



Gradu Amaierako Lana / Trabajo de Fin de Grado
Fisikako Gradua/ Grado en Física

Análisis sociofísico de la España vaciada:
Estudio de la evolución de regiones despobladas
mediante el modelo de Axelrod

Egilea/ Autor/a:
Gonzalo Fernández de Barrena Machón
Zuzendariak/Directores/as:
Íñigo Etxebarria Altzaga
Jon Sáenz Agirre

© 2023, «Gonzalo Fernández de Barrena Machón»

Leioa, 2023eko ekainaren 22a / Leioa, 22 de junio de 2023

Resumen

Uno de los mayores problemas de España es la despoblación que sufren sus zonas rurales. Para analizar este fenómeno, los sociólogos han llevado a cabo diversos estudios durante los últimos años. En este trabajo se presenta una metodología novedosa aunando herramientas físicas y conceptos sociológicos para abordarlo desde una perspectiva sociofísica. Para llevarlo a cabo, se empleará el modelo de Axelrod, que consiste en comparar regiones definidas mediante vectores, siendo sus elementos las características propias de cada lugar. En la primera parte se evaluará la efectividad del método escogido para analizar este tipo de situaciones entre los años 2001 y 2022. Posteriormente, se tendrá como objetivo prever la situación de estas regiones en los años venideros mediante la aproximación de las proyecciones del número de habitantes en distintos lugares en el 2030. Además, también se proporcionarán mecanismos científicos para poder implementar los datos y las distintas variables sociales en los algoritmos creados con el mayor rigor posible.

Índice

1. Introducción	3
2. Modelo de Axelrod	6
2.1. Descripción del modelo	6
2.2. Cadenas de Márkov	7
3. Adaptación del modelo de Axelrod para fenómenos migratorios	9
3.1. Fundamento teórico	9
3.2. Implementación computacional	12
3.3. Factor de sensibilidad	13
4. Resultados	16
4.1. Resultados entre 2001 y 2022	17
4.2. Proyecciones	23
4.3. Factor de sensibilidad	25
4.3.1. Lerma	25
4.3.2. Valdorros	27
5. Conclusiones	30
Referencias	31

1. Introducción

La sociofísica es un campo interdisciplinario que combina la sociología y la física para estudiar y comprender los fenómenos sociales y colectivos desde una perspectiva científica. A pesar del auge de la sociofísica en el panorama científico actual, sus comienzos se remontan siglos atrás [1].

La primera referencia histórica que se tiene es del siglo XVII. El filósofo inglés Thomas Hobbes, tras un encuentro con Galileo Galilei, intentó relacionar los estudios de éste sobre el movimiento con el comportamiento de la sociedad. Así lo reflejó en su obra *De Corpore* (1655), a pesar de hacerlo de manera tangencial [2].

Tras los intentos de Hobbes, no fue hasta 1803 cuando el pensador y economista francés Henri de Saint-Simon le propuso a su alumno Auguste Comte, matemático, la idea de intentar describir los comportamientos sociales mediante las leyes de la física y la biología. El propio Comte, en la revista *Le Producteur*, fue el primero en definir el término *sociofísica* [3]:

La sociofísica es la ciencia que se dedica a estudiar los fenómenos sociales como si se tratasen de fenómenos astronómicos, físicos, químicos y psicológicos. Es decir, que los asuntos sociales también están sujetos a leyes naturales e invariables.

Más adelante, el matemático y sociólogo belga Adolphe Quetelet planteó modelar las conductas sociales usando probabilidad matemática y estadística social. En 1835, en su *Ensayo de sociofísica: El hombre y el desarrollo de sus facultades* mencionó la idea de describir los comportamientos humanos mediante una distribución normal. Tanto Quetelet como Saint-Simon y Comte formaron la primera generación de sociofísicos [4].

A mediados del siglo XX, físicos y astrofísicos como Arthur Iberall y John Q. Stewart demostraron que algunos aspectos de las interacciones sociales podían ser explicadas utilizando analogías con la Ley de Gravitación Universal de Newton. Además, fueron pioneros en comparar los sistemas sociales con los sistemas complejos [5].

El punto álgido de la sociofísica, en cambio, llegó en la segunda mitad del siglo. A partir de la década de 1950, se realizaron avances significativos en la aplicación de modelos físicos en estudios sociales, útiles para analizar las evoluciones y dinámicas culturales o prever las intenciones de voto de los ciudadanos, entre otras. Modelos como el de Ising, Potts o Axelrod fueron de gran importancia para alcanzar este progreso.

El modelo de Ising es un modelo de física estadística que fue propuesto por el físico Ernst Ising en 1925 para estudiar los comportamientos magnéticos de los materiales [6]. Sin embargo, en la década de los cincuenta, los científicos comenzaron a aplicar el modelo a problemas sociales. Su funcionamiento es simple: se representa a los individuos de una población como espines $\pm\frac{1}{2}$, que pueden estar alineados en sentido positivo ($+\frac{1}{2}$) o negativo ($-\frac{1}{2}$). Cada individuo interactúa con sus vecinos cercanos y tiende a alinearse con ellos. La interacción entre los individuos puede modelarse mediante una función de energía, que determina la preferencia de cada individuo por un estado en particular. Dados los dos estados en los que pueden encontrarse los espines, el uso del modelo de Ising preva-

lece en situaciones duales, como la propagación de enfermedades o la polarización política.

Por su parte, el modelo de Potts, desarrollado por el matemático australiano Renfrey Potts, también ha sido utilizado en la sociofísica para estudiar la dinámica de los sistemas sociales [7]. En este modelo, que es una generalización del modelo de Ising, los agentes se distribuyen en una red y pueden tener diferentes estados o características. Cada agente interactúa con sus vecinos y tiende a adoptar el estado de los agentes vecinos con los que comparte características similares [8]. La principal diferencia entre el modelo de Potts y el de Ising es el número de tipos de espín; el primero permite que el espín esté en un número n de estados, mientras que el segundo, dos. El modelo de Potts se suele utilizar para estudiar la migración celular o la segregación espacial en una población.

Finalmente, encontramos al modelo de Axelrod entre los más usados en la sociofísica. Fue propuesto por el científico y politólogo Robert Axelrod para estudiar la expansión de las opiniones o culturas entre personas que comparten un determinado espacio. Se basa en la teoría de juegos y ha sido utilizado para estudiar la cooperación y la competencia en diversos campos, como la biología evolutiva, la economía y la política [9].

Además de los modelos mencionados, la dinámica de poblaciones se ha solido abordar desde la resolución de ecuaciones diferenciales, como se muestra, en un caso simplificado, en la ecuación 1.

$$\frac{dP}{dt} = nP - mP + I - E \quad (1)$$

donde P es la población de cada pueblo; n , la tasa de natalidad; m , la tasa de mortalidad; I , la inmigración, y E , la emigración.

Son notables los casos de las ecuaciones de Lotka-Volterra, donde se describe la dinámica de sistemas biológicos en los que dos especies interactúan [10] o las ecuaciones de Arditi-Ginzburg, que estudia las dinámicas entre depredadores y presas usando proporciones entre ellas [11]. Mediante las correspondientes analogías, se pueden modelizar dinámicas poblacionales.

El motivo principal de la elección del modelo de Axelrod en lugar de estos métodos ya estudiados obedece a que se ha querido abrir una nueva manera de estudiar dinámicas migratorias.

Respecto al futuro de la sociofísica, en los próximos años, se espera que los avances aumenten de manera significativa. Las mejoras computacionales y el incremento de la cantidad de datos sociales disponibles permitirán abordar problemas de mayor magnitud y complejidad, obteniendo resultados de mayor calidad. Además, con la irrupción en los últimos años de otras ramas de estudio, como la econofísica, y su fusión con la sociofísica, se prevén estimaciones aún más concretas.

Mediante este trabajo se quiere implementar un algoritmo computacional escrito en Fortran 90 con el objetivo de realizar un análisis de la evolución poblacional de cualquier sociedad. En este caso, se estudiará el crecimiento y decrecimiento habitacional de

los pueblos afectados por la despoblación de las zonas rurales en la provincia de Burgos, en concreto, en la Comarca del Arlanza. A todas las regiones damnificadas por este fenómeno también se les conoce como la España vaciada. Para llevarlo a cabo, se modificará el modelo de Axelrod y se adaptará a las necesidades que nuestro problema requiere.

Por último, otro de los principales retos que se quiere abordar es la transcripción matemática de la realidad, es decir, definir numéricamente las características de un pueblo, de manera que se puedan valorar los atractivos o carencias de un lugar mediante parámetros cuantificables. El modelo de Axelrod permitirá estimar el impacto que tendría en la evolución habitacional de un pueblo modificar dichos parámetros.

2. Modelo de Axelrod

El modelo de Axelrod es un modelo matemático creado por el politólogo y matemático Robert Axelrod para estudiar la evolución y expansión de ideas, opiniones o culturas. Para ello, a cualquier persona o grupo se le asocia un vector formado por agentes, también llamados rasgos culturales. La idea principal de este modelo es que los vectores con, al menos, un agente en común tenderán a interactuar entre ellos y existirá una probabilidad concreta de que otras de sus características se vuelvan comunes. Juega un rol muy importante en los análisis de diseminación cultural [12].

2.1. Descripción del modelo

En palabras del propio Axelrod, *una persona es más propensa a relacionarse con alguien que hable su mismo idioma, y, a medida que la conversación avance, ambos tenderán a unos patrones discursivos similares*. Su traducción matemática sería la siguiente:

En primer lugar, se define el número de agentes totales f que se tendrán en cuenta para caracterizar a cada individuo. A cada rasgo se le asigna un valor entre 1 y q , siendo q el valor máximo que puede adoptar un rasgo. Cada persona o colectivo i estará asociado a un vector x_i de dimensión f ; habrá q^f opciones culturales distintas. Interactuarán entre ellos los vectores que tengan algún atributo común (llamémosle k al número de rasgos idénticos entre dos vectores), y $\frac{k}{q}$ será la probabilidad de que otra característica más se les vuelva común [13].

En la figura 1 se muestra un esquema de cómo funciona el modelo de Axelrod y, así, se podrán analizar las reglas dinámicas de evolución que se aplican en el modelo. Consideramos que existen $f=3$ rasgos culturales, pudiendo tomar cada uno de ellos $q=10$ valores. El vector característico de uno de los individuos es $(4,5,7)$, mientras que el del otro vector es $(4,6,2)$. El agente que tienen en común es el primero. Como coinciden en uno de los tres agentes, la probabilidad de que alguno de los otros dos rasgos se modifique será de un $\frac{1}{3}$. Por lo tanto, habrá una probabilidad de $\frac{2}{3}$ de que ambos vectores permanezcan iguales a su situación inicial o, por el contrario, en $\frac{1}{3}$ de las veces, uno de los otros dos agentes se convertirá común, es decir, que hay $\frac{1}{6}$ de posibilidades de que se vuelva común el segundo agente, y otro $\frac{1}{6}$, el tercero.

