

# Consenso y polarización en la dinámica espacial de opiniones bajo confianza acotada

Paula Rosa Rodríguez Morales

## Abstract

La formación de opiniones colectivas en sistemas sociales puede dar lugar tanto a un consenso generalizado como a dinámicas de polarización. Aunque estos fenómenos han sido ampliamente estudiados, a día de hoy sigue sin estar claro por qué reglas de interacción aparentemente similares, pueden producir resultados colectivos tan diferentes.

En este paper se presenta un modelo basado en agentes que combina la dinámica de opiniones bajo confianza acotada con mecanismos de relocalización espacial inspirados en el modelo de segregación de Schelling. Mediante simulación se analiza como los agentes mantienen opiniones continuas e interactúan de manera local, actualizando sus creencias únicamente cuando las diferencias de opinión se sitúan por debajo de un determinado umbral de tolerancia. Asimismo, cuando el entorno social inmediato resulta incompatible, los agentes pueden decidir reubicarse en el espacio.

A partir de experimentos de simulación sistemáticos, se observa que las distribuciones iniciales unimodales de opinión tienden a converger hacia un consenso global bajo niveles de tolerancia moderados o elevados. Por el contrario, las condiciones iniciales bimodales en combinación con bajos niveles de tolerancia, dan lugar a estados de polarización persistente acompañados de segregación espacial. Estos resultados en conjunto integran en un único marco computacional, teorías clásicas de influencia social y segregación.

## 1 Introducción

La dinámica de formación de opiniones en sistemas sociales ha sido objeto de amplio interés en sociología, economía, física y ciencia social computacional [1, 4]. Los fenómenos colectivos como el consenso y la polarización influyen de manera decisiva en el comportamiento político, la evolución cultural y, por consiguiente y de forma general, en los procesos de toma de decisiones colectivas [8].

Los primeros modelos formales de dinámica de opiniones se centraron principalmente en la convergencia hacia el consenso, entendiéndolo como un resultado natural de la influencia social repetida. El modelo clásico de DeGroot, mostró que bajo condiciones generales de conectividad, el promedio iterativo de las opiniones conduce a un consenso casi universal [3]. Sin embargo, a pesar de su calidad analítica, estos modelos no explican la persistencia del desacuerdo y la polarización que se observa en sociedades reales.

Para superar esta limitación, se desarrollaron los modelos de confianza acotada, aquellos en los que los agentes solo interactúan con otros cuyas opiniones son suficientemente similares. Estos modelos muestran que la fragmentación y la polarización pueden surgir de manera endógena, incluso sin que exista un conflicto previo entre los agentes [2, 6]. No obstante, la mayoría de estos enfoques no incorpora de forma explícita la estructura espacial.

Por otro lado, la segregación espacial ha sido ampliamente estudiada a partir del modelo de Schelling. Este modelo mostró que incluso preferencias locales débiles pueden generar patrones marcados de segregación a nivel agregado [9], pero, a pesar de su cercanía conceptual, la dinámica de opiniones y la segregación espacial se han abordado en gran medida como problemas independientes.

En este trabajo, se combinan ambas tradiciones de modelado para analizar cómo la influencia social y la relocalización espacial interactúan y determinan los resultados colectivos.

## 2 Desafío

Un desafío en el estudio de la dinámica de opiniones es entender por qué algunas sociedades a veces convergen hacia opiniones compartidas, mientras que bajo condiciones aparentemente similares permanecen persistentemente polarizadas. Este problema se apoya en tres regularidades empíricas.

En primer lugar, el consenso tiende a emerger cuando las opiniones iniciales son relativamente cercanas entre sí. En segundo lugar, la polarización a menudo persiste incluso en contextos donde los individuos están dispuestos a revisar sus puntos de vista. En tercer lugar, la organización espacial y social tiende a reforzar y estabilizar las divisiones ideológicas.

