

Instituto Tecnológico y de Estudios Superiores de Monterrey



MA2004B.201 - OPTIMIZACIÓN ESTOCÁSTICA

SEMESTRE AGOSTO 2024 - DICIEMBRE 2024

GRUPO 201

Reto: Entregable 3

Estudiantes

Pablo Pérez Sandoval	A01710355
Juan Pablo Guzmán Segura	A01039810
Valeria Mariane Cárdenas Rodríguez	A01721814
Rodrigo Leal Torres	A00836930
Máximo Caballero Vargas	A01571607

Profesores

Dr. Fernando Elizalde
Rafael Muñoz

21 de octubre de 2024
Monterrey, Nuevo León

Link al código: <https://drive.google.com/drive/folders/16iJzZIDd0ybX2jc-0DE10IS5Ya5nt0Ra?usp=saring>

1. Introducción al problema

El cambio climático es uno de los desafíos más urgentes de la actualidad, afectando todos los aspectos de la vida cotidiana. Una de las principales causas de este fenómeno es la deforestación, que elimina los bosques que actúan como sumideros naturales de dióxido de carbono (CO_2). Ante esta situación, la reforestación se presenta como una solución viable para mitigar los efectos del cambio climático. Esta práctica consiste en repoblar áreas previamente deforestadas, restaurando la biodiversidad y mejorando la calidad del suelo y el aire. Sin embargo, la implementación de proyectos de reforestación enfrenta múltiples desafíos, como la logística, la incertidumbre sobre las especies existentes y la posible introducción de especies invasoras. En este contexto, el uso de herramientas de optimización resulta fundamental para planificar y ejecutar proyectos de reforestación de manera eficiente, abordando la incertidumbre y minimizando los riesgos asociados.

2. Definición del problema

2.1. Problema

El presente trabajo se centra en optimizar un plan de reforestación para una zona específica definida por nuestro socio formador. El proyecto enfrenta múltiples desafíos logísticos, como la compra excesiva de plantas de ciertas especies y la incertidumbre sobre la presencia de especies no removidas en el terreno. Estos problemas generan pérdidas innecesarias y complicaciones en la planificación. Por consiguiente, se tiene como objetivo aplicar técnicas de optimización estocástica, como las simulaciones de Montecarlo y algoritmos genéticos, para modelar la incertidumbre y proponer una distribución de las especies a ser plantadas de tal forma que se minimice la competencia y, por ende, se maximice el porcentaje de supervivencia en el área de reforestación. Este enfoque permitirá mejorar la toma de decisiones en el proceso de reforestación, garantizando un uso más eficiente de los recursos disponibles.

2.2. Elementos básicos de un modelo matemático

Como primer paso, es esencial comprender la estructura matemática detrás del problema en cuestión. Esto comienza con la identificación y definición de los elementos básicos de un modelo matemático: conjuntos, variables de decisión, parámetros, restricciones y función objetivo. Los conjuntos agrupan los elementos relacionados en el problema, mientras que los parámetros son las características o valores conocidos de antemano que describen el entorno del problema. En contraste, las variables de decisión son aquellos factores controlables dentro del sistema que se está modelando y para los cuales se busca determinar los valores óptimos, de acuerdo con la función objetivo. Las restricciones son las limitaciones que condicionan los valores que las variables de decisión pueden asumir, basándose en las características del problema o el entorno. Finalmente, la función objetivo responde a la pregunta central del modelo, relacionando parámetros, variables de decisión y restricciones, con el propósito de maximizar o minimizar un valor específico. A continuación, se definen estos elementos en el contexto del problema de reforestación.

3. Modelo matemático

3.1. Conjuntos

Para nuestro modelo se consideran dos conjuntos clave. E es el conjunto que contiene todas las especies:

$$E = \{1, 2, 3, 4, 5, 6 \dots n\}, \quad i, k \in E$$

Donde n es el número total de especies.

Asimismo, se tiene un conjunto de todas las posiciones posibles en el terreno donde se pueden plantar las especies:

$$P = \{0, 1, 2, \dots, m\}, \quad j \in P$$

Donde m es el número total de posiciones en el terreno.