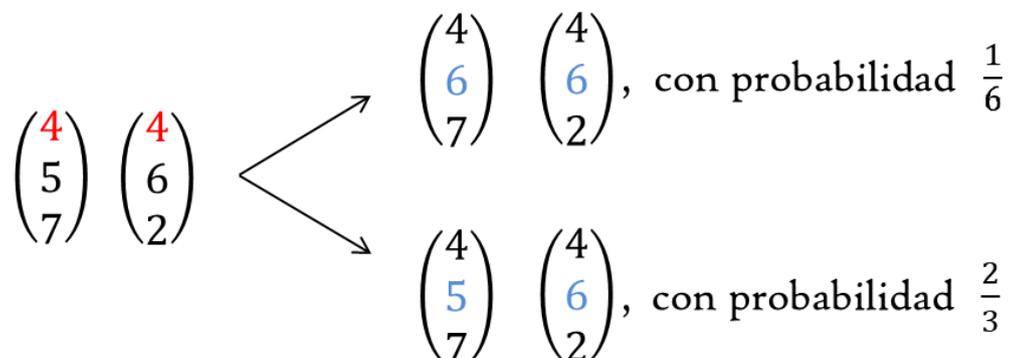


Figura 1: Interacción de dos vectores en una simulación del modelo de Axelrod con $f=3$ y $q=10$.

El estudio de una red de n individuos requiere de los siguientes pasos:

- Seleccionar a un ciudadano o colectivo r aleatoriamente.
- Seleccionar a uno de los **vecinos** del pueblo del ciudadano o colectivo r , denominado s , también al azar. En este caso, **vecino** es un término más amplio, pues se considera vecino a aquel que esté dentro de un determinado radio, con centro en el ciudadano seleccionado en el paso anterior. Si el radio es muy grande, se tendrá en cuenta a toda la red, por lo que sería completamente aleatorio.
- r y s interactuarán entre ellos y se estudiará la probabilidad de que se vuelva común un agente que no lo era. Ésta será $\frac{n_{rs}}{f}$, siendo n_{rs} el número de agentes que tienen en común los vectores r y s .
- En caso de que se realizara el cambio, el vector r adoptaría el mismo valor en alguno de sus agentes que el vector s . Este agente sería escogido de manera arbitraria.

Este proceso pertenece a un tipo de algoritmo de evolución denominado cadena de Márkov.

Un ejemplo de la evolución de una simulación del modelo de Axelrod en una red cuadrada $L \times L$ ($= n$) se puede ver reflejada en la figura 2, tomada de [14].

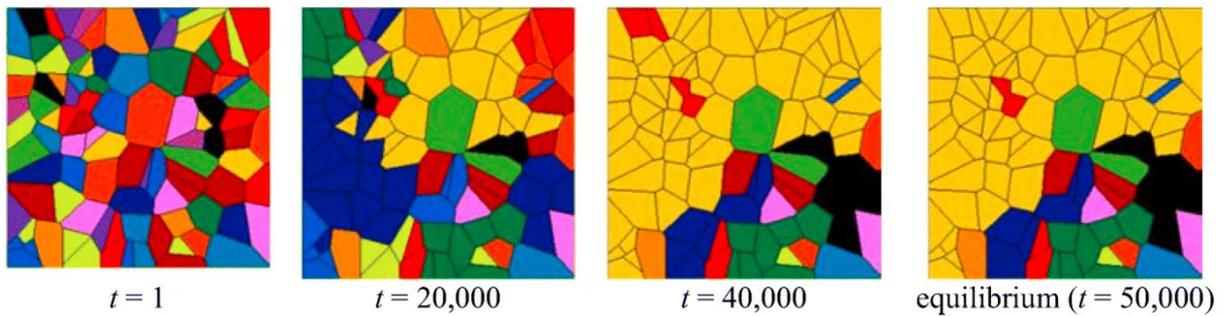


Figura 2: Evolución en una simulación del modelo de Axelrod con $f=10$ y $q=15$ en distintos pasos en la evolución de un sistema [14]. El radio de interacción es 1.

2.2. Cadenas de Márkov

Una cadena de Márkov es un tipo de sucesión que se utiliza para describir la evolución de un proceso estocástico en el tiempo. La característica principal de este modelo es que la probabilidad de modificación de un estado en un instante determinado únicamente depende del paso anterior, y no de los previos al anterior. Su transcripción matemática para un tiempo discretizado se puede expresar mediante la propiedad de Márkov (ecuación 2), siendo P la probabilidad de cambio y X_0, X_1 y X_2 , las secuencias de variables aleatorias, por lo que X_n es el estado en el instante n [15].

$$P[X_{n+1} | X_0, X_1, X_2 \dots, X_n] = P[X_{n+1} | X_n] \quad (2)$$

De un modo más sencillo, se puede entender que las cadenas de Márkov no tienen memoria, ya que no tienen en cuenta el historial del proceso para determinar la probabilidad de cambio de un estado a otro. Esta es una de sus ventajas, pues reduce el coste computacional de las simulaciones, así como el tiempo de ejecución.

Por esta razón, las aplicaciones de las cadenas de Márkov en la ciencia son numerosas. Por un lado, en física encontramos su utilidad en las cadenas de Ehrenfest, que sirven para modelar el intercambio de moléculas de gas entre dos urnas [16]. También tiene diversos usos en física estadística [17], meteorología [18] y modelos epidemiológicos [19].

En la sociofísica, una aplicación común de las cadenas de Márkov es analizar la intención de voto y mejorar los pronósticos electorales. Además, se está implementando en nuevos campos de estudio, entre ellos, la prevención de secuestros [20] y criminalidad [21].

Aún así, a pesar de los ejemplos anteriores, al realizar analogías entre la sociología y la física, hay que tener en cuenta que los elementos físicos están sometidos a limitaciones que las personas no lo están. Por ejemplo, elementos clásicos como un cilindro en un plano inclinado se rige bajo gradientes de energía potencial, o, en casos de partículas relativistas, están limitados por la velocidad de la luz. En cambio, a la hora de analizar fenómenos sociofísicos, hay que tener en cuenta la espontaneidad inherente al ser humano, que da pie a cambios espontáneos más difíciles de prever.

Esto es otro punto a favor de las cadenas de Márkov, pues hay circunstancias inesperadas que llevan a tomar decisiones que son complicadas de vaticinar. Entre ellas, se puede considerar una enfermedad repentina que requiera a un ciudadano acudir a un centro médico con frecuencia. Esto implicaría una necesidad de cercanía de domicilio a una instalación sanitaria. Por lo tanto, los procesos que utilizan este tipo de sucesiones, donde no se tiene en cuenta todo el desarrollo evolutivo sino solamente el paso previo, paliar el efecto de este tipo de circunstancias inesperadas.

En definitiva, las cadenas de Márkov son modelos matemáticos que se utilizan para describir la evolución de procesos estocásticos en el tiempo. Su aplicación en la sociofísica es amplia, ya que se pueden utilizar para modelar una amplia variedad de fenómenos sociales y físicos.

3. Adaptación del modelo de Axelrod para fenómenos migratorios

De los modelos explicados, ninguno destaca por su capacidad innata para analizar fenómenos migratorios. Por lo tanto, se ha modificado y adaptado el modelo de Axelrod para poder llevar a cabo dicha empresa. La razón por la que se ha escogido el modelo de Axelrod frente al modelo de Ising y al modelo de Potts es la aleatoriedad del primero al comparar individuos o colectivos; el modelo de Ising y el modelo de Potts estudian una posibilidad de cambio entre un individuo y sus vecinos, mientras que el modelo de Axelrod selecciona a dos ciudadanos aleatorios, independientemente de su localización. De una manera más científica, se podría decir que el modelo de Ising y el modelo de Potts obedecen al principio de localidad, es decir, que dos objetos alejados uno de otro no pueden influirse mutuamente de manera instantánea. Esto es un punto clave en la elección, porque, teniendo en cuenta el estudio que se va a llevar a cabo en el presente trabajo, carece de sentido abordar este problema partiendo de la premisa de que un ciudadano sólo puede emigrar a los pueblos de alrededor de su pueblos de origen.

3.1. Fundamento teórico

Antes de nada, es conveniente anticiparse a las ventajas y desventajas que se pueden intuir al aplicar el modelo de Axelrod al estudio de fenómenos migratorios.

Por un lado, la ya mencionada no localidad con la que se trabaja en el modelo de Axelrod acerca a la realidad las simulaciones realizadas, siendo este uno de los puntos a favor del modelo de Axelrod. Otra de sus ventajas es que permite una gran flexibilidad a la hora de decidir los valores q de los agentes f , así como el número de éstos. Esto es de gran utilidad, porque, dependiendo del estudio que se quiera llevar a cabo y de la exactitud que se requiera, el algoritmo permite reducir o desprestigiar los parámetros que así se consideren de manera muy sencilla. En este mismo trabajo, para realizar el análisis evolutivo de la Comarca del Arlanza, se han definido a los pueblos mediante vectores de 6 elementos, pues se han considerado suficientes para describir el comportamiento que se quería estudiar del sistema. En caso de que querer hacer un análisis más exhaustivo, bastaría con cambiar la dimensión de los vectores. Esto permite ahorrar coste computacional a la hora de hacer estudios más generales.

Relacionado con este último punto, también se encuentra uno de los mayores inconvenientes del algoritmo: el tiempo de ejecución. Este modelo tan solo permite estudiar una posible migración en cada paso, por lo que, si el número de habitantes a estudiar es muy amplio, serían necesarios muchos pasos para realizar un análisis con éxito; de lo contrario, se quedaría sin estudiar la situación de muchos ciudadanos. Por otro lado, también se puede considerar esta no simultaneidad como una pérdida del realismo del algoritmo, pues, en la realidad, sí que son múltiples los movimientos que se hacen en el mismo instante, aunque se han buscado las maneras de solventar este problema como se verá a lo largo de este trabajo.

En cuanto a su proceso teórico, primero de todo, se asocia a cada pueblo i un vector $v_i = (x_1, x_2, x_3, \dots, x_n)$. Cada uno de los elementos que forman el vector serán rasgos característicos de cada región en cuanto a sus fortalezas y debilidades: sanidad, ocio, empleo,

educación, comunicación de transporte y arraigo. Estos datos también se podrían conseguir mediante estudios demoscópicos, aunque, hasta la actualidad, no se han realizado respecto a los temas que se tratan en este trabajo.

Para una obtención de información más amplia, cada pueblo estará dividido en tres grupos: jóvenes (entre 0 y 19 años), adultos (entre 20 y 51) y ancianos (entre 52 y 99). Los ciudadanos r del pueblo i estarán compuestos por el vector complementario, es decir, $v_{ri} = (q - x_1, q - x_2, q - x_3, \dots, q - x_n)$.

La idea sociofísica que representa el vector complementario juega un papel muy importante en el trabajo, por lo que es necesario explicar de qué trata. El significado sociológico que le corresponde al vector complementario de un pueblo equivale a las necesidades de los ciudadanos; los pueblos con agentes altos tendrán vectores complementarios bajos y viceversa. Dicho de otra manera, la preocupación de los ciudadanos respecto a una variable será mínima si es variable está dotada de un valor alto, mientras que las variables con valores bajos tendrán valores mayores en el vector complementario, refiriéndose así al grado de preocupación de los ciudadanos. Mediante un ejemplo se puede exponer que un pueblo A cuyo agente de sanidad es de 9 sobre 10 tendrá un agente complementario de 1 sobre 10, siendo esto un valor muy bajo, y siendo la sanidad un agente que no preocupa a los ciudadanos del pueblo A, puesto que ya tienen acceso a ella.