Los modelos clásicos de promediado explican adecuadamente el primer patrón, pero no los otros dos. Los modelos de confianza acotada ayudan a explicar el desacuerdo persistente, aunque a menudo omiten la estructura espacial explícita. Los modelos de segregación, en cambio, reproducen con éxito el agrupamiento espacial, pero suelen asumir identidades de grupo fijas. Por ello, para explicar de manera endógena tanto el consenso como la polarización es necesario un enfoque de modelado integrado.

## 3 Estado del arte

Los modelos de confianza acotada han sido ampliamente estudiados mediante enfoques analíticos y computacionales. Una contribución fundamental fue realizada por Deffuant et al. [2], quienes introdujeron un modelo estocástico de interacción por pares y demostraron que la fragmentación de opiniones depende críticamente de los niveles de tolerancia de los agentes. Poco después, Hegselmann y Krause [6] propusieron una formulación determinista multiagente que exhibe dinámicas cualitativas similares.

Investigaciones posteriores ampliaron este marco en varias direcciones, incluyendo la introducción de límites de confianza heterogéneos [7], estructuras de interacción basadas en redes [5] y espacios de opinión multidimensionales [4]. A través de estas variantes, la tolerancia emerge de manera consistente como un parámetro clave que gobierna los resultados colectivos.

En paralelo, el modelo de segregación de Schelling ha sido generalizado a entornos espaciales continuos y múltiples grupos, reforzando la idea de que la segregación puede surgir incluso a partir de preferencias individuales relativamente débiles [9]. Contribuciones más recientes enfatizan la importancia de considerar conjuntamente la influencia social y el ordenamiento espacial para explicar la emergencia y persistencia de la polarización [4].

## 4 Brechas de investigación

A pesar de la amplia investigación existente, sigue existiendo una falta de modelos integrados mínimos que combinen simultáneamente dinámica continua de opiniones, interacción bajo confianza acotada, estructura espacial explícita y relocalización socialmente motivada. Muchos enfoques disponibles o bien abstraen completamente la estructura espacial o dependen de movimiento aleatorio y ruido, lo que limita su interpretabilidad.

En este trabajo abordamos esta brecha introduciendo un modelo parsimonioso basado en agentes que acopla la actualización de opiniones con confianza acotada y la relocalización espacial al estilo de Schelling. El modelo alcanza estados estacionarios estables mediante mecanismos puramente endógenos, sin recurrir a ruido externo ni a supuestos ad hoc.

## 5 Experimentación

### 5.1 Diseño del modelo

Consideramos una población de  $N$  agentes ubicados en una red cuadrada bidimensional con condiciones de contorno periódicas, donde cada sitio de la red puede estar ocupado por como máximo un agente. Cada agente se caracteriza por una opinión continua  $o_i \in [0, 1]$ .

El tiempo evoluciona en pasos discretos según un *esquema de actualización asíncrona*. En cada paso, los agentes son seleccionados uno a uno en orden aleatorio. Para un agente dado, se llevan a cabo secuencialmente dos procesos: primero, la actualización de la opinión mediante interacción social local, y segundo, una evaluación de la satisfacción espacial que puede conducir a una relocalización.

Este orden refleja el supuesto de que los agentes revisan inicialmente sus opiniones en respuesta a la influencia social local y posteriormente evalúan si su entorno espacial sigue siendo aceptable.

## 5.2 Dinámica de opiniones

La actualización de opiniones sigue el mecanismo de confianza acotada introducido por Deffuant et al. [2]. Cuando se selecciona un agente  $i$ , este elige aleatoriamente un único compañero de interacción  $j$  entre sus vecinos espaciales dentro de un radio de interacción fijo. La interacción tiene lugar únicamente si

$$|o_i(t) - o_j(t)| \leq \varepsilon, \quad (1)$$

donde  $\varepsilon$  denota el umbral de tolerancia. Si esta condición se cumple, el agente  $i$  actualiza su opinión según

$$o_i(t+1) = o_i(t) + \mu(o_j(t) - o_i(t)), \quad (2)$$

con  $\mu \in (0,1)$  representando la tasa de aprendizaje. Si no hay vecinos compatibles disponibles, la opinión del agente permanece sin cambios durante ese paso temporal.

