \frac{\text{CostoDeterminístico}_{simheurística} - \text{Costo}_{modelo}}{\text{Costo}_{modelo}} \quad (18)$$

$$RPD \text{ estocástico} = \frac{\text{CostoEsperado}_{simheurística} - \text{CostoEsperado}_{modeloSimulado}}{\text{CostoEsperado}_{modeloSimulado}} \quad (19)$$

$$DA = | \text{Tardanza}_{simheurística} - \text{Tardanza}_{modelo} | \quad (20)$$

Las ecuaciones (17) y (18) permiten calcular la desviación porcentual relativa (RPD) de los costos para el escenario determinístico y estocástico entre el modelo (simulado, determinístico) y la simheurística. La ecuación (19) captura la magnitud de la diferencia entre la tardanza obtenida del modelo y la tardanza obtenida de la simheurística.

Los resultados muestran que la simheurística encuentra, para el caso determinístico del problema, resultados muy cercanos al óptimo encontrado en la mayoría de las instancias pequeñas, y existe una diferencia relativamente baja en comparación con los resultados de las soluciones óptimas del problema. Para el caso de los costos el gap promedio es de 0.26% y para la tardanza es de 0.004 horas (14.4 segundos).

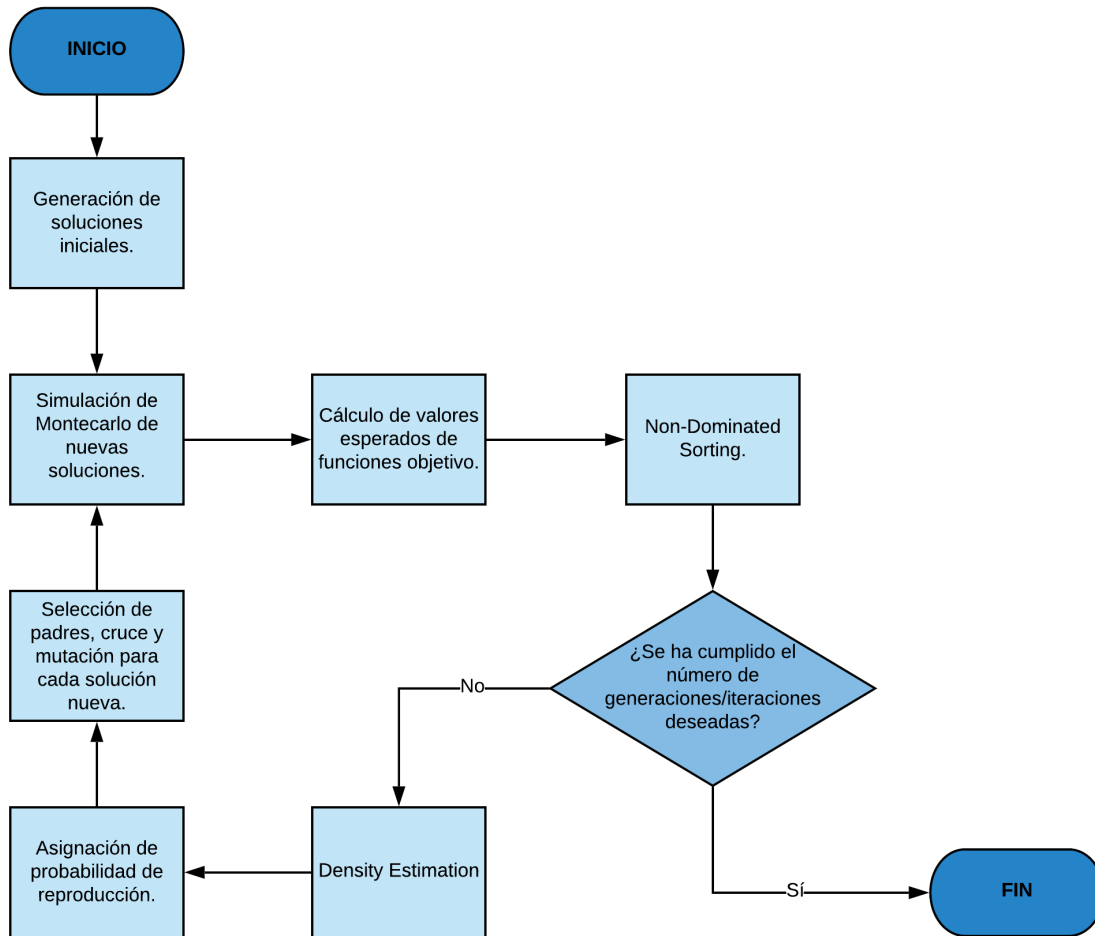


Figura 6. Resumen del proceso de la simheurística diseñada.

Posteriormente, se evaluó el desempeño de la simheurística para el caso estocástico. Por tanto, se simularon las soluciones óptimas arrojadas por el modelo matemático de 6 instancias pequeñas (tamaños de 6 y 8 clientes), para cada una de las dos funciones objetivo, teniendo en cuenta tres coeficientes de variación diferentes y dos distribuciones de probabilidad. Es importante tener en cuenta que las soluciones dadas por el modelo matemático, al ser simuladas posteriormente no necesariamente dan soluciones factibles, dependiendo de las variaciones de la demanda. En tanto que la simheurística está programada para sólo tener en cuenta las soluciones que al ser simuladas siempre son factibles. Por esta razón, se calculó el porcentaje de réplicas donde las soluciones del modelo matemático resultaron factibles para cada simulación. Para poder analizar y concluir sobre estos resultados es importante tener en cuenta que las soluciones arrojadas por la simheurística son factibles en todas las réplicas de cada simulación. Los resultados de estas simulaciones fueron comparados con los resultados de las dos soluciones extremo de la frontera de Pareto dadas por la simheurística para las mismas instancias y los mismos coeficientes de variación. Adicionalmente, se calculó el porcentaje de casos en los que

las simulaciones de la simheurística tuvieron un mejor desempeño que las simulaciones de las soluciones del modelo matemático. En la Tabla 6 y en la Tabla 7 se muestran los resultados.

