

# MODELO TERRITORIAL PARA LA PRIORIZACIÓN DE INVERSIÓN EN INFRAESTRUCTURA

---

INFORME EJECUTIVO

**Observatorio de Ciudades UC**

Facultad de Arquitectura, Diseño y Estudios Urbanos  
Pontificia Universidad Católica de Chile

## INDICE

Introducción.....	3
I. Metodología del modelo de agrupamiento.....	4
II. Trabajo de campo .....	6
III. Resultados del modelo de agrupamiento rural.....	7
Macrozona Norte.....	8
Macrozona Centro .....	9
Macrozona Sur .....	10
Macrozona Austral .....	11
IV. Focalización de la inversión .....	12
V. Consideraciones .....	15

## Introducción

El modelo territorial surge como respuesta a las persistentes brechas de infraestructura básica que afectan a las zonas rurales de Chile, especialmente en materia de agua potable, saneamiento, conectividad vial y digital, resultado de una planificación históricamente centrada en lo urbano y de la limitada capacidad institucional para abordar la diversidad territorial. Estas carencias no solo deterioran la calidad de vida y las oportunidades de desarrollo de la población rural, sino que también profundizan las desigualdades entre territorios, generando la necesidad de contar con herramientas más precisas y coherentes para orientar la inversión pública.

Ante esta necesidad, el Ministerio de Obras Públicas, a través de la Dirección de Planeamiento, impulsa un enfoque de planificación territorial que reconozca la heterogeneidad geográfica, social y ambiental del mundo rural, y que permita priorizar la inversión en infraestructura de forma más eficiente, equitativa y transparente. El modelo territorial para la priorización de inversión en infraestructura nace precisamente para responder a este desafío, ofreciendo un marco analítico y operativo que articula datos espaciales, criterios de política pública y metodologías avanzadas de análisis.

El objetivo central del modelo es integrar información espacial, socioeconómica y ambiental para identificar y priorizar zonas rurales donde la inversión en infraestructura pueda contribuir, desde un enfoque de equidad territorial, a reducir brechas de conectividad y mejorar el acceso a servicios básicos. Para ello, el estudio desarrolla una metodología de agrupamiento territorial basada en algoritmos de clusterización por densidad, y sobre estos resultados construye un modelo de focalización territorial que evalúa la desigualdad o similitud de las condiciones estructurales de vida entre unidades espaciales (agrupamientos de viviendas, distritos, comunas, entre otros).

Un paso fundamental para definir esta metodología fue la revisión bibliográfica y la comparación entre tres algoritmos de clusterización basados en densidad: DBSCAN, HDBSCAN y OPTICS. Esta evaluación constituyó el punto de partida del modelo, dado que el algoritmo seleccionado permite comprender como se distribuye la población rural y, por tanto, cómo se interpretan sus patrones de concentración. En territorios altamente diversos, con asentamientos que varían desde micro-densificaciones hasta núcleos concentrados y zonas extremadamente dispersas, contar con un método adecuado fue esencial para asegurar que el agrupamiento reflejara fielmente la realidad territorial.

A partir del análisis inicial, el estudio optó por utilizar HDBSCAN como algoritmo central del modelo territorial. Su carácter jerárquico y su capacidad para representar múltiples niveles de densidad permitieron capturar de manera más robusta la diversidad de formas de poblamiento rural presentes en el país. DBSCAN fue descartado debido a la complejidad de calibrar simultáneamente sus parámetros en un territorio tan

heterogéneo, mientras que OPTICS presentó menor evidencia de uso operativo en contextos comparables, lo que reducía su conveniencia para un modelo nacional.

Sobre esta base metodológica, se construyó posteriormente el modelo de focalización territorial, integrado en un esquema multicriterio que articula factores territoriales, de infraestructura y de política pública. Tanto la metodología como los resultados fueron validados con analistas del MOP y contrastados mediante trabajo de campo. Finalmente, todas las herramientas desarrolladas —toolbox geoespacial (.tbx), dashboard interactivo y geodatabase estructurada bajo estándares IDE-MOP— quedan disponibles para consulta y análisis institucional.

## I. Metodología del modelo de agrupamiento

La metodología desarrollada para definir agrupamientos de viviendas rurales se basó en la construcción de un Modelo de Agrupamiento de Viviendas Rurales orientado a asegurar una representación completa, precisa y territorialmente coherente del poblamiento rural del país, a partir de una serie de procesos geoespaciales implementados en Python. El proceso combinó una revisión crítica de las bases existentes, la integración de nuevas fuentes de datos y la implementación del algoritmo HDBSCAN para identificar agrupamientos de viviendas con criterios de densidad espacial.