Acto seguido, se selecciona otro pueblo j al azar y, a su vez, uno de sus ciudadanos s (con su respectivo bloque generacional), al que se le designará el vector v_{sj} . La compatibilidad entre un pueblo y otro viene dada por la ecuación 3, donde $v_{n,ri}$ y $v_{n,sj}$ son los vectores que describen los ciudadanos de los pueblos i y j respectivamente, mientras que $\alpha_{n,sj}$ indica el porcentaje que se le atribuye al rasgo n para un ciudadano j que pertenece a un determinado bloque de edad. El factor α_n , que indica la importancia de un agente para cada bloque generacional, depende únicamente del ciudadano j , que es sobre el que hará el estudio migratorio. El ciudadano i sólo se selecciona para compararlo con j , pero nunca para analizar si se muda de pueblo o no.

$$\text{compatibilidad} = \sum_{n=1}^f (v_{n,ri} - v_{n,sj}) \alpha_{n,sj} \quad (3)$$

Dependiendo del resultado de la compatibilidad, se presentan dos opciones:

- Si $\text{compatibilidad} \geq 0$, el ciudadano j no se moverá de su pueblo natal s .
- Si $\text{compatibilidad} < 0$, existirá la posibilidad de cambio del ciudadano j al pueblo r . Esta posibilidad es dependiente del valor de la compatibilidad. Para ello, se selecciona un número aleatorio c entre 0 y $f \cdot q$.
 - Si $c < |\text{compatibilidad}|$, el ciudadano j se mudará al pueblo r .
 - Si $c \geq |\text{compatibilidad}|$, el ciudadano j permanecerá en el pueblo s .

Es trivial ver que, si $\text{compatibilidad} \geq 0$, el ciudadano j permanecerá en el pueblo s , porque indica, mediante los vectores complementarios, que las necesidades de los habitantes i del pueblo r serán mayores que las de los ciudadanos j del pueblo s , y, por ende,

los servicios del pueblo r serán peores. Lo que no es tan intuitivo de comprender es que, si el número aleatorio $c \geq |\textit{compatibilidad}|$, el ciudadano j permanecerá en el pueblo s . Se puede suponer que existe una recta entre 0 y $f \cdot q$. Se dividirá en dos, y el tamaño de cada parte dependerá del valor absoluto de $\textit{compatibilidad}$. Se usa el valor absoluto porque el signo de la compatibilidad es ya indiferente, pues sólo se realiza este paso si, en primera instancia, la compatibilidad era negativa. El primer fragmento corresponde a la compatibilidad del ciudadano j con el pueblo s ; el segundo, con la del pueblo r (la compatibilidad con el pueblo r sería la contraria: $1-|\textit{compatibilidad}|$). Por lo tanto, al seleccionar un número aleatorio c entre 0 y $f \cdot q$, hay más probabilidades de que se encuentre en la parte que más compatibilidad tenga. De manera ilustrativa, se puede ver cómo sería la recta para un caso en el que $|\textit{compatibilidad}| = 7$, $f = 3$ y $q = 10$ en la figura 3.

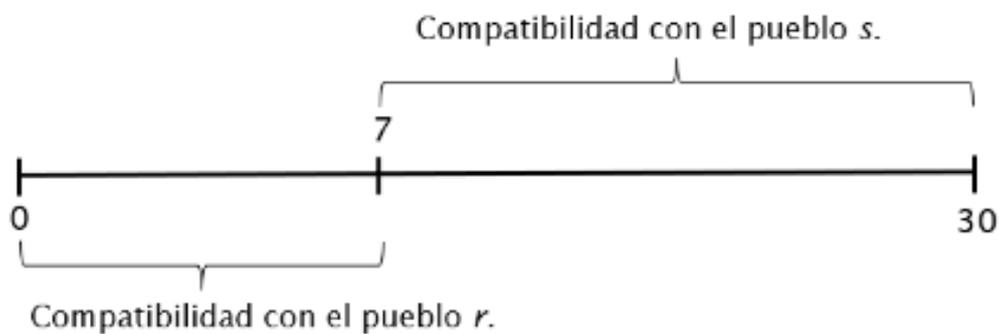
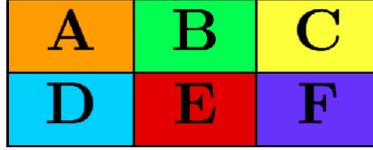


Figura 3: Ilustración de la recta de compatibilidad de los pueblos r y s para una situación de $|\textit{compatibilidad}| = 7$, $f = 3$ y $q = 10$.

Este proceso se puede realizar un número n de veces, y, así, hacer la media de todos los pasos con el fin de obtener un resultado mucho más exacto.

Esta explicación se puede entender más fácil mediante el esquema de la figura 4, donde se representa una red de seis pueblos con $f = 3$ y $q = 10$.

Finalmente, aunque no se haya utilizado para hacer las posteriores simulaciones dada la cercanía entre los pueblos que se han estudiado, también se ha querido tener en cuenta la distancia entre municipios. Éste es un factor clave a la hora de que un ciudadano se planteé mudarse de su pueblo de origen a otro, por lo tanto, no se puede pasar por alto. Para tenerlo en cuenta, hay que estudiar cuáles son los núcleos atractivos de una región. A partir de una determinada distancia, la transición es inviable, por lo que los ciudadanos de un pueblo sólo podrán irse a los que se encuentren dentro del rango predeterminado.



↓

De manera aleatoria, se escogen dos pueblos.

Pueblo B	Pueblo F
$\begin{pmatrix} 3 \\ 7 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 5 \\ 2 \end{pmatrix}$

↓

Con el vector complementario se consigue el vector de los ciudadanos.

Pueblo B	Pueblo F
$\begin{pmatrix} 7 \\ 3 \\ 8 \end{pmatrix}$	$\begin{pmatrix} 6 \\ 5 \\ 8 \end{pmatrix}$

↓

Aleatoriamente se elige el bloque generacional del ciudadano F ; es joven. Se utiliza la ecuación 2 para calcular la compatibilidad. Se supone $\alpha_n=1$.

$$compatibilidad = (7 - 6) \cdot \alpha_{1,joven} + (3 - 5) \cdot \alpha_{2,joven} + (8 - 8) \cdot \alpha_{3,joven}$$

$$compatibilidad = -1$$

↓

$compatibilidad < 0$. Hay probabilidad de cambio.

Seleccionamos un número aleatorio c : 21.

↓

Como $c \geq |compatibilidad|$, el ciudadano F permanecerá en el pueblo F .

Figura 4: Esquema de la adaptación del modelo de Axelrod para modelos migratorios.

3.2. Implementación computacional

El programa se ha escrito en Fortran 90. A la hora de implementar computacionalmente el modelo, se ha intentado respetar con la mayor exactitud posible sus cimientos teóricos. A pesar de ello, se han tenido que hacer algunas concesiones.

Antes de nada, el problema que hay que abordar desde la perspectiva computacional

es el coste de ejecución. El modelo de Axelrod, y, en consecuencia, cualquier algoritmo que lo emplee, es un método que, en cada paso computacional, tan solo estudia el movimiento de un vector. Por lo tanto, para estudiar una red formada por una gran cantidad de vectores son necesarios un alto número de pasos. Esto puede ser un inconveniente, por eso, para solucionarlo, se pueden llevar a cabo varias vías. La aquí tomada consiste en que las regiones con un número muy elevado de habitantes solamente reciben ciudadanos. Dicho con otras palabras y haciendo el símil con la realidad, en las grandes urbes se pueden despreciar los habitantes que emigran de ellas, pues es un valor muy inferior al del total de sus ciudadanos, así que pueden ser denominadas como regiones únicamente atractivas. De este modo, no se hacen pasos en vano, porque la probabilidad de que un ciudadano de una gran ciudad se mude a un pueblo más pequeño y con peores incentivos es ínfima.

Partiendo de la misma base, otro de los aspectos a tener en cuenta a la hora de implementar un algoritmo que usa el modelo de Axelrod es que no considera la simultaneidad, es decir, en cada paso computacional, sólo se estudia el caso de un único individuo, dejando de lado al resto de ciudadanos o pueblos. En la realidad no sucede así; pueden darse más de una migración a la vez en la misma unidad de tiempo. Para solucionar esto, se harán los pasos computacionales determinados un número n veces y se hará la media paso a paso. Esto, aunque ralentice el programa, se asemeja más a las situaciones reales, que es el objetivo principal de este trabajo.

Tanto para el cómputo total de los ciudadanos de cada pueblo como para simular la cantidad de jóvenes, adultos y ancianos que viven en cada pueblo se ha mantenido el realizar la media tras realizar los respectivos n procesos para obtener los resultados óptimos. Además, gracias a ello, las gráficas siguen tendencias más constantes, eliminando, así, cambios aleatorios.

Además de todo esto, el programa también está diseñado para obtener más información de la que se le presupone al modelo de Axelrod. Dada la ecuación 3, también se permite obtener datos segregados dependiendo del bloque generacional que se quiera estudiar.

Finalmente, también se ha añadido al programa un método para que se tenga en cuenta el factor de la sensibilidad a la hora de introducir los datos que formarán los vectores que definen cada uno de los pueblos. Esto se hará obteniendo la raíz del error cuadrático medio (RECM) de la variable que se quiera analizar. El valor de la variable escogida irá alternando, para que, finalmente, se calcule la raíz del error cuadrático medio con cada uno de ellos. De esta manera, el valor que minimice el error será el que más se ciña a la realidad.

3.3. Factor de sensibilidad

Cada vez son más los sociólogos que aúnan sus conocimientos con herramientas matemáticas. Esto, además de proporcionar nuevas maneras de estudio, también ha dado pie a grandes retos. Uno de ellos se presenta en este trabajo: cuantificar la realidad.

Los vectores que definen cada pueblo están formados por agentes con poco rigor teórico, y su elección se basa en encuestas susceptibles de errores difíciles de evaluar, no es fácil

determinar la calidad de los servicios de un pueblo en una escala numérica. Esta inexactitud puede hacer que los resultados de las simulaciones no sean los esperados, por lo que la definición de los vectores es un aspecto del que dependerá el buen desarrollo del algoritmo.

Una de las maneras para abordar este inconveniente es mediante la raíz del error cuadrático medio. La raíz del error cuadrático medio (RECM) es una medida utilizada para evaluar la precisión de un modelo de predicción. Representa la diferencia cuadrática promedio entre los valores predichos por el modelo y los valores reales o verdaderos. Su fórmula queda reflejada en la ecuación 4.

$$RECM = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2} \quad (4)$$

donde \hat{Y}_i es la predicción del modelo; Y_i , el resultado real esperado, y n es el número de predicciones.

Como se muestra en la ecuación 4, el error cuadrático medio y su raíz elevan los errores al cuadrado. Por lo tanto, se puede deducir que se penalizan más los errores grandes, lo que significa que el ECM y la RECM dan mayor importancia a los errores más significativos.