Modelo Determinístico - Instancias Pequeñas									
Instancia	Número de clientes	nCamiones	Capacidad	FOC (\$) Modelo Matemático	FOT (h) Modelo Matemático	FOC (\$) Heurística	FOT (h) Heurística	RPD FOC	DA FOT
1	3	2	70	57,195.25	0	57,195.50	0	0.00%	-
2	4	2	70	58,172.25	0	58,172.40	0	0.00%	-
3	5	2	70	62,418.75	0	62,418.80	0	0.00%	-
4	6	2	70	68,206.25	0	68,206.30	0	0.00%	-
5	7	2	70	109,010.25	0.35	109,010.30	0.43	0.00%	0.08
6	8	2	70	110,011.75	0.85	110,047.00	0.85	0.03%	- 0.00
7	3	2	70	52,862.00	0	52,862.00	0	0.00%	-
8	4	2	70	58,908.00	0	58,908.10	0	0.00%	-
9	5	2	70	104,757.25	0	104,757.30	0	0.00%	-
10	6	2	70	104,805.50	0	104,805.50	0	0.00%	-
11	7	2	70	110,841.75	1.97	110,977.00	1.97	0.12%	- 0.00
12	8	2	70	110,851.00	1.98	111,552.00	1.98	0.63%	0.00
13	3	2	70	57,354.25	0	57,354.30	0	0.00%	-
14	4	2	70	63,479.75	0.23	63,479.80	0.23	0.00%	0.00
15	5	2	70	109,151.00	0.23	109,151.00	0.23	0.00%	0.00
16	6	2	70	109,281.00	1.30	109,654.00	1.30	0.34%	- 0.00
17	7	3	70	116,027.00	0.08	119,880.00	0.07	3.32%	- 0.00
18	8	3	70	116,568.00	0.08	116,832.00	0.07	0.23%	- 0.00
Promedio								0.26%	0.004

Tabla 5. Comparación Modelo vs Simheurística.

Al comparar los valores esperados de las simulaciones de la distribución LogNormal se encontró una diferencia del 2.07% y 0.91 horas para las funciones objetivo de costo y tardanza respectivamente. Por otro lado, se observó que la simheurística tuvo un mejor desempeño en el 38.89% de los casos para la función objetivo de costos y en el 55.56% en la función objetivo de tardanza. Adicionalmente, se encontró que las simulaciones de las soluciones del modelo matemático fueron factibles en el 87.43% de los casos de la función objetivo de costos y del 94.78% de la función objetivo de tardanza.

Los resultados de los valores esperados de las simulaciones de la distribución uniforme muestran una diferencia del 0.03% y -0.20 horas para las funciones objetivo de costo y tardanza respectivamente (donde la simheurística tuvo un mejor desempeño general en la función objetivo de tardanza). Por otro lado, se observó que la simheurística obtuvo mejores resultados en el 50.00% de los casos para la función objetivo de costos y en el 61.11% en la función objetivo de tardanza. Adicionalmente, se encontró que las simulaciones de las soluciones del modelo matemático fueron factibles en el 80.06% de los casos de la función objetivo de costos y del 84.96% de la función objetivo de tardanza.

Finalmente, se tomaron datos de los tiempos de ejecución de la simheurística de las 6 instancias pequeñas utilizadas en las simulaciones para el caso determinístico, y se compararon contra los tiempos de ejecución del modelo matemático para esas mismas instancias. Teniendo en cuenta que el modelo matemático solo puede resolver el problema con una función objetivo a la vez (mientras que la simheurística calcula las dos funciones objetivo a la vez), se reportan los tiempos de cada función objetivo por separado, pero se compara el resultado de la suma de los tiempos de ejecución del modelo matemático de ambas funciones objetivo. Los resultados se muestran en la Tabla 8.

Distribución LogNormal - Instancias Pequeñas		Simheurística		Modelo Matemático				RPD - FOC	DA - FOT
Instancia	Coefficiente	Valor esperado FOC(\$)	Valor esperado FOT(h)	Valor esperado FOC(\$)	Valor esperado FOT(h)	%Factibles FOC	%Factibles FOT		
1	0.1	68136.00	0.00	68242.09	0.00	100.0%	100.0%	-0.16%	0.00
	0.15	69524.50	0.00	68258.46	0.04	50.0%	100.0%	1.85%	-0.04
	0.2	73230.60	0.00	70419.71	0.45	83.3%	100.0%	3.99%	-0.45
2	0.1	110609.00	0.91	110029.86	0.96	100.0%	100.0%	0.53%	-0.05
	0.15	115394.00	3.54	110181.38	0.92	100.0%	100.0%	4.73%	2.62
	0.2	119663.00	4.95	110383.10	0.97	100.0%	100.0%	8.41%	3.97
3	0.1	105323.00	0.00	109174.48	0.13	83.3%	100.0%	-3.53%	-0.13
	0.15	105727.00	0.00	115275.03	0.30	100.0%	100.0%	-8.28%	-0.30
	0.2	105181.00	0.00	113254.34	0.04	57.1%	100.0%	-7.13%	-0.04
4	0.1	116748.00	3.69	110885.68	2.52	100.0%	100.0%	5.29%	1.17
	0.15	118207.00	9.90	117631.94	3.26	100.0%	100.0%	0.49%	6.65
	0.2	121993.00	8.81	120240.94	4.02	100.0%	100.0%	1.46%	4.79
5	0.1	108917.00	0.71	109700.98	0.69	100.0%	82.2%	-0.71%	0.02
	0.15	108829.00	0.75	109490.23	1.72	100.0%	68.9%	-0.60%	-0.97
	0.2	104144.00	0.00	109713.50	2.16	0.0%	55.0%	-5.08%	-2.16
6	0.1	131482.00	0.00	116594.59	0.84	100.0%	100.0%	12.77%	-0.84
	0.15	128272.00	0.00	116130.11	0.64	100.0%	100.0%	10.46%	-0.64
	0.2	132382.00	3.02	117331.63	0.28	100.0%	100.0%	12.83%	2.74
Promedio						87.43%	94.78%	2.07%	0.91
							Casos de mejora	38.89%	55.56%

Tabla 6. Comparación Instancias pequeñas. Distribución Lognormal.