## 5.3 Relocalización espacial

Sea  $\mathcal{N}_i$  el conjunto de vecinos del agente  $i$  dentro del radio de interacción. Definimos la fracción de vecinos socialmente compatibles como

$$f_i = \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \mathbf{1}(|o_i - o_j| \leq \varepsilon), \quad (3)$$

donde  $\mathbf{1}(\cdot)$  es la función indicadora. Un agente  $i$  se considera insatisfecho cuando  $f_i < \theta$ , siendo  $\theta$  el umbral de satisfacción. Los agentes insatisfechos se reubican en un sitio vacío seleccionado aleatoriamente en la red, mientras que los agentes satisfechos permanecen en su ubicación actual.

## 5.4 Protocolo experimental y parámetros

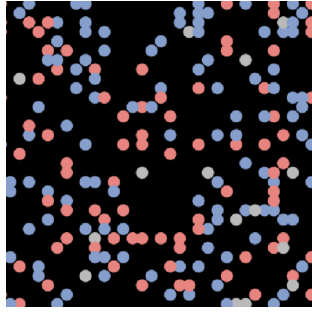
Consideramos distribuciones iniciales de opinión unimodales y bimodales. Salvo que se indique lo contrario, las simulaciones se realizan utilizando los parámetros resumidos en la Tabla 1. Todas las

simulaciones fueron implementadas y ejecutadas utilizando la plataforma de modelado basado en agentes NetLogo. Los resultados cualitativos son robustos frente a variaciones moderadas de estos parámetros.

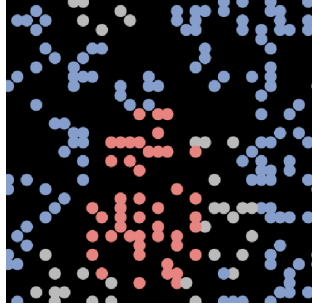
Table 1: Parámetros base de simulación.

Parámetro	Valor
Número de agentes $N$	400
Tamaño de la red	$25 \times 25$
Tolerancia $\varepsilon$	0.2–0.4
Tasa de aprendizaje $\mu$	0.5
Umbral de satisfacción $\theta$	0.5
Radio de interacción	1.5
Condiciones de contorno	Periódicas

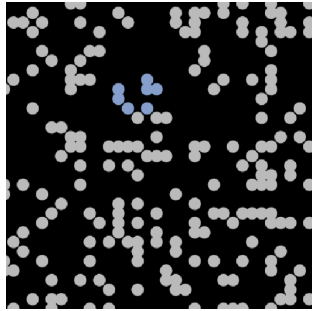
Cada simulación se ejecuta hasta que el sistema alcanza un estado estacionario. La convergencia se define operacionalmente como el punto en el que tanto la varianza global de opiniones como el número de agentes que se relocalizan permanecen estables durante una ventana temporal suficientemente larga (normalmente 500 pasos temporales consecutivos).



(a) Estado inicial ( $t = 0$ )

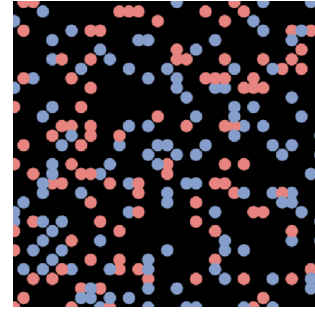


(b) Estado intermedio

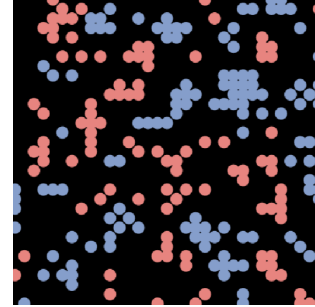


(c) Estado final (consenso)