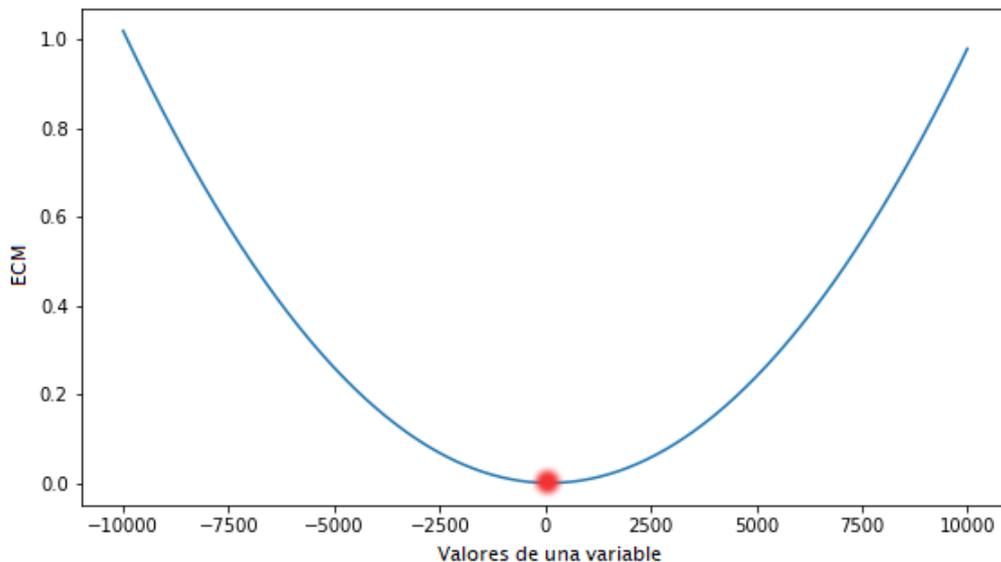


Figura 5: Gráfica teórica del error cuadrático medio frente a distintas predicciones de una variable cuando el resultado real esperado es 0 [22].

Desde un punto de vista teórico, la gráfica resultante al representar el error cuadrático medio frente a los distintos valores que toma la variable que se analiza es una parábola; se magnifican los errores lejanos del valor real esperado. Para un caso en el que el resultado real esperado es 0, se muestra en la figura 5 cómo es su curva. Como era de esperar, los errores cuadráticos medios más alejados del 0 aumentan considerablemente, mientras que los valores de las predicciones del modelo que son tendentes a cero, hacen que disminuya el error. Así, se considera que el valor que minimiza el ECM en el caso de la figura 5 es 0,

y será el que haga que las simulaciones tengan una mayor tendencia a la realidad.

4. Resultados

Al simular el programa, se han tenido en cuenta varios municipios de la Comarca del Arlanza, doce, en concreto: Cilleruelo de Abajo, Covarrubias, Lerma, Mecerreyes, Nebreda, Presencio, Quintanilla del Agua, Santa María del Campo, Tordómar, Valdorros, Villafruela y Villalmanzo (figura 6). Además, a pesar de formar parte de otra comarca, también se ha tenido en cuenta a la ciudad de Burgos, pues es notable la influencia que tiene en los pueblos arlanceños dada su cercanía. Recordemos que cualquier simulación realizada con el modelo de Axelrod es un proceso lento, ya que en cada paso se estudia la transición de un solo individuo. Por lo tanto, para optimizar el coste computacional, se le ha adjudicado a la ciudad de Burgos un número insignificante de ciudadanos. De esta manera, Burgos seguirá siendo un núcleo altamente atractivo en la región, pero se evitará que el algoritmo tenga que recorrer todos los habitantes de la ciudad en vano; se asume que la gran mayoría de los habitantes de la ciudad permanecerían en ella debido a sus incentivos.



Figura 6: Mapa de la Comarca del Arlanza [23].

Respecto a cómo se ha definido cada pueblo, se ha optado por hacerlo mediante la intuición y sentido común, siendo éste un primer acercamiento a lo que sería una simulación hecha con datos seleccionados por físicos o matemáticos. Se han considerado seis los agentes suficientes para definir cada pueblo ($f = 6$): sanidad, ocio, empleo, educación, movilidad y arraigo.

Además, con el fin de comprobar la efectividad del programa, primero se estudiará el periodo entre 2001 y 2022. Así, se pueden comparar los datos reales con los obtenidos en la simulación. A continuación, se proyectarán los habitantes de cada pueblo entre los años

2022 y 2030. Por último, se analizará la RECM de las variables, con el fin de mejorar los valores que se han asignado a los vectores de los pueblos.

4.1. Resultados entre 2001 y 2022

Gracias al Instituto Nacional de Estadística (INE) [24], se puede obtener la evolución en el número de habitantes de cualquier municipio español entre los años 2001 y 2022. Así, en la gráfica 7 se representa la evolución real de los pueblos seleccionados.

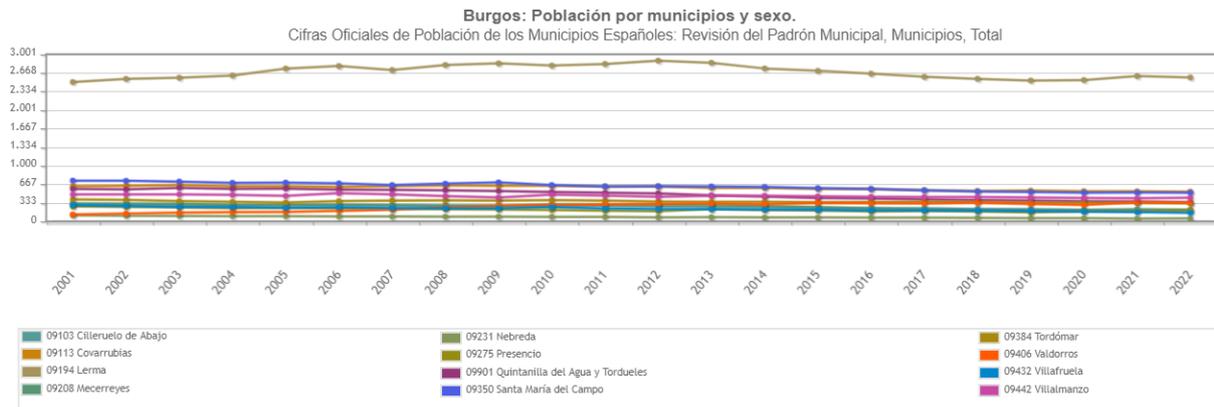


Figura 7: Datos reales de la evolución habitacional de distintos pueblos entre 2001 y 2022 [24].

Para hacer un análisis más concreto y, posteriormente, poder comparar los datos reales con los simulados, en la Tabla 1 se facilitan los valores de la gráfica 7.

Pueblos	Población (2001)	Población (2022)	Variación (%)
Cilleruelo de Abajo	314	210	-33,1
Covarrubias	633	541	-14,5
Lerma	2502	2537	+1,4
Mecerreyes	320	103	-39,7
Nebreda	110	53	-51,8
Presencio	268	184	-31,3
Quintanilla del Agua	583	360	-38,2
Santa María del Campo	732	517	-29,4
Tordómar	395	334	-15,4
Valdorros	123	290	+141,5
Villafruela	290	176	-39,3
Villalmanzo	483	420	-13,0

Tabla 1: Población de distintos pueblos en 2001 y en 2022, con su correspondiente variación.

Ahora, se va a realizar la simulación con el algoritmo creado partiendo de los datos reales de 2001. Para ello, se ha considerado que 21 años equivalen a 50000 pasos computacionales. Esta aproximación se obtiene mediante una regla de tres, que da como

resultado que 1 año virtual corresponde a 2381 pasos computacionales. Además, se realizará este proceso 5 veces, siendo el resultado final la media de todos ellos. En la Tabla 2 están los vectores que definen cada pueblo, y, en la Tabla 3, los coeficientes $\alpha_{n,sj}$ que se han utilizado para asignar la importancia que le da cada bloque generacional a cada una de las características que conforman el municipio. Hay que tener en cuenta que, en el año 2007 (paso computacional 13636) se inauguró en Valdorros un gran polígono industrial, dotando de gran afluencia de nuevos vecinos en los años venideros al municipio. Por lo tanto, a partir del paso 13636, Valdorros variará su vector original; de asignarle un 20 de empleo, pasará a tener un 60 sobre 100.

Pueblo	Sanidad	Ocio	Empleo	Educación	Movilidad	Arraigo
Cill. de Abajo	4	5	10	3	3	15
Covarrubias	5	5	10	5	5	12
Lerma	30	15	10	20	5	10
Mecerreyes	4	5	8	1	4	12
Nebreda	5	5	5	1	1	25
Presencio	4	5	10	1	3	23
Quint. del Agua	5	5	5	1	4	5
S. M. del Campo	5	5	8	5	3	5
Tordómar	5	5	10	1	4	25
Valdorros	5	10	20	7	4	20
Villafruela	4	5	10	3	3	15
Villalmanzo	5	10	15	3	4	20
Burgos	55	65	55	65	75	65

Tabla 2: Vectores que definen distintos pueblos en 2001.

	Jóvenes	Adultos	Ancianos
Sanidad	0,5	0,5	0,9
Ocio	0,8	0,5	0,7
Empleo	1,0	1,0	0,5
Educación	0,8	1,0	0,5
Movilidad	0,8	0,6	0,8
Arraigo	0,8	0,8	1,0

Tabla 3: Valores seleccionados para el coeficiente $\alpha_{n,sj}$ en cada caso.

Con todo lo aportado en la Tabla 2 y en la Tabla 3, se ha procedido a simularlo. Su resultado queda reflejado en la figura 8.

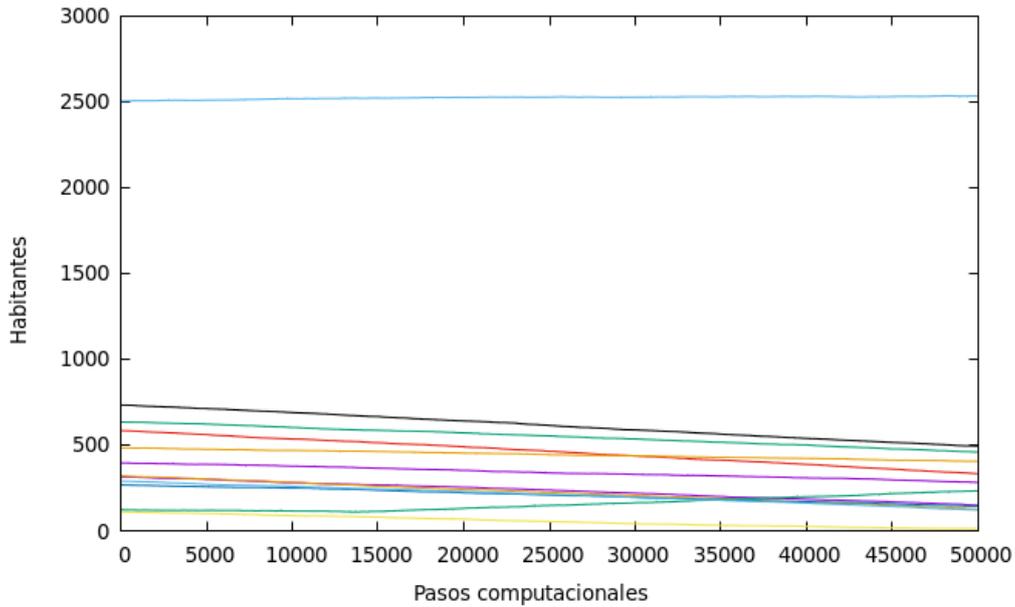


Figura 8: Evolución habitacional de distintos pueblos tras 50000 pasos computacionales.