Distribución Uniforme - Instancias Pequeñas		Simheurística		Modelo Matemático				RPD - FOC	DA - FOT
Instancia	Coefficiente	Valor esperado FOC(\$)	Valor esperado FOT(h)	Valor esperado FOC(\$)	Valor esperado FOT(h)	%Factibles FOC	%Factibles FOT		
1	0.1	73,182.68	-	68,267.50	-	100.0%	87.5%	7.20%	-
	0.15	67,266.60	-	69,879.88	-	100.0%	53.8%	-3.74%	-
	0.2	69,685.40	-	73,187.87	-	50.0%	59.8%	-4.79%	-
2	0.1	115,131.00	0.82	110,038.30	0.88	100.0%	100.0%	4.63%	- 0.06
	0.15	115,017.00	1.20	112,525.15	0.94	100.0%	100.0%	2.21%	0.25
	0.2	113,576.00	0.36	109,983.11	0.93	100.0%	100.0%	3.27%	- 0.58
3	0.1	110,243.00	0.26	104,918.18	0.13	28.6%	54.5%	5.08%	0.13
	0.15	110,422.00	0.74	114,759.96	-	12.5%	40.2%	-3.78%	0.74
	0.2	105,851.00	-	113,181.19	0.16	100.0%	100.0%	-6.48%	- 0.16
4	0.1	111,266.00	1.72	110,821.61	2.05	100.0%	100.0%	0.40%	- 0.33
	0.15	110,515.00	1.51	114,695.72	2.30	100.0%	92.7%	-3.65%	- 0.79
	0.2	110,402.00	1.94	113,271.32	2.57	100.0%	100.0%	-2.53%	- 0.63
5	0.1	108,975.00	1.15	109,244.60	1.33	100.0%	100.0%	-0.25%	- 0.19
	0.15	108,926.00	0.77	109,415.22	1.34	100.0%	78.6%	-0.45%	- 0.57
	0.2	108,732.00	0.54	109,374.86	1.32	50.0%	62.0%	-0.59%	- 0.78
6	0.1	120,920.00	0.46	116,517.12	0.38	100.0%	100.0%	3.78%	0.08
	0.15	116,593.00	-	116,582.12	0.21	50.0%	100.0%	0.01%	- 0.21
	0.2	116,704.00	-	116,548.23	0.47	50.0%	100.0%	0.13%	- 0.47
Promedio						80.06%	84.96%	0.03%	- 0.20
							Casos de mejora	50.00%	61.11%

Tabla 7. Comparación Instancias pequeñas. Distribución Uniforme.

Tamaño de instancia (Número de clientes)	Instancia	Tiempo FOC - Modelo (s)	Tiempo FOT - Modelo (s)	Tiempo total - Modelo (s)	Tiempo - Heurística (s)
6	1	0.5	0.2	0.7	7.55
	2	0.9	0.5	1.4	6.05
	3	3	1.6	4.6	7.39
8	4	40	31	71	8.26
	5	54	43.6	97.6	7.43
	6	1103.1	173	1276.1	7.11

Tabla 8. Resultados tiempos de ejecución.

Los resultados muestran que, para las instancias de tamaño de 6 clientes, el modelo matemático tuvo un tiempo de ejecución menor, sin embargo, es una diferencia de unos pocos segundos en comparación con la simheurística. Para el caso de las instancias de tamaño de 8 clientes, se puede ver como agregando únicamente 2 nodos, el tiempo de ejecución del modelo matemático aumenta drásticamente, mientras que los tiempos de ejecución de la simheurística tienen un aumento muy pequeño, y en los tres casos resultan mucho menores a los tiempos del modelo matemático.

7. Evaluación de la simheurística con respecto al modelo matemático en instancias grandes

En esta sección se evalúa el desempeño de la simheurística en instancias grandes. Para ello se ejecutaron 6 instancias grandes (tamaños de 20 y 30 clientes) con un tiempo máximo de 2 horas de ejecución en el modelo matemático y la mejor solución encontrada hasta ese momento fue simulada con las mismas distribuciones de probabilidad y coeficientes de variación que las simulaciones de las instancias pequeñas. En este caso también se calculó el porcentaje de réplicas de cada simulación de las soluciones del modelo matemático que resultaron factibles. Luego se ejecutó la simheurística para dichas instancias y se compararon los resultados de los valores esperados de las funciones objetivo contra los resultados simulados del modelo matemático. Adicionalmente, se calculó el porcentaje de casos donde la simheurística tuvo un mejor desempeño que el modelo matemático para cada función objetivo. En la Tabla 9 se muestran los resultados de la comparación para el caso determinístico. En la Tabla 10 y Tabla 11 se encuentran los resultados para las simulaciones de las distribuciones LogNormal y uniforme respectivamente.

Es importante resaltar que el modelo matemático no halló ninguna solución óptima dentro del tiempo de ejecución aceptable. Además, se intentó realizar esta comparación con una instancia más grande (tamaño de 50 clientes) pero el modelo matemático no pudo obtener una solución factible después de transcurridas las 2 horas límite.

En el caso determinístico, la simheurística logró encontrar soluciones con valores de función objetivo de costos muy cercanas a las arrojadas por el modelo matemático con una diferencia de 1.45% y es importante resaltar que la simheurística encontró mejores soluciones en el 66.67% de las instancias utilizadas. Respecto a la función objetivo de tardanza, la simheurística halló mejores soluciones en todas las instancias tenidas en cuenta, con un porcentaje de mejora de 77.38% en comparación al modelo matemático.

En el caso de las simulaciones de la distribución LogNormal se encontró una diferencia de 23.07% y -0.08% para la función objetivo de costos y función objetivo de tardanza respectivamente (resaltando que la simheurística obtuvo un mejor desempeño general en la función objetivo de tardanza que las soluciones del modelo matemático). En el caso de la función objetivo de costos, el modelo matemático obtuvo mejores resultados en todas las instancias y coeficientes de variación tenidos en cuenta, pero únicamente el 9.06% de las simulaciones resultaron ser factibles. Por otro lado, en el caso de la función objetivo de tardanza, la simheurística tuvo mejor desempeño en el 55.56% de los casos.