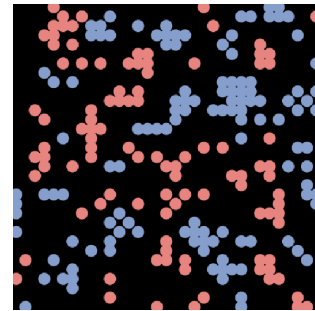
Figure 1: Evolución temporal del sistema bajo condiciones iniciales unimodales. Los agentes ocupan sitios de la red y están coloreados según su valor de opinión, que varía desde azul (opiniones bajas) hasta rojo (opiniones altas). El sistema converge hacia un consenso global.



(a) Estado inicial ( $t = 0$ )



(b) Estado intermedio



(c) Estado final (consenso)

Figure 2: Evolución temporal bajo condiciones iniciales bimodales. Dos clústeres de opinión persisten y se segregan espacialmente. Los colores indican valores de opinión y el agrupamiento espacial refleja la dinámica de segregación.

Para asegurar la reproducibilidad y facilitar futuras extensiones, el código fuente completo del modelo está disponible públicamente en un repositorio de GitHub<sup>1</sup>.

## 6 Resultados

Para cuantificar los resultados de opinión, calculamos la varianza de opiniones definida como

<sup>1</sup><https://github.com/paularodriguez3/dinamica-opiniones-abm>

$$\text{Var}(o) = \frac{1}{N} \sum_{i=1}^N (o_i - \bar{o})^2, \quad (4)$$

donde  $\bar{o} = \frac{1}{N} \sum_{i=1}^N o_i$  denota la opinión media.

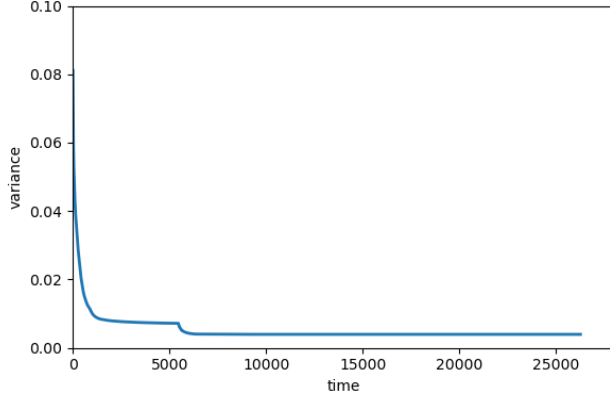


Figure 3: Varianza de opiniones a lo largo del tiempo para condiciones iniciales unimodales. La curva representa el promedio sobre múltiples ejecuciones independientes de la simulación. El sistema converge rápidamente hacia un consenso global.

En contraste, las condiciones iniciales bimodales combinadas con baja tolerancia producen una polarización persistente. La varianza de opiniones permanece alta y estable a lo largo del tiempo, y emerge la segregación espacial.

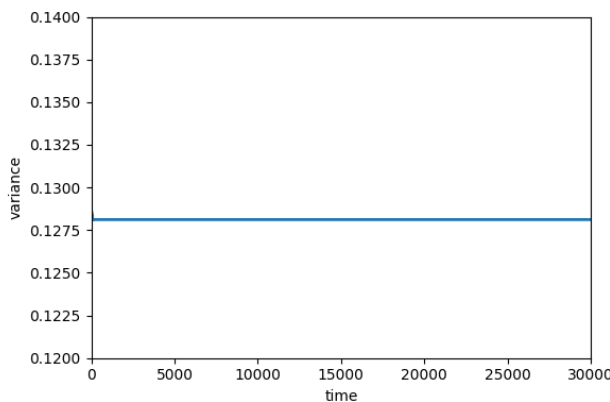


Figure 4: Varianza de opiniones a lo largo del tiempo para condiciones iniciales bimodales. La varianza permanece alta y estable, lo que indica polarización persistente. Las curvas representan promedios sobre múltiples ejecuciones de la simulación.