Al igual que en el caso anterior, sus valores numéricos aparecen en la Tabla 4.

Pueblos	Población (2001)	Población (2022)	Variación (%)
Cilleruelo de Abajo	314	151	-51,8
Covarrubias	633	477	-24,6
Lerma	2502	2541	+1,6
Mecerreyes	320	118	-63,2
Nebreda	110	13	-87,8
Presencio	268	135	-49,7
Quintanilla del Agua	583	334	-42,7
Santa María del Campo	732	511	-30,1
Tordómar	395	288	-27,1
Valdorros	123	242	+96,8
Villafruela	290	125	-57,0
Villalmanzo	483	395	-18,3

Tabla 4: Simulación de los habitantes de distintos pueblos tras 50000 pasos computacionales, con su correspondiente variación.

La efectividad del programa se puede ver en la Tabla 5, donde se comparan las variaciones reales y las obtenidas mediante simulación. Para poder analizarla, hay que definir el concepto de efectividad. A pesar de su ambigüedad, se pueden distinguir dos tipos según el estudio que se quiera llevar a cabo. Por un lado, en el caso de que sólo interesen los datos iniciales y finales, basta con comparar la variación entre los datos simulados y los reales. En cambio, si se quiere analizar la efectividad en distintos periodos entre 2001 y 2022, se ha optado por realizarla calculando la raíz del error cuadrático medio y estudiando los valores donde se minimiza. El análisis de la RECM se hará más adelante.

Pueblos	Variación real	Variación simulada	Var. comparativa
Cill. de Abajo	-33,1	-51,8	18,7
Covarrubias	-14,5	-24,6	10,1
Lerma	+1,4	+1,6	0,2
Mecerreyes	-39,7	-63,2	23,5
Nebreda	-51,8	-87,8	36,0
Presencio	-31,3	-49,7	18,4
Quint. del Agua	-38,2	-42,7	4,4
S. M. del Campo	-29,4	-30,1	0,8
Tordómar	-15,4	-27,1	11,6
Valdorros	141,5	+96,8	44,7
Villafruela	-39,3	-57,0	17,6
Villalmanzo	-13,0	-18,3	5,2
Media			15,9

Tabla 5: Variaciones respecto a los habitantes en 2001 y 2022 en la realidad y en la simulación tras 50000 pasos computacionales.

De los resultados recogidos en la Tabla 5 se puede concluir que la tendencia de la realidad es la misma que la de las simulaciones; los pueblos que pierden ciudadanos también lo hacen en las simulaciones, mientras que los que tienen una variación positiva en la realidad, también la tienen en los datos obtenidos.

Es obvio que Lerma, dada la diferencia del número de habitantes respecto al resto de pueblos y, en consecuencia, en la escala vertical, imposibilita cualquier tipo de análisis visual. Por lo tanto, se prescinde de ella para facilitar el estudio gráfico. En las gráficas 9 y 10 se muestra.

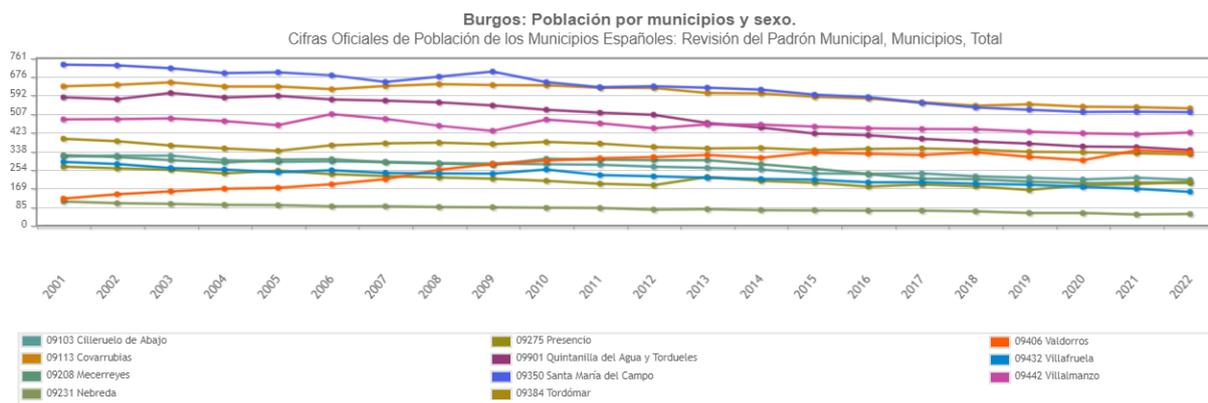


Figura 9: Datos reales de la evolución habitacional de distintos pueblos entre 2001 y 2022.

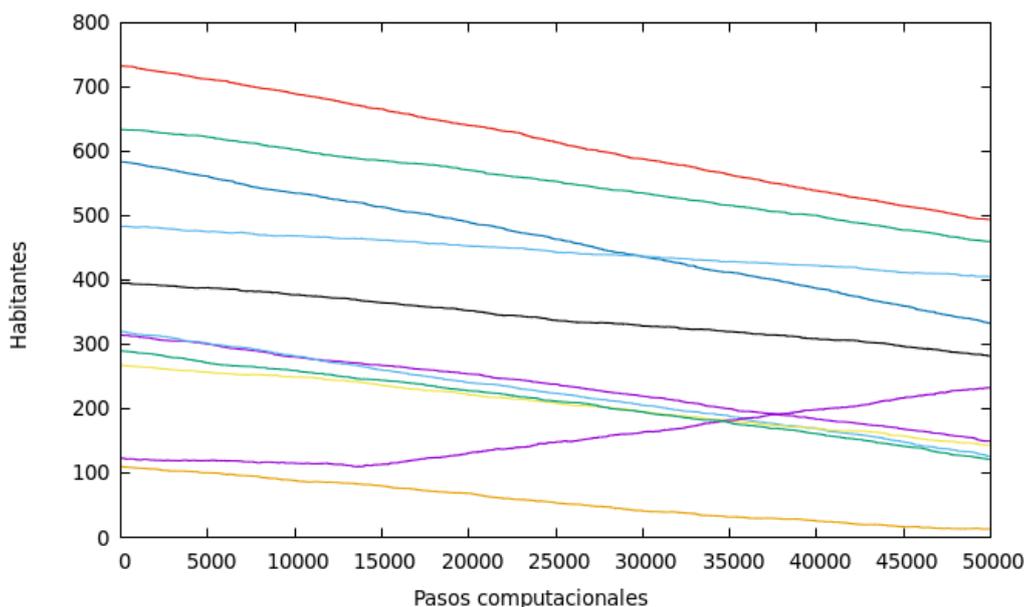


Figura 10: Evolución habitacional de distintos pueblos tras 50000 pasos computacionales.

Otro de los ámbitos que permite estudiar el algoritmo desarrollado es la segregación por franjas de edades. En este caso, por ejemplo, se han seleccionado los pueblos de Lerma y de Valdorros. Los vectores que definen cada uno de los dos pueblos se recogen en la Tabla 6.

Pueblo	Sanidad	Ocio	Empleo	Educación	Movilidad	Arraigo
Lerma	30	15	10	20	5	10
Valdorros	5	10	20	7	5	20

Tabla 6: Vectores que definen Lerma y Valdorros en el 2001.

A simple vista, estos datos pueden llevar a ciertas conclusiones. Por un lado, Lerma tiene sus mayores atractivos en la sanidad y en la educación, por lo que sería esperable que ancianos y adultos se sintieran más atraídos a este pueblo. Por su parte, el mayor incentivo de Valdorros es su oferta de empleo; esto hará que ciudadanos de corta y mediana edad se decidan por mudarse al municipio valdorrense.

Como datos iniciales también se ha supuesto que existen en cada pueblo el mismo número de jóvenes, adultos y ancianos, es decir, un tercio de sus respectivos habitantes totales. Esto es, en Lerma, con sus 2502 habitantes en 2001, se presumen 834 jóvenes, adultos y ancianos; Valdorros, con 123 ciudadanos, se ha supuesto que partía de un total de 41 miembros de cada franja. En la figura 11 y en la figura 12 se puede ver la evolución de población de cada pueblo con sus respectivos bloques generacionales.

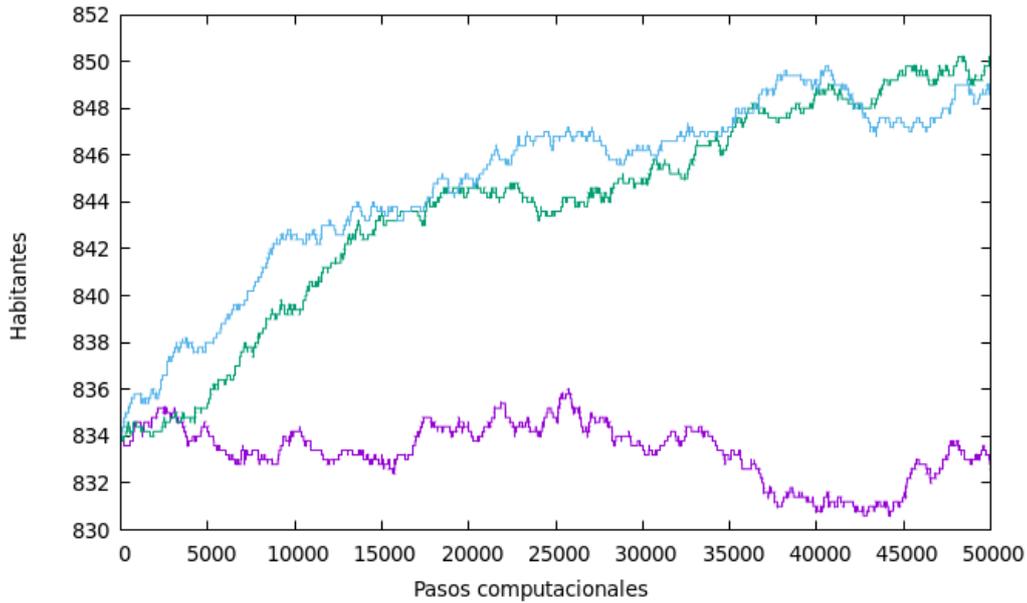


Figura 11: Evolución habitacional de los jóvenes (morado), adultos (verde) y ancianos (azul) en Lerma tras 50000 pasos computacionales.