Caso Determinístico - Instancias Grandes							
Instancia	Número de clientes	Modelo Matemático		Simheurística		Dif. FOC	Dif. FOT
		FOC (\$)	FOT (min.)	FOC (\$)	FOT (min.)		
1	20	173,217.00	135.51	227,556.00	5.75	31.37%	-95.75%
2	30	405,045.00	115.54	381,564.00	32.60	-5.80%	-71.79%
3	20	232,418.25	37.54	233,142.00	3.19	0.31%	-91.49%
4	30	368,311.00	94.02	341,322.00	12.45	-7.33%	-86.76%
5	20	253,849.50	12.73	243,417.00	9.21	-4.11%	-27.61%
6	30	413,959.25	149.65	390,066.00	13.68	-5.77%	-90.86%
Promedio						1.45%	-77.38%
Casos de mejora						66.67%	100.00%

Tabla 9. Comparación Caso determinístico - Instancias Grandes

Distribución LogNormal - Instancias Grandes										
Instancia	Número de clientes	Coeficiente	Simheurística		Modelo Matemático				Dif. FOC	Dif. FOT
			FOC (\$)	FOT (min.)	FOC (\$)	FOT (min.)	%Factibles FOC	%Factibles FOT		
1	20	0.1	231,210.00	5.80	173,574.56	137.23	0.00%	0.00%	33.21%	-95.77%
		0.15	260,154.00	27.34	176,183.29	139.90	0.00%	0.05%	47.66%	-80.46%
		2	267,520.00	46.94	174,301.39	140.65	0.00%	0.00%	53.48%	-66.62%
2	30	0.1	439,389.00	78.82	408,353.80	114.33	0.00%	0.00%	7.60%	-31.06%
		0.15	453,766.00	144.52	410,522.82	119.44	33.00%	0.00%	10.53%	20.99%
		2	513,556.00	149.28	414,244.85	117.29	0.00%	0.00%	23.97%	27.27%
3	20	0.1	281,696.00	10.54	234,231.95	38.44	0.00%	0.00%	20.26%	-72.58%
		0.15	313,141.00	38.52	234,055.33	38.62	0.00%	9.00%	33.79%	-0.28%
		2	316,661.00	44.15	234,178.36	42.39	0.00%	8.00%	35.22%	4.14%
4	30	0.1	403,914.00	88.33	368,352.91	93.46	0.00%	0.00%	9.65%	-5.49%
		0.15	418,278.00	99.19	369,812.46	103.34	0.00%	47.00%	13.11%	-4.01%
		2	420,137.00	89.90	366,530.60	102.16	0.00%	52.00%	14.63%	-12.00%
5	20	0.1	298,102.00	16.22	254,910.90	11.89	0.00%	0.00%	16.94%	36.49%
		0.15	290,920.00	18.97	253,297.01	15.46	50.00%	4.00%	14.85%	22.69%
		2	330,222.00	61.97	267,625.36	17.35	14.00%	3.00%	23.39%	257.15%
6	30	0.1	459,976.00	114.96	413,650.00	150.43	33.00%	33.00%	11.20%	-23.58%
		0.15	486,395.00	167.85	409,812.45	152.04	33.00%	57.00%	18.69%	10.40%
		2	523,795.00	169.63	412,367.66	152.52	0.00%	33.00%	27.02%	11.21%
Promedio						9.06%	13.67%	23.07%	-0.08%	
Casos de mejora						0%	55.56%			

Tabla 10. Comparación Instancia Grande, Distribución LogNormal

Distribución Uniforme - Instancias Grandes										
Instancia	Número de clientes	Coeficiente	Simheurística		Modelo Matemático				Dif. FOC	Dif. FOT
			FOC (\$)	FOT (min.)	FOC (\$)	FOT (min.)	%Factibles FOC	%Factibles FOT		
1	20	0.1	227,169.00	7.42	173,260.47	134.50	0.00%	0.00%	31.11%	-94.48%
		0.15	240,666.00	4.68	174,905.48	137.41	6.56%	0.00%	37.60%	-96.59%
		2	233,906.00	5.50	172,933.65	133.32	0.00%	0.00%	35.26%	-95.88%
2	30	0.1	412,283.00	45.21	405,173.64	115.33	0.00%	0.00%	1.75%	-60.80%
		0.15	389,236.00	44.99	405,178.53	117.61	0.00%	14.00%	-3.93%	-61.74%
		2	400,012.00	46.72	411,510.99	119.09	0.00%	33.00%	-2.79%	-60.77%
3	20	0.1	226,323.00	2.65	234,136.23	38.46	100.00%	70.00%	-3.34%	-93.11%
		0.15	268,805.00	4.13	232,104.77	37.24	0.00%	0.00%	15.81%	-88.91%
		2	244,196.00	3.15	234,896.04	38.57	0.00%	0.00%	3.96%	-91.83%
4	30	0.1	329,614.00	2.90	365,138.57	94.41	0.00%	0.00%	-9.73%	-96.93%
		0.15	346,290.00	16.79	367,929.78	94.42	0.00%	0.00%	-5.88%	-82.21%
		2	346,383.00	9.40	370,204.75	91.52	0.00%	0.00%	-6.43%	-89.73%
5	20	0.1	255,152.00	7.17	253,890.97	12.73	0.00%	0.00%	0.50%	-43.65%
		0.15	279,552.00	4.49	254,006.20	12.64	4.50%	0.00%	10.06%	-64.51%
		2	285,528.00	14.05	252,197.05	14.16	0.00%	0.00%	13.22%	-0.74%
6	30	0.1	386,277.00	14.74	408,733.48	149.55	0.00%	50.00%	-5.49%	-90.14%
		0.15	402,012.00	20.45	412,585.08	149.98	0.00%	50.00%	-2.56%	-86.36%
		2	408,881.00	25.49	408,962.36	148.27	0.00%	40.00%	-0.02%	-82.81%
Promedio							6.17%	14.28%	6.06%	-76.73%
							Casos de mejora	50.00%	100.00%	

Tabla 11. Comparación Instancia Grande, Distribución Uniforme.