3.2. Parámetros

- $C_{i,k}$ es la matriz de competencia que indica cuánta competencia existe entre las especies i y k . La competencia se pondera en una escala del 1 al 100, donde 100 representa la competencia más intensa posible entre las especies.
- a_i es la disponibilidad de cada especie i .
- $d_{j,j'}$ es la distancia entre las posiciones j y j' (posición adyacente a j) en el terreno, considerando un diseño de plantación en tresbolillo. Cabe resaltar que este parámetro solamente tiene tres valores posibles, dado que se busca penalizar a las distancias más cercanas para la ejecución correcta de nuestra función objetivo.

$$d_{j,j'} = \begin{cases} 2, & \text{si la distancia entre las posiciones es de 4 metros.} \\ 1, & \text{si la distancia entre posiciones es de 6.4 metros.} \\ 0, & \text{si las posiciones no son adyacentes entre sí.} \end{cases}$$

3.3. Variables

Se definen dos variables de decisión que relacionan una planta con su posición en el terreno:

$$x_{ij} = \begin{cases} 1, & \text{si la especie } i \text{ está plantada en la posición } j. \\ 0, & \text{si no.} \end{cases}$$
$$y_{kj'} = \begin{cases} 1, & \text{si la especie } k \text{ está plantada en la posición } j'. \\ 0, & \text{si no.} \end{cases}$$

3.4. Restricciones

En el contexto del problema, se establecen dos restricciones importantes:

- **Disponibilidad de especies:** Se debe respetar la disponibilidad específica de cada especie, limitando la cantidad de veces que puede ser plantada:

$$\sum_{j=0}^m x_{ij} = a_i \quad \forall i \in E$$

- **Evitar monocultivos:** Se debe evitar plantar especies iguales en posiciones adyacentes:

$$\sum_{j'=0}^m x_{i,j} + y_{k,j'} \leq 1 \quad \forall i = k, \quad \forall j, j' \in P \quad \text{adyacentes}$$

3.5. Función Objetivo

La función objetivo minimiza la suma total de la competencia entre todas las especies a ser plantadas, considerando los parámetros y restricciones establecidos previamente. El objetivo es obtener una distribución de plantación que minimice la competencia y promueva la supervivencia de las plantas, evitando el monocultivo.

$$\min z = \sum_{i \in S} \sum_{j \in P} \sum_{j' \in P} \sum_{k \in S} x_{ij} Y_{kj'} C_{i,k} d_{jj'}$$

3.6. Modelo matemático completo

$$\min \sum_{i \in S} \sum_{j \in P} \sum_{j' \in P} \sum_{k \in S} x_{ij} Y_{kj'} C_{i,k} d_{jj'}$$

s.a.:

$$\sum_{j=0}^m x_{ij} = a_i \quad \forall i \in E$$

$$\sum_{j'=0}^m x_{i,j} + y_{k,j'} \leq 1 \quad \forall i = k, \quad \forall j, j' \in P \quad \text{adyacentes}$$

4. Grafo de competencia

A continuación se muestra un grafo representativo de la competencia entre las diferentes especies. En él, los nodos corresponden a las especies y las aristas (líneas) que conectan los nodos indican la magnitud de la competencia entre ellas, ponderada por los números que aparecen junto a cada arista. Es importante resaltar que, en este caso, la competencia se encuentra en una escala de 1-10. Asimismo, todas las especies están conectadas entre

sí, lo que implica que no hay especies completamente aisladas del resto en términos de competencia. Esto es indicativo de un ecosistema donde todas las especies afectan, en mayor o menor medida, a las demás.