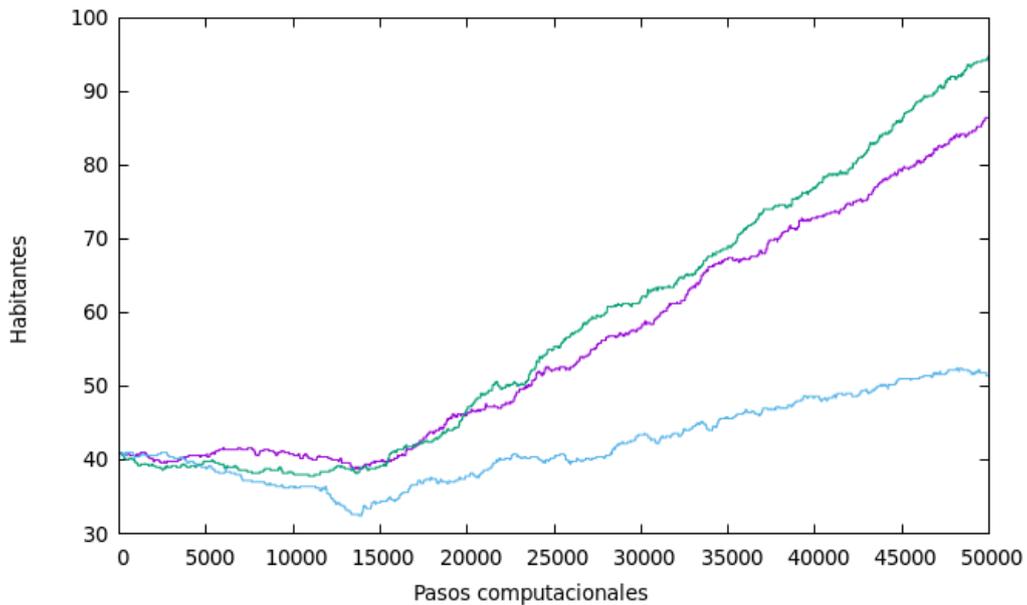


Figura 12: Evolución habitacional de los jóvenes (morado), adultos (verde) y ancianos (azul) en Valdorros tras 50000 pasos computacionales.

Como se ha comentado con anterioridad, en Lerma, la subida en el número de adultos y ancianos que se muestra en la gráfica 11 es consecuencia de la buena sanidad y educación de la localidad, así que se pueden considerar unos resultados satisfactorios. Los de Valdorros, por otro lado, también. Eran de esperar los resultados obtenidos en la gráfica 12, pues la amplia oferta de trabajo de Valdorros tiene un gran atractivo para jóvenes y adultos.

Al igual que en el resto de casos, para un análisis más sencillo y un mejor estudio de la efectividad del programa, los datos de la figura 11 y la figura 12 se han recogido en la Tabla 7.

Lerma						Valdorros					
Jóvenes		Adultos		Ancianos		Jóvenes		Adultos		Ancianos	
2001	2022	2001	2022	2001	2022	2001	2022	2001	2022	2001	2022
834	838	834	848	834	855	41	92	41	92	41	58

Tabla 7: Habitantes de los jóvenes, adultos y ancianos en Lerma y Valdorros tras 50000 pasos computacionales, el equivalente a 2001 y 2022.

Por lo tanto, los resultados de la Tabla 7 son fácilmente interpretables: a los pueblos con mejor sanidad acuden los ciudadanos de edad más avanzada y evitan los lugares sin ella; si tienen una amplia oferta educacional, será atractivo para las personas de mediana edad porque será más fácil educar a sus hijos en el propio pueblo sin necesidad de transporte; finalmente, tanto los jóvenes como los adultos se verán atraídos por municipios que tengan una gran oferta laboral.

4.2. Proyecciones

Otra de las utilidades del algoritmo creado consiste en proyectar la evolución habitacional en el futuro; basta con aumentar el número de pasos simulados. Su objetivo principal es tener una estimación del número de ciudadanos que habrá en un pueblo o región en un tiempo futuro. Para poder realizar el símil con la realidad, hay que recordar que 2381 pasos computacionales equivalen a 1 año. Se han simulado 69049 pasos, con el fin de conocer los habitantes de distintos pueblos de la Comarca del Arlanza en el año 2030.

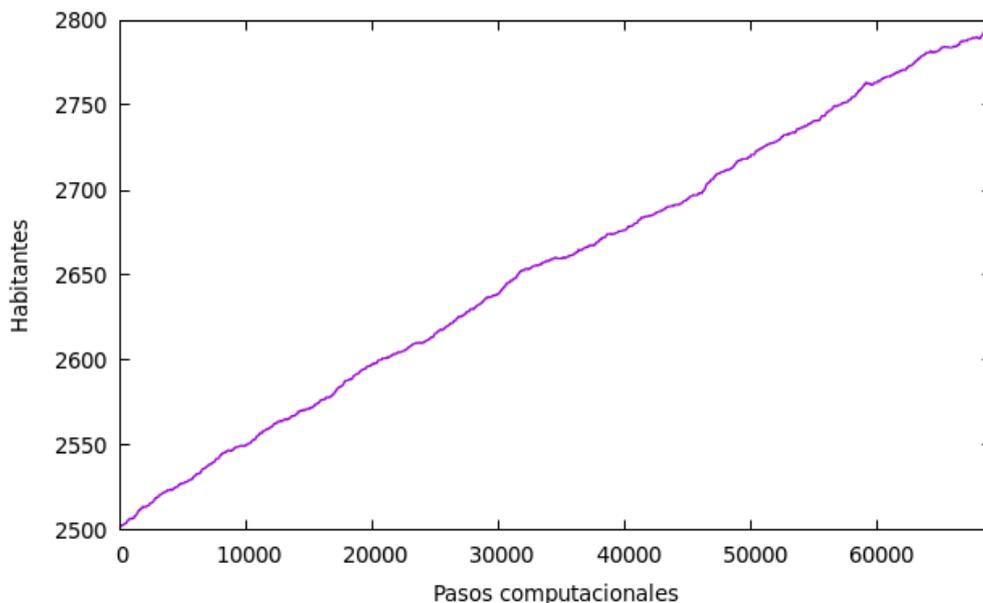


Figura 13: Proyección de Lerma desde 2001 hasta 2030 (69049 pasos computacionales).

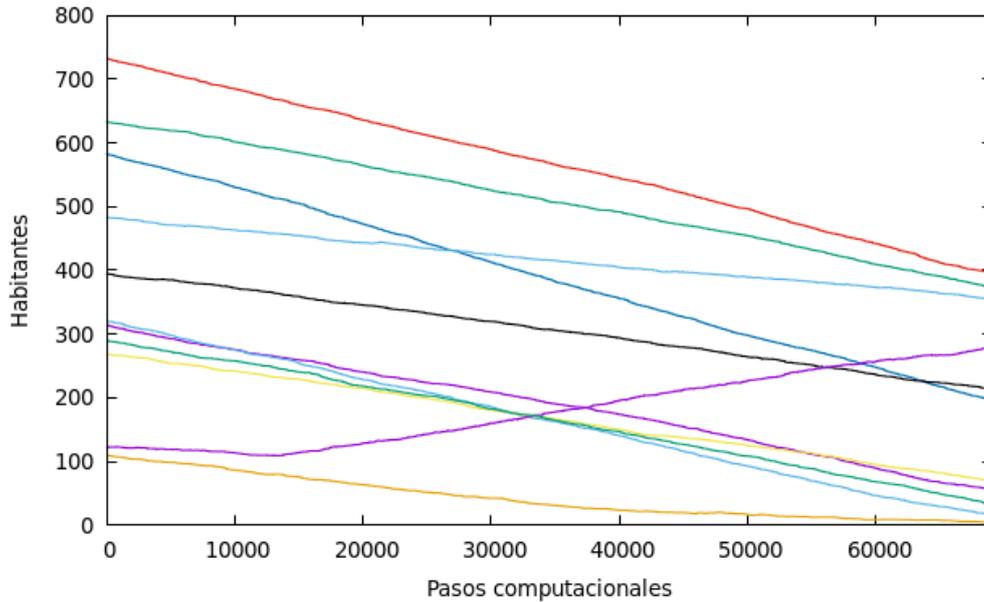


Figura 14: Proyección de varios pueblos desde 2001 hasta 2030 (69049 pasos).

Al igual que ha sucedido en el resto del trabajo, dada la diferencia en el eje vertical entre los habitantes de Lerma y el resto de pueblos, se ha considerado dividir la gráfica de todos los pueblos analizados en dos: la proyección de Lerma (gráfica 13) y la proyección del resto de pueblos (gráfica 14). Se han recogido estos datos en la Tabla 8.

A simple vista, examinar estos datos es complicado, porque escasean estudios exhaustivos de proyecciones de habitantes en los pueblos que conforman la denominada España vaciada. Aún así, se pueden prever las tendencias habitacionales de la Comarca del Arlanza mediante otras informaciones más generales. Por ejemplo, según portales estadísticos [25], la provincia de Burgos tendrá el 90% de los 178864 habitantes que tenía en 2011, es decir, 160977. Además, en una noticia del *Diario de León* [26], basado en un estudio del Instituto Nacional de Estadística, coloca a Castilla y León como la tercera comunidad autónoma que más habitantes habrá perdido entre 2014 y 2029, con una pérdida del 9%.

	2001	2030	Variación (%)
Cilleruelo de Abajo	314	56	-82,1
Covarrubias	633	373	-41,1
Lerma	2502	2797	+11,8
Mecerreyes	320	17	-94,6
Nebreda	110	6	-94,5
Presencio	268	70	-73,9
Quintanilla del Agua	583	196	-66,3
Santa María del Campo	732	396	-46,0
Tordómar	395	214	-45,8
Valdorros	123	278	+125,8
Villafruela	290	33	-88,6
Villalmanzo	483	355	-26,5

Tabla 8: Proyecciones de distintos pueblos tras 69049 pasos, el equivalente al año 2030.

Teniendo en cuenta estos datos, y viendo que la tendencia que toman la mayoría de los pueblos es la pérdida de población (gráficas 13 y 14), parece que los resultados obtenidos en las simulaciones son semejantes a lo que se le presupone al futuro. A pesar de su tendencia realista, parece que la pérdida de habitantes de las simulaciones es excesiva. De todas maneras, cabe recordar que los datos con los que se han hecho las comparaciones tampoco son exactos, sino aproximaciones que se han realizado teniendo en cuenta lo que sucede alrededor de la Comarca del Arlanza; lo ideal sería tener los estudios sociológicos acerca de las proyecciones de cada uno de los pueblos con los incentivos reales de cada uno de ellos, con la mejoría o empeoramiento de éstos a lo largo de los años.

4.3. Factor de sensibilidad

Una de las tareas más complejas a la hora de ejecutar el algoritmo es la definición de los vectores característicos de cada uno de los pueblos. Muchos de estos datos se pueden aproximar a la realidad mediante el uso de la intuición, aunque no dejarían de tener un grado de inexactitud. Para intentar atenuarlo, se ha calculado la raíz del error cuadrático medio (RECM) de una variable con distintos valores; entre todos ellos, el que minimice ésta será el valor más probable de la variable.

En la Tabla 5 se ve la variación entre los datos reales y la simulación, por lo que se puede tener una primera aproximación de la efectividad. Además, ésta también se puede abordar desde otra perspectiva: mediante el cálculo de la raíz del error cuadrático medio en varios puntos entre el año 2001 y 2022, por lo que no sólo servirá para estudiar la exactitud de las variaciones iniciales y finales sino para la evolución completa. En este caso, se estudiará la evolución en 5 puntos a lo largo del recorrido: 2001, 2005, 2010, 2015 y 2022.

4.3.1. Lerma

El primer pueblo a analizar será Lerma. Para satisfacer la ecuación 4, es necesario conseguir los habitantes reales en los años elegidos (2001, 2005, 2010, 2015 y 2022) y los correspondientes a las simulaciones en su análogo de pasos computacionales.

Variable	2001	2005	2010	2015	2022
Real	2502	2744	2798	2703	2584
0	2502	2503	2502	2511	2518
10	2502	2520	2531	2541	2560
20	2502	2541	2560	2579	2610
30	2502	2549	2584	2623	2666
40	2502	2558	2592	2635	2691
50	2502	2573	2620	2665	2730
60	2502	2593	2648	2702	2782
70	2502	2598	2667	2736	2830
80	2502	2618	2687	2765	2870
90	2502	2641	2717	2800	2914

Tabla 9: Habitantes en Lerma variando la movilidad en distintos años reales y virtuales.