Con respecto a las simulaciones de la distribución uniforme, se observó un comportamiento muy parecido al caso determinístico, donde el modelo matemático tuvo un mejor desempeño en la función objetivo de costos con una diferencia de 6.06%, y la simheurística tuvo un mejor desempeño en la función objetivo de tardanza, obteniendo una diferencia del -76.73% en comparación con el modelo matemático. Sin embargo, la simheurística obtuvo mejores resultados en el 50.00% de los casos en la función objetivo de costos y en el 100% en la función objetivo de tardanza. También se debe recalcar que en este caso el porcentaje de soluciones factibles de las simulaciones del modelo matemático es relativamente bajo con un 6.17% en el caso de la función objetivo de costos y 14.28% para el caso de la función objetivo de tardanza.

Posteriormente, se tomaron los datos de los tiempos de ejecución de la simheurística para el caso determinístico de las 6 instancias grandes que fueron comparadas con el modelo matemático. Los resultados se muestran en la Tabla 12.

Número de clientes	Instancia	Tiempo - Heurística (s)
20	1	11.58
	2	11.92
	3	10.77
30	4	12.86
	5	12.99
	6	12.21
Promedio		12.06

Tabla 12. Tiempos de ejecución.

Los resultados muestran que la simheurística logró encontrar soluciones cercanas, e incluso mejores en algunos casos, a las soluciones arrojadas por el modelo matemático en un tiempo significativamente menor, pues la simheurística requirió 12.06 segundos en promedio para encontrar las soluciones, mientras que el modelo matemático se ejecutó durante 2 horas, y no había logrado encontrar una solución óptima.

Finalmente, se corrió la simheurística para para instancias más grandes (50, 70 y 90), con ambas distribuciones de probabilidad y 3 coeficientes de variación; y se tomaron los tiempos de ejecución para cada prueba. Los resultados se muestran en la Tabla 13.

Número de clientes	Instancia	Distribución	Coef. Variación	Tiempo (s)	Tiempo promedio (s)
50	1	2	0.00	22.91	75.83
		1	0.00	31.45	
		2	0.10	43.54	
		1	0.10	141.06	
		2	0.15	49.50	
		1	0.15	144.01	
		2	0.20	53.22	
	3	1	0.20	120.98	99.30
		2	0.00	26.08	
		1	0.00	43.90	
		2	0.10	74.01	
		1	0.10	214.58	
		2	0.15	73.53	
		1	0.15	158.61	
70	2	2	0.00	25.19	104.13
		1	0.00	44.66	
		2	0.10	67.90	
		1	0.10	249.41	
		2	0.15	75.40	
		1	0.15	167.14	
		2	0.20	100.91	
	4	1	0.20	102.47	63.62
		2	0.00	18.29	
		1	0.00	27.74	
		2	0.10	43.28	
		1	0.10	117.93	
		2	0.15	45.52	
		1	0.15	119.76	
90	5	2	0.00	31.52	141.06
		1	0.00	61.11	
		2	0.10	103.80	
		1	0.10	286.24	
		2	0.15	117.01	
		1	0.15	181.71	
		2	0.20	127.46	
	6	1	0.20	219.62	146.67
		2	0.00	31.26	
		1	0.00	62.49	
		2	0.10	98.51	
		1	0.10	342.35	
		2	0.15	138.65	
		1	0.15	238.57	
		2	0.20	106.03	
		1	0.20	155.49	

Tabla 13. *Tiempos de ejecución Instancias Grandes*

Los resultados muestran que incluso en instancias grandes, los tiempos de ejecución de la simheurística son bastante aceptables, pues tarda entre 1 y 3 minutos aproximadamente para hallar soluciones y simular sus resultados, mientras que el modelo matemático puede durar más de 2 horas corriendo una instancia del mismo tamaño, sin encontrar una solución factible.

8. Análisis Experimental de la simheurística para diferentes distribuciones de probabilidad y coeficientes de variación

Se utilizó la herramienta de diseño de experimentos (DOE por sus siglas en inglés) para analizar los resultados de la simheurística cuando se tienen en cuenta las variables estocásticas de demanda de los clientes y tiempos de entrega. Se realizaron pruebas preliminares para parametrizar la simheurística, y se escogieron los parámetros que fueron ingresados y arrojaron los mejores resultados de las diferentes instancias probadas en el indicador de desempeño MID (*Mean ideal distance*) propuesto por (Karimi, Zandieh, & Karamooz, 2010). Este indicador mide la capacidad de hallar soluciones con valores de funciones objetivo cercanas a cero en una heurística que resuelve problemas bi-objetivo y se calcula con la siguiente fórmula (21):

$$MID = \frac{\sum_{i=1}^n \sqrt{f_{1i}^2 + f_{2i}^2}}{n} \quad (21)$$

Donde f_{1i} es el valor de la función objetivo 1 de la i -ésima solución no-dominada hallada por la simheurística, y n es el número de soluciones no-dominadas halladas por la simheurística. En la Tabla 14 se muestran los parámetros escogidos para la simheurística.

Parámetro	Valor
Número de generaciones	150
Número de soluciones por generación	100
Probabilidad de mutación	0.5

Tabla 14. Parámetros de la simheurística

Información del factor	
Factor	Niveles Valores
Distribución	2 LogNormal; Uniforme
Coficiente	4 0,00; 0,10; 0,15; 0,20

Figura 7. Resumen del diseño del experimento.

Análisis de Varianza					
Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Modelo	18	5,23793E+12	2,90996E+11	820,79	0,000
Bloques	11	5,22314E+12	4,74831E+11	1339,31	0,000
Lineal	4	11125949107	2781487277	7,85	0,000
Distribución	1	5868509808	5868509808	16,55	0,000
Coficiente	3	5257439299	1752479766	4,94	0,003
Interacciones de 2 términos	3	3662623852	1220874617	3,44	0,021
Distribución*Coficiente	3	3662623852	1220874617	3,44	0,021
Error	77	27299107306	354533861		
Total	95	5,26523E+12			

Figura 8. ANOVA. Función Objetivo de Costos (FOC).

Análisis de Varianza					
<u>Fuente</u>	<u>GL</u>	<u>SC Ajust.</u>	<u>MC Ajust.</u>	<u>Valor F</u>	<u>Valor p</u>
Modelo	18	17592289	977349	42,99	0,000
Bloques	11	16667710	1515246	66,66	0,000
Lineal	4	642388	160597	7,06	0,000
Distribución	1	307173	307173	13,51	0,000
Coefficiente	3	335215	111738	4,92	0,004
Interacciones de 2 términos	3	282191	94064	4,14	0,009
Distribución*Coefficiente	3	282191	94064	4,14	0,009
Error	77	1750375	22732		
Total	95	19342664			

Figura 9. ANOVA. Función objetivo de tardanza (FOT).