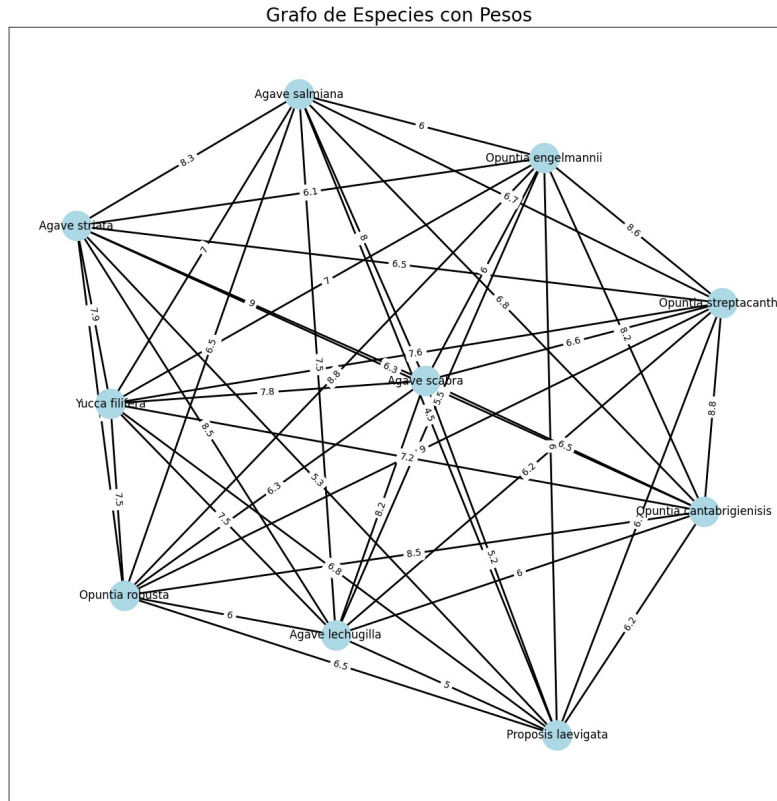


Figura 1: Grafo de competencia entre especies

5. Distribución de probabilidad de plantas por hectárea

Para este análisis, se consideraron varios polígonos con características ambientales y de suelo diferenciadas, y se evaluó la densidad de plantas por hectárea en cada uno de ellos. A partir de estos datos, se generaron las distribuciones de probabilidad que describen el comportamiento de la vegetación en cada área específica.

	Polígono 1	Polígono 2	Polígono 3	Polígono 4	Polígono 5	Polígono 6	Polígono 7	Polígono 8	Polígono 9	Polígono 10
Densidad por hectárea	128,90625	130,1204819	127,0710059	132,6086957	125,375	131,3299233	128,9330922	135,4609929	19,56011666	16,60014912
Probabilidad de plantas por hect	0,04008761622	0,04046522137	0,03951688703	0,04123901284	0,03898945849	0,04084133673	0,04009596369	0,042126028	0,006082858276	0,005162359521
	Polígono 11	Polígono 12	Polígono 13	Polígono 14	Polígono 15	Polígono 16	Polígono 17	Polígono 18	Polígono 19	Polígono 20
Densidad por hectárea	5218,226218	14426,22732	14926,04935	5601,21787	17039,8622	16384,13374	12606,33355	12702,38421	85348,54717	92051,34729
Probabilidad de plantas por hect	0,04008761622	0,04046522137	0,03951688703	0,04123901284	0,03898945849	0,04084133673	0,04009596369	0,042126028	0,006082858276	0,005162359521
	Polígono 21	Polígono 22	Polígono 23	Polígono 24	Polígono 25	Polígono 26	Polígono 27	Polígono 28	Polígono 29	Polígono 30
Densidad por hectárea	134017,7606	363275,6171	384922,0173	139897,8154	444229,719	408931,8775	320851,1669	307660,7965	14067024,49	17869843,09
Probabilidad de plantas por hect	0,04008761622	0,04046522137	0,03951688703	0,04123901284	0,03898945849	0,04084133673	0,04009596369	0,042126028	0,006082858276	0,005162359521

Figura 2: Distribución de probabilidad de plantas en los polígonos

En la Figura 2 presentada, se observa la distribución de la probabilidad de plantas para cada grupo de polígonos. Estos gráficos permiten identificar patrones de distribución y variabilidad en la densidad de plantas a lo largo de diferentes áreas. Por ejemplo, los polígonos con una alta densidad de plantas muestran una probabilidad de éxito mayor, lo que podría indicar que las condiciones de suelo y clima son óptimas para el crecimiento de la vegetación. En cambio, aquellos polígonos donde la probabilidad es más baja podrían estar afectados por factores limitantes como la disponibilidad de agua, la calidad del suelo, o la competencia entre especies.

6. Distribución de probabilidad de ocurrencia por especie

A continuación, se muestra la distribución de probabilidad de ocurrencia por cada una de las especies en la hectárea, la cual detalla la probabilidad de encontrar cada una de las 10 especies seleccionadas en el área planeada. La probabilidad se calculó considerando la proporción de cada especie respecto a un total de treinta polígonos antes proporcionados por el socio formador, la cual nos ayudó a obtener la probabilidad de la preexistencia de cada especie en nuestros modelos realizados posteriormente.