El punto de partida consistió en el preprocesamiento de los datos de entrada: las viviendas rurales en formato punto levantados por el INE en la Actualización Cartográfica Precensal 2023 (APC23). Dado que esta cobertura no incluye viviendas ubicadas al interior de Aldeas INE —representadas únicamente mediante polígonos— se identificó un sesgo relevante en zonas rurales con mayor concentración. Para corregirlo, se incorporaron huellas de edificios extraídas desde Overture Maps mediante procesos automatizados en Python, y posteriormente validadas según el número oficial de viviendas por Aldea. Esta integración permitió reconstruir el universo de viviendas dentro de polígonos de Aldea, ajustando duplicidades mediante muestreo aleatorio en los casos donde las huellas superaban el total reportado por INE.

En paralelo, se detectó un segundo tipo de inconsistencia en la APC23: edificaciones rurales duplicadas y superpuestas que afectaban la precisión del cálculo de distancias. Este problema fue corregido mediante procedimientos combinados en Rstudio, lo que permitió corregir 3.018 registros (0,1% del total) y asegurar la integridad geométrica de la base final.

El resultado de esta etapa fue una base consolidada de 1.657.970 viviendas rurales, unificada y estandarizada. Sobre esta capa se aplicó un conjunto de tratamientos técnicos —reproyección cartográfica, incorporación de campos con coordenadas UTM y variables complementarias— con el fin de garantizar la compatibilidad entre los datos, el algoritmo HDBSCAN y los flujos automatizados construidos en Python.

Con el objetivo de representar de forma más precisa la dinámica de los asentamientos rurales, el algoritmo HDBSCAN se aplicó a escala comunal. Este cambio metodológico permitió evitar distorsiones observadas en análisis regionales previos, asegurando que los agrupamientos reconocieran límites y barreras naturales propias de cada territorio y que reflejaran de mejor forma la estructura local del poblamiento.

Para definir el parámetro de mínimo de viviendas por agrupamiento (MinPts) se realizaron pruebas con valores de 20, 50, 100 y 200, considerando recomendaciones de la DIRPLAN y criterios asociados los Servicios Sanitarios Rurales (SSR) como infraestructura pública de referencia para determinar el mínimo de viviendas que son beneficiadas en zonas rurales. Tras las pruebas ejecutadas, el trabajo de campo y talleres de discusión con analistas regionales se determinó que un MinPts de 20 ofrece un equilibrio óptimo entre sensibilidad para identificar pequeñas densificaciones y robustez en los resultados. La implementación del algoritmo se realizó en Python por la compatibilidad con ArcGIS Pro, considerando que posteriormente el método de agrupamiento se implementaría mediante una herramienta toolbox en ArcGIS Pro.

Después de todas las pruebas, HDBSCAN fue ejecutado sobre las 312 comunas -con viviendas rurales- del país de forma eficiente y reproducible, con un Mínimo de Puntos (MinPts) de 20 y determinando la densidad a comunal.

Complementariamente, a partir de las viviendas clusterizadas se crearon polígonos que representan la extensión de cada agrupamiento, de modo de contar con una unidad espacial operativa para la fase de focalización. Para evaluar la calidad morfológica de estas unidades se construyó una clasificación de consistencia basada en la homogeneidad de la densidad local (distancia mutua de alcance, mrd). Así, se consideraron consistentes aquellos clústeres cuyas viviendas presentan densidades semejantes y dan lugar a polígonos regulares y compactos; y no consistentes aquellos en que la densidad interna varía de forma significativa y la geometría adopta formas alargadas o con “puntas”, típicas de los métodos jerárquicos del algoritmo.

Con el fin de caracterizar el tamaño de los clústeres según su cantidad de viviendas, se tomó inicialmente la tipología de operadores SSR (menor, mediano, mayor). Sin embargo, dado que cerca del 90% de los clústeres quedó en la categoría “menor”, se subdividió esta clase para mejorar la representatividad estadística. La tipología final quedó definida como: muy bajo (menos de 50 viviendas), bajo (entre 50 y 100), medio (entre 101 y 300), alto (entre 301 y 600) y muy alto (más de 601 viviendas). Esta clasificación permite distinguir con mayor precisión la estructura del poblamiento y orientar decisiones de infraestructura.

Para asegurar agrupamientos coherentes y homogéneos, se utilizó la distancia de alcance mutuo (MRD) como indicador de densidad local y se aplicó un filtro intra-clúster que elimina las viviendas con MRD atípicamente alto (percentil 90). Este depurado reduce los casos en que el efecto jerárquico de HDBSCAN incorpora puntos con

densidad insuficiente (MRD alto) y, en consecuencia, mejora la compacidad territorial de los clústeres.

A partir de este filtrado se definieron dos categorías analíticas de las viviendas ruidos (no clusterizadas):

1. Ruido permanente: viviendas que no cumplen los criterios del modelo en ninguna corrida o ajuste de parámetros, por lo que nunca califican para integrar un clúster.
2. Ruido satélite: viviendas que inicialmente fueron incorporadas por el algoritmo, pero que quedaron excluidas tras el filtro por presentar una densidad local discordante con la del clúster.

Esta distinción permite identificar bordes de expansión y zonas de transición del poblamiento que podrían integrarse ante ajustes futuros de parámetros, sin confundirlas con el ruido estructural que el modelo descarta de forma consistente.