Se realizó la comprobación de supuestos, pero no se cumplió el supuesto de normalidad de los residuos. Por esta razón, se aplicó la prueba de Friedman para corroborar los resultados del ANOVA y poder concluir si los resultados del ANOVA son correctos. En la Figura 10 y en la Figura 11 se muestran los resultados de la prueba de Friedman para cada variable de respuesta.

Teniendo en cuenta que con una significancia del 5% las dos pruebas de Friedman dieron significativas, se realizó la prueba de Nemenyi, para lograr identificar los tratamientos en los que existen diferencias significativas al compararse con los demás tratamientos.

Prueba			
Hipótesis nula	H ₀ : Todos los efectos del tratamiento son cero		
Hipótesis alterna	H ₁ : No todos los efectos del tratamiento son cero		
<u>Método</u>	<u>GL</u>	<u>Chi-cuadrada</u>	<u>Valor p</u>
No ajustado para empates	7	20,59	0,004
Ajustado para empates	7	20,65	0,004

Estadísticas descriptivas			
<u>Tratamiento</u>	<u>N</u>	<u>Mediana</u>	<u>Suma de clasificaciones</u>
LogNormal0,00	12	232190	49,5
LogNormal0,10	12	234043	49,0
LogNormal0,15	12	245889	73,0
LogNormal0,20	12	252411	81,0
Uniforme0,00	12	232012	44,5
Uniforme0,10	12	232594	47,0
Uniforme0,15	12	232386	42,0
Uniforme0,20	12	234226	46,0
General	96	236969	

Figura 10. Resultados prueba Friedman, función objetivo de costos (FOC).

Prueba			
Hipótesis nula	H ₀ : Todos los efectos del tratamiento son cero		
Hipótesis alterna	H ₁ : No todos los efectos del tratamiento son cero		
Método	GL	Chi-cuadrada	Valor p
No ajustado para empates	7	30,22	0,000
Ajustado para empates	7	33,51	0,000

Estadísticas descriptivas			
Tratamiento	N	Mediana	Suma de clasificaciones
LogNormal0,00	12	120,229	49,5
LogNormal0,10	12	118,182	48,0
LogNormal0,15	12	141,293	78,0
LogNormal0,20	12	195,087	86,0
Uniforme0,00	12	119,774	42,5
Uniforme0,10	12	112,256	43,5
Uniforme0,15	12	117,299	43,0
Uniforme0,20	12	121,493	41,5
General	96	130,702	

Figura 11. Resultados prueba de Friedman, función objetivo de tardanza (FOT).

Los resultados cuya diferencia entre tratamientos es significativa se presentan en la Tabla 15:

Comparación de Medias N° Prueba		Comparación de Medias Prueba		Ri	Rj	Valor Crítico	Ri - Rj	Binario Ri - Rj > V.C	Análisis
i	j	i	j						
1	3	LogNormal0,00	LogNormal0,15	49.5	73	10.508	23.5	1	Con 95% de confianza, se puede concluir que EXISTE una diferencia entre los rankings de las pruebas analizadas.
1	4	LogNormal0,00	LogNormal0,20	49.5	81	10.508	31.5	1	
2	3	LogNormal0,10	LogNormal0,15	49	73	10.508	24	1	
2	4	LogNormal0,10	LogNormal0,20	49	81	10.508	32	1	
3	5	LogNormal0,15	Uniforme0,00	73	44.5	10.508	28.5	1	
3	6	LogNormal0,15	Uniforme0,10	73	47	10.508	26	1	
3	7	LogNormal0,15	Uniforme0,15	73	42	10.508	31	1	
3	8	LogNormal0,15	Uniforme0,20	73	46	10.508	27	1	
4	5	LogNormal0,20	Uniforme0,00	81	44.5	10.508	36.5	1	
4	6	LogNormal0,20	Uniforme0,10	81	47	10.508	34	1	
4	7	LogNormal0,20	Uniforme0,15	81	42	10.508	39	1	
4	8	LogNormal0,20	Uniforme0,20	81	46	10.508	35	1	

Tabla 15. Resultados prueba Nemenyi, función objetivo de tardanza (FOT).

Según lo anterior, se observa que entre más grande sea la diferencia entre los coeficientes de variación mayor será la diferencia entre el comportamiento de los comparados a pesar de que pueda ser la misma distribución analizada, pues el coeficiente de variación genera soluciones con diferencias más grandes. Por ejemplo, se recalca como entre la distribución LogNormal con coeficiente de variación 0 y LogNormal con coeficiente de variación 0,1 no hubo diferencia significativa entre su comportamiento del valor de la función objetivo de costos y de la función objetivo de tardanza. Tampoco existe diferencia significativa entre los resultados con coeficiente de variación 0 para la distribución LogNormal y la distribución Uniforme. Lo anterior implica que a mayor coeficiente de variación mayor es el valor esperado de las funciones objetivo y por ende mayor el indicador del MID en la frontera de Pareto. Además, se observa que bajo un mismo coeficiente de variación entre las dos distribuciones LogNormal y uniforme hay diferencia significativa, lo cual implica que el tipo de distribución tiene también una influencia en los resultados. De lo anterior se deduce la importancia a nivel práctico de reducir al máximo la variabilidad de los parámetros estocásticos en lo que esté bajo el control de las empresas. Sin embargo, habrá variables que no se pueden controlar, por lo cual es muy importante recabar datos muy ajustados para estimar adecuadamente las distribuciones de probabilidad de los mismos y así poder dar soluciones ajustadas a su comportamiento.

9. Limitaciones, conclusiones y recomendaciones

En este proyecto se solucionó un VRP con ventanas de tiempo y demandas estocásticas cuyo objetivo constaba en minimizar la tardanza y los costos asociados a las rutas obtenidas. Para solucionarlo se desarrolló una simheurística que hibridiza la metaheurística NSGA-II y la simulación de Montecarlo. Los costos por minimizar corresponden a costos variables por tiempo de recorrido del camión, costo fijo del uso del camión y penalización por el incumplimiento de llegada dentro de las ventanas de tiempo de cada cliente. Por su parte, la minimización de tardanza corresponde a encontrar rutas en las que en las ventanas de tiempo superiores (Soft Windows), el camión no llegue después de mucho tiempo a atender cada cliente.