En este caso, se estudiará la variable de *movilidad y transporte*, a la que, en la Tabla 2, se le ha dotado con un 5 sobre 100, pues Lerma tiene una estación de autobuses, algo que, a pesar de su poca afluencia, no es habitual en los pueblos de la comarca. Esta variable tomará valores entre 0 y 90, en saltos de 10 en 10. Esta información se plasma en la Tabla 9.

Con estos datos, se puede calcular la raíz del error cuadrático medio para cada uno de los valores de la variable de movilidad. En la Tabla 10, se recogen las RECM con distintos valores para la variable de movilidad, mientras que su representación gráfica se muestra en la figura 15.

Variable	0	10	20	30	40	50	60	70	80	90
RECM	132,9	106,2	93,1	67,5	44,0	37,0	42,4	58,6	68,8	85,4

Tabla 10: Distintos valores de movilidad con sus correspondientes errores cuadráticos medios en Lerma.

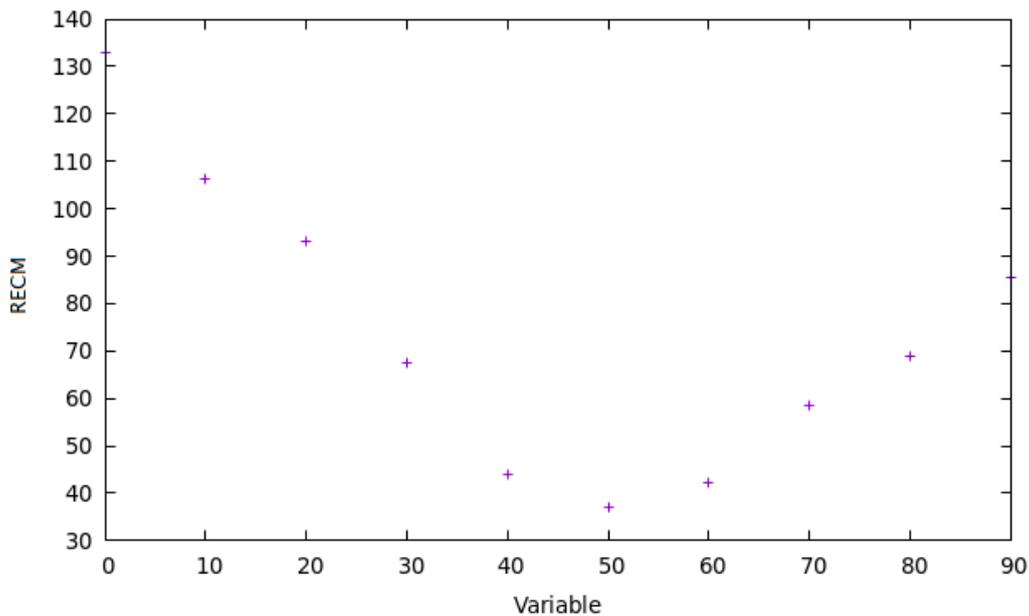


Figura 15: Distintos valores de movilidad con sus correspondientes RECM en Lerma.

Tanto de la Tabla 10 como de la gráfica 15 se pueden discutir varios puntos. Por un lado, que el valor en la variable de *movilidad y transporte* que minimiza la raíz del error cuadrático medio es 50, con un error de 37,0. Esto significa que el valor asignado a la movilidad en primera instancia es bajo. También se percibe en la gráfica 15 uno de los principios del error cuadrático medio y su raíz: el error aumenta considerablemente con valores alejados del real, mientras que en puntos cercanos al vértice, apenas hay variación. En este caso, es notable que el error aumenta a medida que nos alejamos de 50, valor que minimiza el error.

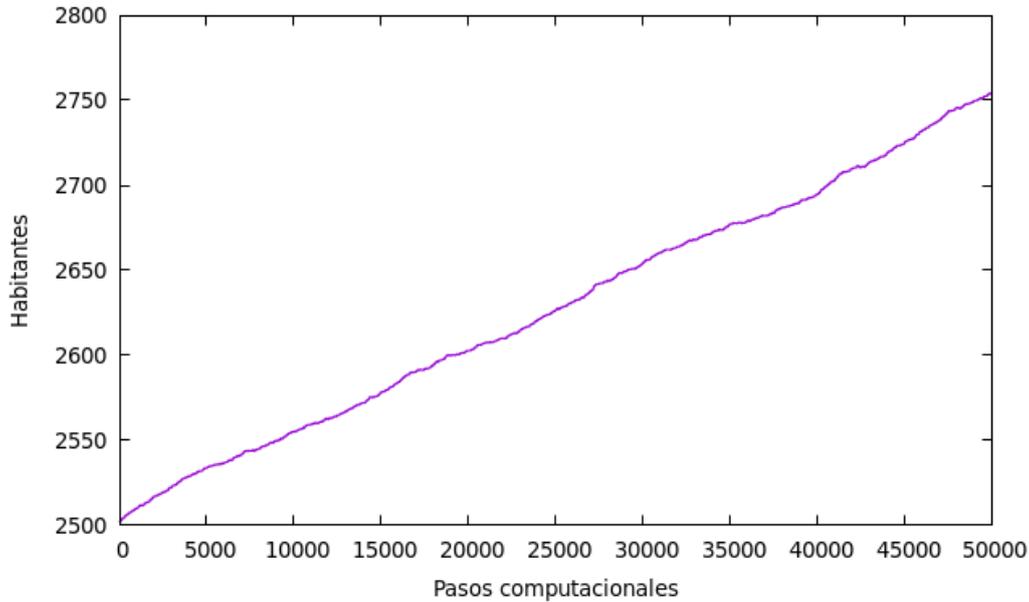


Figura 16: Evolución habitacional de Lerma con la variable de movilidad = 50 tras 50000 pasos computacionales.

Aunque es trivial ver que el número de habitantes ha aumentado comparándolo con la simulación realizada con 5 sobre 100 como valor de la variable (Tabla 4), la gráfica 16 se ciñe más a los datos reales a lo largo de los años. Esto se debe a que, a pesar del alejamiento del valor real de población en el año 2022, se han tenido en cuenta varios puntos a lo largo de la evolución completa. Así, se adapta mejor a los valores reales en todo el recorrido evolutivo, obteniendo más precisión total, en lugar de tener en cuenta únicamente el punto final.

4.3.2. Valdorros

Debido a la particularidad que tiene Valdorros al aumentar considerablemente su población a partir del año 2007 (paso computacional 13636), también se ha querido estudiar la RECM para identificar el valor óptimo de la variable asociada al *ocio*. Al igual que para analizar la variable de movilidad en Lerma, también requerimos de los datos simulados y datos reales de población en los años 2001, 2005, 2010, 2015 y 2022. Éstos se conseguirán corriendo la variable de *ocio* de 0 a 90, de 10 en 10. Los datos reales se obtendrán del Instituto Nacional de Estadística [27].

Mediante la ecuación 4 y los datos de la Tabla 11, se ha calculado la raíz del error cuadrático medio para los valores de la variable de *ocio* comprendidos entre 0 y 90. Estos valores y sus correspondientes RECM se han recogido en la Tabla 12.

Variable	2001	2005	2010	2015	2022
Real	123	172	295	332	333
0	123	108	136	154	183
10	123	115	148	175	223
20	123	133	178	221	278
30	123	144	190	235	300
40	123	158	210	264	348
50	123	164	224	287	378
60	123	176	249	315	411
70	123	188	264	357	471
80	123	192	277	351	489
90	123	208	299	389	523

Tabla 11: Habitantes en Valdorros con distintos valores para la variable de ocio en varios años reales y virtuales.

Variable	0	10	20	30	40	50	60	70	80	90
RECM	129,4	111,0	78,0	66,8	49,5	42,8	41,5	64,8	71,5	90,1

Tabla 12: Distintos valores de ocio con sus correspondientes errores cuadráticos medios en Valdorros.

La representación gráfica de los datos de la Tabla 12 se muestra en la figura 17.

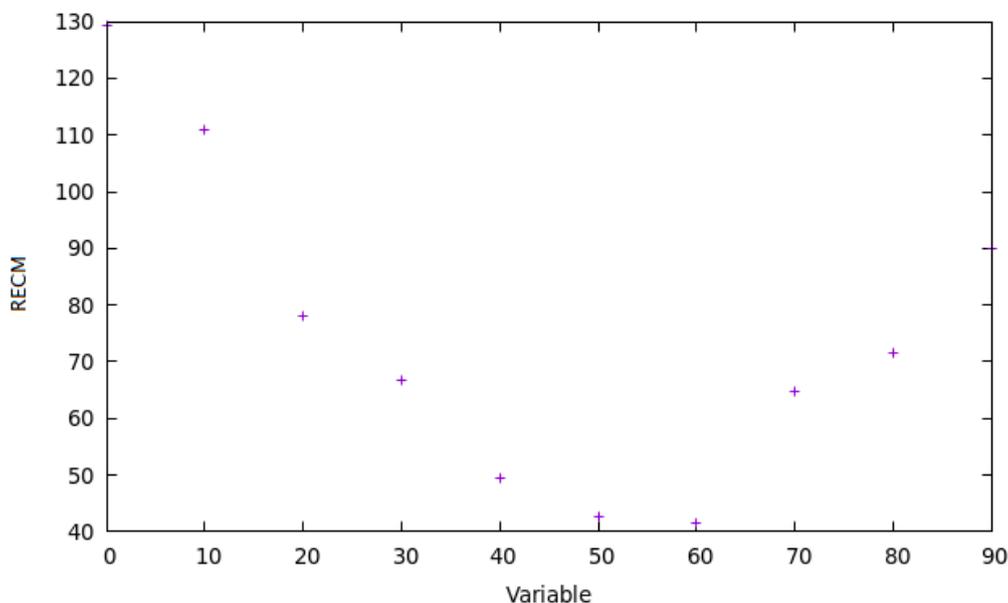


Figura 17: Distintos valores de ocio con sus correspondientes RECM en Valdorros.

Si se compara con el resultado teórico del error cuadrático medio (gráfica 5), se puede considerar que los resultados obtenidos se ajustan con éxito a lo esperado. Para el caso de Valdorros, la curva se centra en 60, que es el valor que minimiza el error.

Por último, se ha vuelto a simular el caso de Valdorros, cambiando la variable de *ocio* a 60 sobre 100 entre los años 2001 y 2022, el equivalente a 50000 pasos computacionales. Su resultado se muestra en la figura 18.