#	Especie	Distribución de probabilidad
0	Agave lechuguilla	0.06 %
1	Agave salmiana	0.29 %
2	Agave scabra	0.06 %
3	Agave striata	0.06 %
4	Opuntia cantabrigiensis	0.08 %
5	Opuntia engelmannii	0.06 %
6	Opuntia robusta	0.12 %
7	Opuntia streptacantha	0.10 %
8	Prosopis laevigata	0.13 %
9	Yucca filifera	0.04 %
Plantas totales		658

Cuadro 1: Especies y su demanda en una hectárea

Como se observa en el Cuadro 1, la especie con mayor probabilidad de ocurrencia es Agave salmiana con un 0.29 %, seguida por Prosopis laevigata y Opuntia robusta. Estas especies muestran una mayor representatividad en el área de estudio, lo que puede indicar su adaptación a las condiciones locales. En contraste, especies como Yucca filifera y Agave lechuguilla presentan una menor probabilidad de ocurrencia, lo cual sugiere una presencia más limitada. Estos resultados proporcionan una visión clara sobre la composición de la vegetación en la hectárea estudiada y son un punto de partida para la planificación de actividades de conservación y manejo de las plantas.

7. Método metaheurístico: Algoritmo Genético

Los algoritmos genéticos son técnicas de optimización estocástica inspiradas en los principios de la evolución natural. Simulan la "supervivencia del más apto", donde cada solución factible se representa como un cromosoma codificado por genes. A través de la selección y recombinación de cromosomas, estos algoritmos exploran diversas partes de la región factible, evolucionando gradualmente hacia mejores soluciones según una función objetivo o "fitness". Los operadores genéticos permiten preservar información valiosa mientras promueven la diversificación de soluciones. A continuación se muestra la información del algoritmo genético generado para el contexto del problema.

- **Tipo de cromosoma:** Numérico
- **Longitud:** $n = 658$
- **Criterio de inicialización:** 131 plantas fijas asignadas de manera pseudoaleatoria con cierta distribución de probabilidad de cada especie basada en datos de 30 polígonos, y el resto (no fijas) generadas de forma aleatoria con cierta distribución o demanda.
- **Criterio de infactibilidad:** Si la suma de la presencia de cada especie no es igual a su respectiva demanda, se elimina ya que es una solución infactible
- **Criterio de paro:** 100 generaciones
- **Función fitness:** Función Objetivo (minimizar la suma de la competencia de cada nodo con sus respectivos vecinos en toda la hectárea)
- **Criterio de selección:** Torneo. Se seleccionan los dos mejores cromosomas como padres.
- **Tamaño de población:** 5
- **Probabilidad de cruce:** 0
- **Puntos de cruce:** Ninguno
- **Probabilidad de mutación:** 0.5
- **Condiciones en la mutación:** Se toma un número aleatorio entre 0 y 1. En el cual si dicho número era menor a 0.5, se itera gen por gen generando un número aleatorio cual si es menor a 0.5 dicho gen intercambiaba lugar con otro de los lugares posibles a ser intercambiados. Pero si el primer número aleatorio es mayor a 0.5 se intercambian de manera aleatoria el 10 por ciento de los peores genes.
- **Criterio de reemplazo:** Se reemplazan los dos cromosomas con la peor función fitness si estos son superados por los hijos mutados

Con dichas condiciones se buscará y obtendrá una población final con soluciones efectivas y eficientes que nos permitirán reducir la competencia entre especies en cada uno de los cromosomas, dando un resultado preciso en nuestro contexto del problema.

8. Simulación de Montecarlo

La simulación de Monte Carlo es una herramienta eficaz para modelar la incertidumbre y generar predicciones utilizando distribuciones de probabilidad. Funciona iterando un modelo matemático miles de veces, generando combinaciones posibles y proporcionando un rango de resultados, en lugar de un valor fijo. Aunque su precisión aumenta con el número de iteraciones (comúnmente 10,000), esto implica una mayor demanda computacional. La efectividad de esta técnica depende de la calidad de los inputs, por lo que es crucial elegir correctamente las distribuciones de probabilidad y un modelo adecuado. En este caso, se utilizó la simulación de Monte Carlo para determinar las distribuciones de probabilidad de cada especie en las posiciones posibles dentro de la hectárea, identificando tendencias y frecuencias de aparición en las soluciones generadas.