La simheurística fue evaluada con instancias pequeñas utilizando un coeficiente de variación cero para emular el caso determinístico y así poder comparar con los resultados dados por el modelo matemático en instancias pequeñas, concluyendo con un gap de 0.26% para la función objetivo de costos, y de 0.004 horas (DA) para la que la función objetivo de tardanza, que la simheurística presenta resultados muy competitivos y en menores tiempos computacionales a medida que crece la cantidad de nodos. En cuanto a la simheurística propuesta, los resultados muestran que se obtiene un excelente desempeño al ser ejecutada con instancias grandes al compararse con el modelo de programación lineal utilizado, logrando encontrar soluciones con valores de funciones objetivo cercanas a las mejores encontradas por el modelo, y cumpliendo con el 100% de factibilidad en todas las réplicas, mientras que las simulaciones de las soluciones arrojadas por el modelo de programación lineal obtuvieron bajos porcentajes de factibilidad en las instancias grandes. De la misma manera se puede concluir que la simheurística permitió encontrar resultados satisfactorios para las funciones objetivo de costo y tardanza, en tiempos de ejecución aceptables para instancias considerablemente grandes brindando así un acercamiento a la solución óptima en tiempos mucho menores, pues se encontró una diferencia promedio de 119.8 minutos entre los tiempos de ejecución de la simheurística y el modelo matemático para las instancias compradas. Adicionalmente, los resultados de las pruebas estadísticas permiten concluir que: i) los dos factores tratados (distribución y coeficiente de variación de los datos) afectan significativamente los resultados de las soluciones del problema planteado; ii) cuando los datos siguen una distribución uniforme, el coeficiente de variación no afecta significativamente los resultados; y iii) la distribución de probabilidad LogNormal afecta en mayor medida los resultados de ambas funciones objetivo haciéndolas aumentar a medida que el coeficiente de variación es mayor, a pesar de que los coeficientes que se probaron son relativamente pequeños.

Los resultados de este estudio muestran que llevar un control de los tiempos de desplazamiento para la entrega de bienes y servicios a los clientes puede afectar significativamente los resultados del desempeño de una compañía, pues dependiendo de la distribución de los datos y la variabilidad de estos, los resultados pueden tener diferencias significativas. Por esta misma razón, es recomendable contar con herramientas que ayuden a medir estos parámetros para poder determinar con certeza una distribución de probabilidad que se ajuste adecuadamente a los datos y así poder implementar metodologías y herramientas como las propuestas en este estudio para obtener soluciones que representen la realidad y poder obtener mejoras en los resultados en la práctica.

Dentro de las limitaciones encontradas en el desarrollo del experimento, se pudo resaltar que para los niveles del factor coeficiente de variación, se definieron valores relativamente pequeños (siendo el máximo 0.20) debido a que, al aumentar el coeficiente de variación, aumenta la probabilidad de que se incumplan las restricciones de las ventanas de tiempo y de capacidad de los camiones, haciendo más difícil encontrar soluciones factibles para el problema.

Para futuros estudios se recomienda analizar otras distribuciones de probabilidad y coeficientes de variación para este mismo problema. También se sugiere colocar más parámetros estocásticos al problema implementando la simheurística NSGA-II propuesta. Finalmente se sugiere hacer comparaciones con simheurísticas que involucren otras metaheurísticas diferentes como *Iterated Greedy Algorithm*, PSO, Búsqueda Tabú, GRASP, entre otras.

10. Referencias

- Balaprakash, P., Birattari, M., Stützle, T., & Dorigo, M. (2015). Estimation-based metaheuristics for the single vehicle routing problem with stochastic demands and customers. *Computational Optimization and Applications*, 61(2), 463–487. <https://doi.org/10.1007/s10589-014-9719-z>
- Baradaran, V., Shafaei, A., & Hosseinian, A. H. (2019). Stochastic vehicle routing problem with heterogeneous vehicles and multiple prioritized time windows: Mathematical modeling and solution approach. *Computers and Industrial Engineering*, 131(March), 187–199. <https://doi.org/10.1016/j.cie.2019.03.047>
- Biesinger, B., Hu, B., & Raidl, G. R. (2018). A Genetic Algorithm in Combination with a Solution Archive for Solving the Generalized Vehicle Routing Problem with Stochastic Demands. *Transportation Science*, 52(3), 673–690. <https://doi.org/10.1287/trsc.2017.0778>
- Braekers, K., Hartl, R. F., Parragh, S. N., & Tricoire, F. (2016). A bi-objective home care scheduling problem: Analyzing the trade-off between costs and client inconvenience. *European Journal of Operational Research*, 248(2), 428–443. <https://doi.org/10.1016/j.ejor.2015.07.028>
- Clarke, G., & Wright, J. W. (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research*, 12(4), 568–581. <https://doi.org/10.1287/opre.12.4.568>
- Dantzig, G. B., Fulkerson, D. R., Johnson, S. M., Chvátal, V., & Cook, W. (2010). Solution of a large-scale traveling-salesman problem. In *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art* (pp. 7–28). https://doi.org/10.1007/978-3-540-68279-0_1
- Deb, K., Agrawal, S., Pratap, A., & Meyarivan, T. (2000a). A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II. In *CEUR Workshop Proceedings* (Vol. 1133, pp. 849–858). https://doi.org/10.1007/3-540-45356-3_83
- Eksioglu, B., Vural, A. V., & Reisman, A. (2009, November). The vehicle routing problem: A taxonomic review. *Computers and Industrial Engineering*, Vol. 57, pp. 1472–1483. <https://doi.org/10.1016/j.cie.2009.05.009>
- Erera, A. L., Morales, J. C., & Savelsbergh, M. (2010). The Vehicle Routing Problem with Stochastic Demand and Duration Constraints. *Transportation Science*, 44(4), 474–492. <https://doi.org/10.1287/trsc.1100.0324>
- Errico, F., Desaulniers, G., Gendreau, M., Rei, W., & Rousseau, L. M. (2018). The vehicle routing problem with hard time windows and stochastic service times. *EURO Journal on Transportation and Logistics*, 7(3), 223–251. <https://doi.org/10.1007/s13676-016-0101-4>
- Framinan, J. M., & Perez-Gonzalez, P. (2015). On heuristic solutions for the stochastic flowshop scheduling problem. *European Journal of Operational Research*, 246(2), 413–420. <https://doi.org/10.1016/j.ejor.2015.05.006>
- Fu, J., & Wen, X. H. (2017). Model-based multi-objective optimization methods for efficient management of subsurface flow. *Society of Petroleum Engineers - SPE Reservoir Simulation Conference 2017*, (April), 159–178. <https://doi.org/10.2118/182598-pa>
- Goodson, J. C., Thomas, B. W., & Ohlmann, J. W. (2016). Restocking-based rollout policies for the vehicle routing problem with stochastic demand and duration limits. *Transportation Science*, 50(2), 591–607. <https://doi.org/10.1287/trsc.2015.0591>
- Gutierrez, A., Dieulle, L., Labadie, N., & Velasco, N. (2018). A hybrid metaheuristic algorithm for the