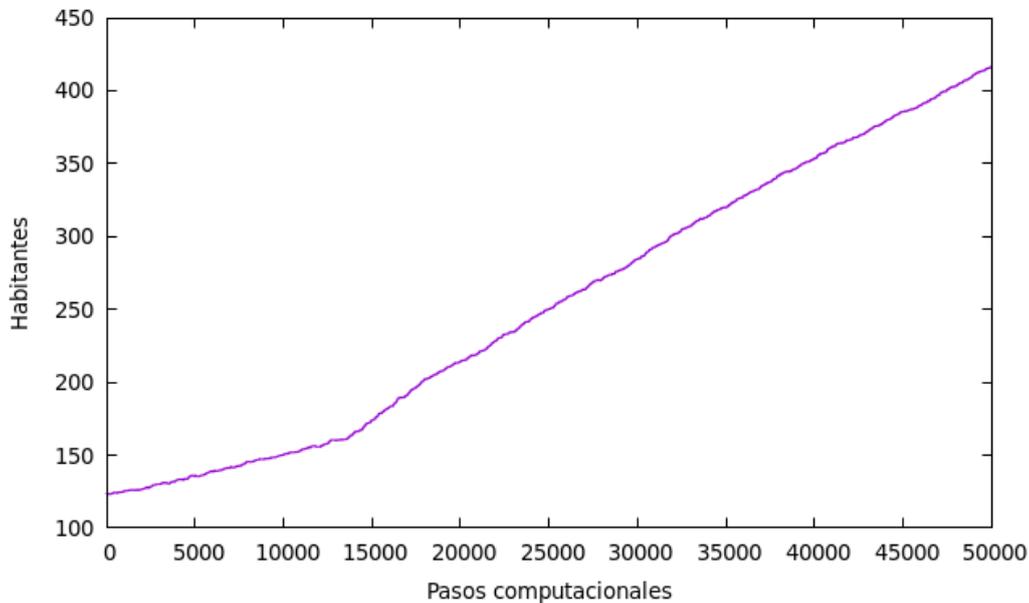


Figura 18: Evolución habitacional de Valdorros con la variable de ocio = 60 tras 50000 pasos computacionales.

Al comparar el resultado de esta simulación con el de la gráfica 9, se da el mismo efecto que en el análisis de Lerma: dista más del valor final real, pero se acopla mejor a la realidad en el resto del recorrido temporal.

Tras analizar los casos de Lerma y de Valdorros, se puede considerar que los resultados obtenidos mediante la minimización de la raíz del error cuadrático medio son satisfactorios. Su objetivo principal era que los datos obtenidos mediante las simulaciones se ciñeran a los habitantes reales a lo largo de los distintos pasos computacionales, y, como se muestra en la figura 16 y en la figura 18, al compararlas con las gráficas 7 y 9, se ha cumplido. También se observa que se han obtenido resultados satisfactorios en otro de los fines del trabajo: conseguir los valores de los parámetros del sistema (*movilidad y transporte* en Lerma y *ocio* en Valdorros) que producen mejores resultados que los escogidos inicialmente.

5. Conclusiones

Al ser un campo de estudio novedoso, son varias las conclusiones que se pueden observar. Para un análisis óptimo, conviene recordar los objetivos de este trabajo.

El fin principal de este trabajo era estudiar mediante herramientas físicas y matemáticas los fenómenos migratorios, en concreto, las dinámicas poblacionales de la Comarca del Arlanza entre los años 2001 y 2022, además de predecir su evolución habitacional en los años venideros. También se han querido proponer metodologías para cuantificar la realidad de un modo más exacto.

En cuanto a las simulaciones realizadas para estudiar las evoluciones de los distintos pueblos entre 2001 y 2022, se puede considerar que los resultados obtenidos son satisfactorios. Al comparar con los datos proporcionados por el Instituto Nacional de Estadística [24], se aprecian tendencias muy parecidas a las de las simulaciones; los pueblos con mayores incentivos aumentan el número de habitantes, mientras que los que tienen pocos atractivos muestran un decrecimiento poblacional. Aún así, también hay que tener en cuenta que algunos resultados difieren considerablemente de los datos reales. Esto puede obedecer a una inexactitud al introducir los vectores definitorios de los pueblos. Otro punto interesante que se puede examinar de estas simulaciones es que las variaciones más grandes vienen de los pueblos más pequeños; esto se debe a que la variación del número de ciudadanos no afecta por igual a núcleos pequeños de población que a lugares más poblados. El caso de Valdorros ilustra, además, que el carácter dinámico de la configuración de un pueblo obligaría a tener en cuenta los cambios con el tiempo de todos los factores en todos los pueblos.

Deducir conclusiones al analizar las proyecciones de habitantes en el año 2030 es más complicado, debido a los pocos estudios que se han realizado en este aspecto. No obstante, hay predicciones a nivel autonómico y provincial que auguran una pérdida de habitantes en toda la región. Suponiendo que la Comarca del Arlanza también tiene una tendencia parecida, puede considerarse que los resultados logrados se ajustan con éxito a la realidad. De todas maneras, hay pueblos que tienen una pendiente más negativa de lo que, en primera instancia, se les presupondría según las fuentes que se han consultado.

En relación al método empleado para cuantificar la realidad de un modo óptimo, cabe mencionar que éste es uno de los puntos claves de la sociofísica, y, en consecuencia, del programa desarrollado. Se ha conseguido que el algoritmo minimice el valor de alguna de las variables, obteniendo, de esta manera, resultados más realistas. Sin embargo, analizando los datos obtenidos, parece que mejorarían su calidad si la raíz del error cuadrático medio se calculara en intervalos de tiempo más cortos para poder adaptarse mejor a las variaciones repentinas de habitantes.

A modo de conclusión general, se puede afirmar que los algoritmos desarrollados han tenido un comportamiento adecuado. De momento, sus limitaciones son evidentes, sobre todo, a la hora de cuantificar la realidad; éste se puede considerar el mayor reto de la sociofísica en los próximos años, confiando en que este trabajo pueda añadir nuevas perspectivas y complementar las técnicas ya existentes.

Referencias

- [1] Gheorghe G Săvoiu. *Sociophysics: A New Science or a New Domain for Physicists in a Modern University?* URL: https://www.academia.edu/1346408/Sociophysics_A_New_Science_or_a_New_Domain_for_Physicists_in_a_Modern_University.
- [2] Stewart Duncan. “Thomas Hobbes”. En: *The Stanford Encyclopedia of Philosophy*. Ed. por Edward N. Zalta. Summer 2022. Metaphysics Research Lab, Stanford University, 2022.
- [3] Georg G. Iggers. “Further Remarks about Early Uses of the Term ”Social Science””. En: *Journal of the History of Ideas* 20.3 (1959), págs. 433-436. ISSN: 00225037, 10863222. URL: <http://www.jstor.org/stable/2708121>.
- [4] *Sociophysics - The last science*. URL: <https://polyhistornaut.wordpress.com/2016/03/19/sociophysics-the-last-science/>.
- [5] *Social physics*. URL: <https://www.eoht.info/page/Social%20physics>.
- [6] Stephen G. Brush. *History of the Lenz-Ising*. October 1967. URL: <https://www.apps.stat.vt.edu/leman/VTCourses/ReviewIsing.pdf>.
- [7] Informs. *Renfrey B. Potts*. URL: <https://www.informs.org/Explore/History-of-0.R.-Excellence/Biographical-Profiles/Potts-Renfrey-B>.
- [8] Laura Beaudin. *A Review of the Potts Model*. 2007. URL: <https://scholar.rose-hulman.edu/cgi/viewcontent.cgi?article=1194&context=rhumj>.
- [9] Dietrich Stauffer. “Opinion Dynamics and Sociophysics”. En: *Encyclopedia of Complexity and Systems Science*. Ed. por Robert A. Meyers. New York, NY: Springer New York, 2009, págs. 6380-6388. ISBN: 978-0-387-30440-3. DOI: 10.1007/978-0-387-30440-3_376. URL: https://doi.org/10.1007/978-0-387-30440-3_376.
- [10] Javier Falcó. *Simulación del Modelo Depredador-Presa de Lotka-Volterra*. URL: <https://www.uv.es/falbe/MatExp/aplicada/modelizacion/Lotka-Volterra/>.
- [11] Yuri Tyutyunov y Lyudmila Titova. “From Lotka-Volterra to Arditi-Ginzburg: 90 Years of Evolving Trophic Functions // Biology Bulletin Reviews”. En: 10 (mayo de 2020), págs. 167-185. DOI: 10.1134/S207908642003007X.
- [12] C. Gracia-Lázaro et al. “Residential segregation and cultural dissemination: An Axelrod-Schelling model”. En: *Phys. Rev. E* 80 (4 oct. de 2009), pág. 046123. DOI: 10.1103/PhysRevE.80.046123. URL: <https://link.aps.org/doi/10.1103/PhysRevE.80.046123>.
- [13] Journal of Artificial Societies y Social Simulation. *Modelo de Axelrod*. URL: <https://www.jasss.org/12/1/6/appendixB/Axelrod1997.html>.
- [14] Andreas Flache. *Between Monoculture and Cultural Polarization: Agent-based Models of the Interplay of Social Influence and Cultural Diversity*. 2018. URL: <https://doi.org/10.1007/s10816-018-9391-1>.
- [15] Jorshua Aráuz Cantarero. *Cadenas de Márkov*. URL: <https://jorshua.wordpress.com/2012/06/11/cadenas-de-markov/>.
- [16] J. Torres. *Procesos de Márkov*. URL: <https://www.ugr.es/~jtorres/leccion3.pdf>.

- [17] Eugene Seneta. “Markov Chains as Models in Statistical Mechanics”. En: *Statistical Science* 31.3 (2016), págs. 399-414. DOI: 10.1214/16-STS568. URL: <https://doi.org/10.1214/16-STS568>.
- [18] Tornike Gongadze Gogvadze. *Diseño y Simulación de un modelo predictivo para la velocidad y dirección del viento*. URL: <https://upcommons.upc.edu/bitstream/handle/2117/99138/TFG-Tornike-Gongadze.pdf?sequence=1&isAllowed=y>.
- [19] Ricardo Ocaña-Riola. “Modelos de Markov aplicados a la investigación en Ciencias de la Salud”. es. En: *Interciencia* 34 (mar. de 2009), págs. 157-162. ISSN: 0378-1844. URL: http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0378-18442009000300004&nrm=iso.
- [20] Luisa R. y Eliana Toro Ocampo Yuleidy Ordóñez. *Aplicación de cadenas de Márkov continuas a las estadísticas del secuestro en Colombia*.
- [21] Adriana Bellorin. *Cadenas de Márkov*. URL: <https://isibellorin.files.wordpress.com/2012/06/cadenas-de-markov2.pdf>.
- [22] Dilane. *Presentation of MSE (Mean Squared Error)*. URL: <https://medium.com/@DilaneKombou/presentation-of-mse-mean-squared-error-a4ee9b6cff49>.
- [23] Oficina de Turismo de Lerma. URL: <https://www.citlerma.com/antiguo/Images/mapaarlanza.gif>.
- [24] Instituto Nacional de Estadística. *Burgos: población por municipios y sexo*. URL: <https://www.ine.es/jaxiT3/Tabla.htm?t=2862&L=0>.
- [25] Microanálisis Inmobiliario. *Proyecciones de población por provincias durante la próxima década*. URL: <https://microanalisisinmobiliario.com/proyecciones-de-poblacion/>.
- [26] Diario de León. *El INE calcula que Castilla y León perderá 224.000 habitantes hasta 2029*. URL: <https://www.diariodeleon.es/articulo/castillayleon/ine-calcula-castilla-leon-perderra-224000-habitantes-2029/201410290500011470961.html>.
- [27] Instituto Nacional de Estadística. *Valdorros: población por municipios y sexo*. URL: <https://www.ine.es/jaxiT3/Datos.htm?t=2862>.