9. Experimentación y resultados

9.1. Características de la PC y software

- **Modelo de la PC:** MacBook Pro (16 pulgadas, 2023)
- **Sistema Operativo:** macOS Sonoma, versión 14.6
- **Capacidad de disco duro:** 2 TB (1.16 TB disponibles)
- **Memoria Ram y capacidad:** 96 GB
- **Tipo de procesador:** Apple M2 Max
- **Número de núcleos:** 12
- **Software utilizado y versión:** Visual Studio Code (versión 1.94.2)

10. Tamaño del problema

- **Número de especies:** 10
- **Plantas por especie y en una hectárea:**

#	Especie	Demanda
0	Agave lechuguilla	42
1	Agave salmiana	196
2	Agave scabra	42
3	Agave striata	42
4	Opuntia cantabrigiensis	49
5	Opuntia engelmannii	38
6	Opuntia robusta	73
7	Opuntia streptacantha	64
8	Prosopis laevigata	86
9	Yucca filifera	26
Plantas totales		658

Cuadro 2: Especies y su demanda en una hectárea

- Número de variables consideradas: 2
- Número de parámetros: 4

11. Experimentación

Se realizaron 1000 simulaciones del proceso de siembra en nuestro algoritmo, obteniendo lo siguiente:

- Valor esperado del total de plantas existentes en una hectárea: 131 plantas
- Valor esperado de plantas por especie existentes en una hectárea:

Especie	Valor esperado
Agave lechuguilla	8
Agave salmiana	39
Agave scabra	8
Agave striata	8
Opuntia cantabrigiensis	10
Opuntia engelmannii	8
Opuntia robusta	15
Opuntia streptacantha	13
Prosopis laevigata	17
Yucca filifera	5

Cuadro 3: Especies y su demanda en una hectárea

- Valor esperado de plantas por especie a suministrar por hectárea:

Especie	Valor esperado
Agave lechuguilla	34
Agave salmiana	157
Agave scabra	34
Agave striata	34
Opuntia cantabrigiensis	39
Opuntia engelmannii	30
Opuntia robusta	58
Opuntia streptacantha	51
Prosopis laevigata	69
Yucca filifera	21

Cuadro 4: Especies y su demanda en una hectárea

- Valor esperado de competencia existente en una hectárea: 48,988.1
- Probabilidad de transición entre plantas:
- Orden en que se debe realizar la plantación:

	0	1	2	3	4	5	6	7	8	9
0	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
1	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
2	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
3	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
4	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
5	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
6	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
7	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
8	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040
9	0.064	0.298	0.064	0.064	0.074	0.058	0.111	0.097	0.131	0.040