- vehicle routing problem with stochastic demands. *Computers and Operations Research*, 99, 135–147. <https://doi.org/10.1016/j.cor.2018.06.012>
- Juan, A. A., Faulin, J., Grasman, S. E., Rabe, M., & Figueira, G. (2015). A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems. *Operations Research Perspectives*, 2, 62–72. <https://doi.org/10.1016/j.orp.2015.03.001>
- Juan, A. A., Faulin, J., Jorba, J., Caceres, J., & Marquès, J. M. (2013). Using parallel & distributed computing for real-time solving of vehicle routing problems with stochastic demands. *Annals of Operations Research*, 207(1), 43–65. <https://doi.org/10.1007/s10479-011-0918-z>
- Juan, A., Faulin, J., Grasman, S., Riera, D., Marull, J., & Mendez, C. (2011). Using safety stocks and simulation to solve the vehicle routing problem with stochastic demands. *Transportation Research Part C: Emerging Technologies*, 19(5), 751–765. <https://doi.org/10.1016/j.trc.2010.09.007>
- Karimi, N., Zandieh, M., & Karamooz, H. R. (2010). Bi-objective group scheduling in hybrid flexible flowshop: A multi-phase approach. *Expert Systems with Applications*, 37(6), 4024–4032. <https://doi.org/10.1016/j.eswa.2009.09.005>
- Lagos, C., Guerrero, G., Cabrera, E., Moltedo-Perfetti, Andr. S., Johnson, F., & Paredes, F. (2018). An improved particle swarm optimization algorithm for the VRP with simultaneous pickup and delivery and time windows. *IEEE Latin America Transactions*, 16(6), 1732–1740. <https://doi.org/10.1109/TLA.2018.8444393>
- Lei, H., Laporte, G., & Guo, B. (2012). The Vehicle Routing Problem with Stochastic Demands and Split Deliveries. *INFOR: Information Systems and Operational Research*, 50(2), 59–71. <https://doi.org/10.3138/infor.50.2.059>
- Lenstra, J. K., & Kan, A. H. G. R. (1981). Complexity of vehicle routing and scheduling problems. *Networks*. <https://doi.org/10.1002/net.3230110211>
- Liebman, J. C. (1970). Mathematical models for solid waste collection and disposal. *38th National Meeting of the Operations Research Society of America Bulletin of the Operations Research Society of America*, 18.
- Marks, D. H., & Striker, R. (1971). Routing for public service vehicles. *Journal of Urban Planning and Development*, 97, 165–171.
- Miranda, D. M., Branke, J., & Conceição, S. V. (2018). Algorithms for the multi-objective vehicle routing problem with hard time windows and stochastic travel time and service time. *Applied Soft Computing*, 70(1), 66–79. <https://doi.org/10.1016/j.asoc.2018.05.026>
- Montero, A., José Miranda-Bront, J., & Méndez-Díaz, I. (2017). An ILP-based local search procedure for the VRP with pickups and deliveries. *Annals of Operations Research*, 259(1–2), 327–350. <https://doi.org/10.1007/s10479-017-2520-5>
- Rajabi-Bahaabadi, M., Shariat-Mohaymany, A., Babaei, M., & Vigo, D. (2019). Reliable vehicle routing problem in stochastic networks with correlated travel times. In *Operational Research*. <https://doi.org/10.1007/s12351-019-00452-w>
- Rincon-Garcia, N., Waterson, B. J., & Cherrett, T. J. (2018). Requirements from vehicle routing software: perspectives from literature, developers and the freight industry. *Transport Reviews*, 38(1), 117–138. <https://doi.org/10.1080/01441647.2017.1297869>
- Salavati-Khoshghalb, M., Gendreau, M., Jabali, O., & Rei, W. (2019). An exact algorithm to solve the vehicle

routing problem with stochastic demands under an optimal restocking policy. *European Journal of Operational Research*, 273(1), 175–189. <https://doi.org/10.1016/j.ejor.2018.07.039>

Sarasola, B., Doerner, K. F., Schmid, V., & Alba, E. (2016). Variable neighborhood search for the stochastic and dynamic vehicle routing problem. *Annals of Operations Research*, 236(2), 425–461. <https://doi.org/10.1007/s10479-015-1949-7>

Sorensen, K., & Marc, S. (2009). *A Practical Approach for Robust and Flexible Vehicle Routing Using Metaheuristics and Monte Carlo Sampling* (pp. 387–407). pp. 387–407. Retrieved from <https://login.ezproxy.javeriana.edu.co/login?url=https%3A%2F%2Fsearch.proquest.com%2Fdocview%2F223647881%3Faccountid%3D13250>

Vidal, T., Teodor, G., Gendreau, M., & Prins, C. (2014). A unified solution framework for multi-attribute vehicle routing problems. *European Journal of Operational Research*, 234(3), 658.

Wilson, N., & Sussman, J. (1971). Implementation of computer algorithms for the dial-a-bus system. *39th National Meeting of ORSA/TIMS*.