## 7 Discusión

Los resultados indican que los resultados colectivos están determinados conjuntamente por los niveles de tolerancia de los agentes y la distribución inicial de opiniones. En particular, la relocalización espacial amplifica los efectos de la confianza acotada al reforzar la exposición selectiva y limitar la interacción entre grupos, lo que a su vez estabiliza configuraciones polarizadas.

Es importante destacar que el modelo alcanza estados estacionarios sin depender de ruido ni de movimiento perpetuo, una característica que mejora su interpretabilidad. Este hallazgo es coherente con argumentos recientes que enfatizan el papel del ordenamiento social en la emergencia y persistencia de la polarización [4].

No obstante, deben reconocerse varias limitaciones del presente modelo. El entorno espacial se representa mediante una red discreta y las opiniones son unidimensionales, mientras que los sistemas sociales reales suelen involucrar espacios continuos y estructuras de creencias multidimensionales. Además, el modelo no incorpora la influencia de los medios ni choques externos, ambos factores que pueden desempeñar un papel significativo en la dinámica de opiniones del mundo real.

## 8 Conclusión

Este trabajo ha presentado un modelo basado en agentes mínimo pero teóricamente fundamentado que integra la dinámica de opiniones con confianza acotada y la segregación espacial al estilo de Schelling. A pesar de su simplicidad, el modelo reproduce resultados colectivos canónicos.

Al reunir dos tradiciones fundamentales de modelado, este trabajo contribuye a una comprensión más clara de cómo la influencia social y la organización espacial configuran conjuntamente la dinámica colectiva de opiniones.

## References

- [1] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. “Statistical physics of social dynamics”. In: *Reviews of Modern Physics* 81.2 (2009), pp. 591–646. DOI: 10.1103/RevModPhys.81.591.

- [2] Guillaume Deffuant et al. “Mixing Beliefs among Interacting Agents”. In: *Advances in Complex Systems* 3.1–4 (2000), pp. 87–98. DOI: 10.1142/S0219525900000078.
- [3] Morris H. DeGroot. “Reaching a Consensus”. In: *Journal of the American Statistical Association* 69.345 (1974), pp. 118–121. DOI: 10.1080/01621459.1974.10480137.
- [4] Andreas Flache et al. “Models of Social Influence: Towards the Next Frontiers”. In: *Journal of Artificial Societies and Social Simulation* 20.4 (2017). DOI: 10.18564/jasss.3521.
- [5] Santo Fortunato. “On the Consensus Threshold for the Opinion Dynamics of Krause–Hegselmann”. In: *International Journal of Modern Physics C* 16.2 (2005), pp. 259–270. DOI: 10.1142/S0129183105007111.
- [6] Rainer Hegselmann and Ulrich Krause. “Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation”. In: *Journal of Artificial Societies and Social Simulation* 5.3 (2002). URL: <http://jasss.soc.surrey.ac.uk/5/3/2.html>.
- [7] Jan Lorenz. “Continuous Opinion Dynamics under Bounded Confidence: A Survey”. In: *International Journal of Modern Physics C* 18.12 (2007), pp. 1819–1838. DOI: 10.1142/S0129183107011789.
- [8] Michael W. Macy et al. “Polarization in dynamic networks: A Hopfield model of emergent structure”. In: *American Sociological Review* 72.2 (2007), pp. 221–254. DOI: 10.1177/000312240707200205.
- [9] Thomas C. Schelling. “Dynamic Models of Segregation”. In: *Journal of Mathematical Sociology* 1.2 (1971), pp. 143–186. DOI: 10.1080/0022250X.1971.9989794.