Cuadro 5: Probabilidad de transición entre plantas

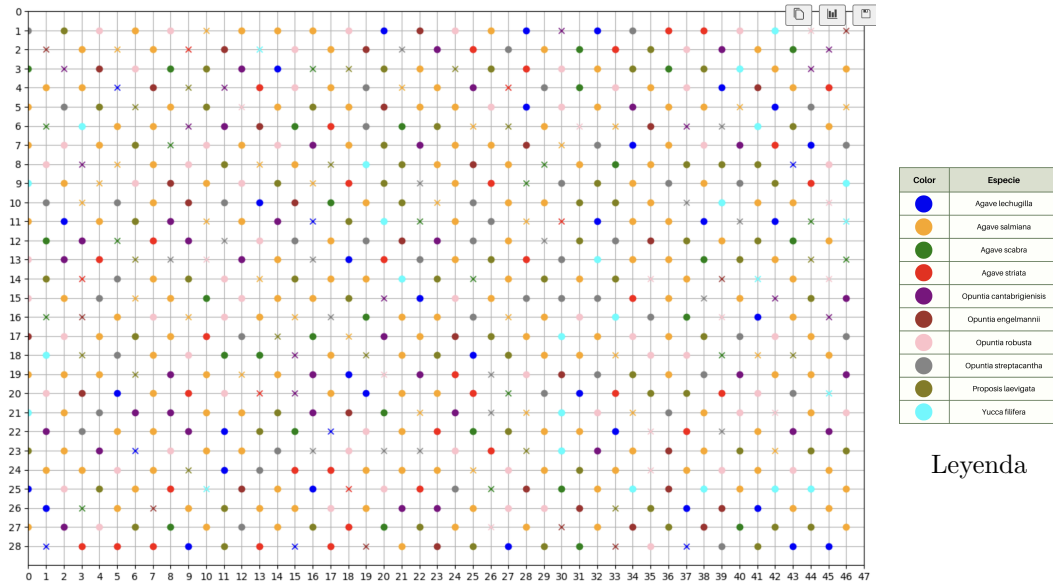


Figura 3: Mejor cromosoma

La figura anterior muestra la distribución óptima para la plantación de especies en nuestro polígono determinado. Cada punto en el mapa representa una especie de planta, codificada por un color específico, mientras que las 'x' indican plantas preexistentes en el sitio, y los puntos indican nuevas especies plantadas de acuerdo con la propuesta optimizada, permitiendo una planificación precisa al reducir los riesgos de monocultivo y asegurando que cada planta tenga el espacio y los recursos necesarios para crecer.

Tipo de generación	Promedio de competencia por nodo	Porcentaje de competencia
Peor	112.5	100 %
Random	87.12	77.54 %
Algoritmo Genético	74.45	66.17 %

Cuadro 6: Comparación de tipos de generación

Tiempo computacional: 754m 30.7s

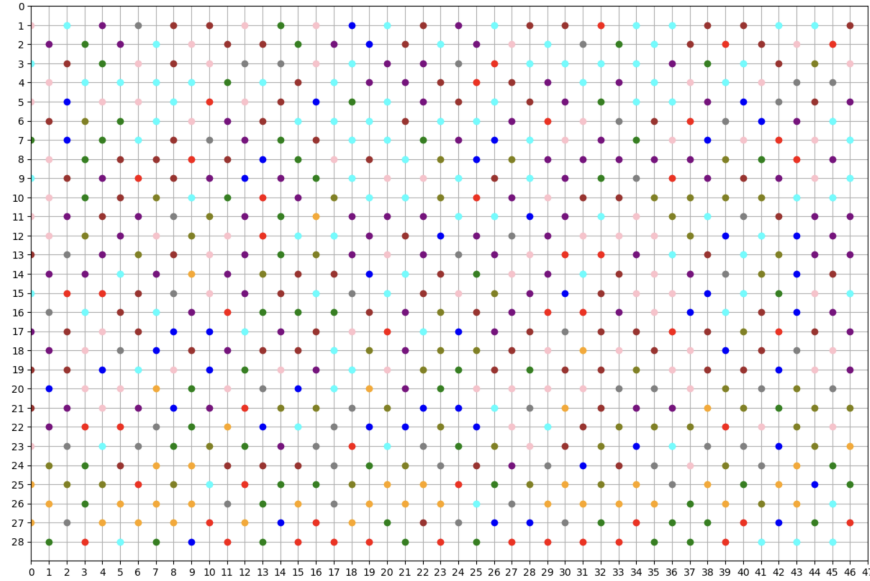


Figura 4: Ocurrencia por especie

En la figura proporcionada, se muestra la ocurrencia más común de cada posición en la hectárea. A través de la simulación de Monte Carlo, se generaron distintas distribuciones iniciales de las especies, lo que permitió obtener la distribución final por posición. Esta figura refleja la especie más común en cada ubicación, según las mejores soluciones obtenidas mediante el algoritmo genético.

12. Conclusiones y recomendaciones

En resumen, el algoritmo genético implementado optimiza el proceso de reforestación, ajustándose a la función objetivo de minimizar la competencia entre las especies, teniendo en cuenta tanto las ponderaciones de competencia como las distancias entre plantas. Además, se respetan las restricciones del número de plantas establecidas y se evita el monocultivo.

El método heurístico del algoritmo genético mostró una mejora constante en el valor de la función fitness a lo largo de las generaciones. Este enfoque aumenta significativamente la tasa de éxito en la reforestación, al considerar tanto la disposición espacial como los niveles de competencia, contribuyendo a los objetivos de restauración ecológica y sostenibilidad.

El resultado final fue la plantación de 658 plantas en una hectárea, cumpliendo con las restricciones de espacio y las especificaciones establecidas por el socio formador.

Si bien el algoritmo demostró un rendimiento positivo, su precisión y capacidad de optimización podrían haberse mejorado con información adicional. Datos más precisos sobre tasas de supervivencia de las especies, calidad del suelo y condiciones climáticas habrían permitido afinar mejor los parámetros del modelo. Actualmente, los datos utilizados fueron proporcionados de manera básica, sin un nivel de detalle profundo.

Además, disponer de datos históricos sobre reforestaciones anteriores habría enriquecido la simulación, permitiendo adaptar el modelo a necesidades específicas de cada área. Esto habría permitido explorar escenarios más realistas y tomar decisiones mejor informadas, optimizando también la forma de mutación de los cromosomas en el algoritmo.

Si bien el algoritmo es útil, presenta limitaciones en cuanto a su demanda computacional, siendo el tiempo de ejecución elevado debido al alto número de iteraciones necesarias para acercarse a una solución óptima. Una recomendación clave sería optimizar este proceso para mejorar la eficiencia sin comprometer la calidad de los resultados.

En conclusión, contar con un algoritmo de este tipo es extremadamente valioso para procesos de reforestación, ya que permite tomar decisiones adecuadas basadas en variables cruciales como el tipo de suelo, el tamaño del terreno y las especies a plantar. Esto es esencial para maximizar la supervivencia de las plantas y contribuir al mejoramiento del aire y la vida natural en las áreas reforestadas.

13. Bibliografía

Referencias

- [1] Amazon Web Services. (s.f.). *¿Qué es la simulación de Monte Carlo?*. Amazon Web Services. <https://aws.amazon.com/es/what-is/monte-carlo-simulation/#:~:text=Una%20simulaci%C3%B3n%20de%20Monte%20Carlo%20es%20un%20modelo%20probabil%C3%ADstico%20que,de%20la%20oficina%20es%20fija>.
- [2] International Business Machines Corporation. (s.f.). *¿Qué es la simulación Monte Carlo?*. IBM. <https://www.ibm.com/mx-es/topics/monte-carlo-simulation>
- [3] Universidad Pública de Navarra. (s.f.). *¿Qué plantas constituyen la flora invasora?*. Universidad Pública de Navarra. https://www.unavarra.es/herbario/invasoras/htm/concepto_i.htm#:~:text=Las%20plantas%20invasoras%20son%20aquellas,extenderse%20en%20un%20amplio%20territorio.
- [4] Badii, M. H., & Castillo, J. (2009). Distribuciones probabilísticas de uso común (Probabilistic distributions of common use). *Daena*, 4(1), 149-178. [http://www.spentamexico.org/v4-n1/4\(1\)%20149-178.pdf](http://www.spentamexico.org/v4-n1/4(1)%20149-178.pdf)
- [5] Del Castillo, B. (8 de diciembre de 2016). *El problema de los cuatro colores*. Revista C2. <https://www.revistac2.com/el-problema-de-los-cuatro-colores/>
- [6] Universidad Militar Nueva Granada. (s.f.). *FAEDIS*. http://virtual.umng.edu.co/distancia/ecosistema/odin/odin_desktop.php?path=Li4vb3ZhcY9pbmdlbml1cmlhX2NpdmlsL2ludmVzdGlnYWNPb25fZGVfb3BlcmFjaW9uZXNfaWkvW5pZGFkXzIv#slide_1

- [7] Guthrie, F. (1880). Note on the coloring of maps. *Proceedings of the Royal Society of Edinburgh*, 10.
- [8] Eos Data Analytics. (1 de diciembre de 2023). *Monocultivo En La Agricultura: Pros Y Contras*. <https://eos.com/es/blog/monocultivo/>
- [9] Iberdrola. (s.f.). *Reforestación: Qué es, Tipos y Casos de Éxito*. Iberdrola. <https://www.iberdrola.com/sostenibilidad/que-es-reforestacion>
- [10] Secretaría de Medio Ambiente y Recursos Naturales. (10 de octubre de 2019). *Monocultivos agrícolas y ganaderos, incompatibles con el entorno social*. Gobierno de México. <https://www.gob.mx/semarnat/articulos/monocultivos-agricolas-y-ganaderos-incompatibles-con-el-entorno-social#:~:text=El%20uso%20intensivo%20de%20la,la%20Alimentaci%C3%B3n%20que%20se%20celebra>