Finalmente, los resultados se consolidaron en una capa nacional que, para cada vivienda, registra el identificador de clúster, la probabilidad de pertenencia, el puntaje de outlier y otras métricas de densidad, junto con la capa poligonal que delimita la extensión de cada agrupamiento y su clasificación por tamaño y consistencia. Estas salidas constituyen el insumo directo para el modelo de focalización y para su visualización en la plataforma WebGIS.

## **II. Trabajo de campo**

Tras el desarrollo metodológico de la primera versión del Modelo de Agrupamiento, se realizó un trabajo de campo para validar sus resultados en las macrozonas norte y sur del país. Esta etapa permitió validar la clusterización realizada por HDBSCAN con la estructura real de los asentamientos rurales, verificar su coherencia territorial y levantar información cualitativa necesaria para ajustar el modelo.

El levantamiento se organizó en dos fases. Primero se realizó un preterreno en la provincia de Quillota, donde se evaluaron los instrumentos de medición y la operatividad del método junto a analistas regionales de DIRPLAN. Esta instancia permitió afinar criterios, corregir los formularios de registro y ajustar la lectura de los clusters en la propia plataforma de terreno. Con estos ajustes, se desarrolló posteriormente una campaña más amplia en la macrozona Sur y otra en la macrozona Norte, abarcando un total de 23 localidades en las regiones de Biobío, Ñuble, Atacama y Coquimbo, en coordinación permanente con equipos regionales de DIRPLAN, Servicios Sanitarios Rurales y Vialidad.

Durante las visitas, el equipo evaluó si la delimitación de los clusters representaba adecuadamente la continuidad del poblamiento y su relación con el territorio. Esto implicó observar directamente la forma y extensión de los asentamientos, verificar la presencia de barreras naturales y antrópicas, y revisar la correspondencia entre los

agrupamientos obtenidos y la infraestructura rural existente. El uso de ArcGIS Field Maps y Survey123 permitió registrar información en tiempo real, integrando observaciones georreferenciadas que complementaron la interpretación del modelo.

Los hallazgos del terreno reforzaron varias decisiones metodológicas y permitieron identificar ajustes necesarios. Se confirmó que el parámetro MinPts debe adaptarse según las características regionales: en la macrozona Sur, donde existen densidades relativamente altas y grandes áreas productivas, valores menores permitieron capturar micro-densificaciones que el modelo omitía con parámetros más rígidos. En la macrozona Norte ocurrió algo similar, especialmente en zonas donde el desierto genera asentamientos dispersos; allí, valores elevados impedían reconocer agrupamientos que sí funcionan como unidades territoriales reales. También se observaron diferencias importantes en el rol de las barreras: mientras en el sur los ríos tienden a actuar como límites comunales, en el norte no necesariamente cumple ese rol; en cambio, rutas principales como la Ruta 5 sí funcionan como divisores efectivos en ambas macrozonas.

Asimismo, se confirmó la necesidad de incluir las Aldeas INE dentro de la base de viviendas, pues su omisión generaba vacíos significativos en zonas semiconcentradas. Asimismo, las observaciones permitieron validar la importancia del filtrado aplicado a la base APC23, ya que múltiples edificaciones clasificadas como viviendas correspondían en realidad a bodegas, galpones u otros usos.

### III. Resultados del modelo de agrupamiento rural

La aplicación del algoritmo HDBSCAN con un MinPts de 20 sobre la base consolidada de 1.657.970 viviendas rurales permitió identificar, en todo el país, un patrón desigual de concentración del poblamiento rural. A nivel nacional, el modelo agrupó cerca de dos tercios de las viviendas rurales (1.108.779 equivalente a un 66,9%), mientras que el tercio restante fue clasificado como ruido debido a su mayor aislamiento o discontinuidad territorial. Esta distribución evidencia que, aunque existe una amplia presencia de asentamientos rurales consolidados, también persiste una proporción significativa de viviendas dispersas que, por sus características, no alcanzan a conformar conglomerados suficientemente densos bajo los criterios del modelo.

Un aspecto relevante del análisis es la distinción entre el ruido permanente y el ruido satélite. En promedio, un 71% del ruido corresponde a viviendas que permanecen aisladas independientemente del ajuste del modelo, lo que confirma su carácter estructuralmente disperso. El 29% restante corresponde a viviendas cercanas a clusters que fueron excluidas por razones metodológicas, principalmente por encontrarse en el rango de densidad atípica definido por la distancia mutua de alcance. Estas últimas representan un conjunto relevante para el análisis territorial, ya que muestran bordes de expansión o transiciones entre patrones de ocupación.

A nivel de clusters, el país presenta 11.807 agrupamientos identificados, predominando ampliamente las categorías de agrupamiento “muy bajo” y “bajo”, lo que refleja la

fragmentación del poblamiento rural y la existencia de numerosos núcleos pequeños a lo largo del territorio. Ahora bien, esta tendencia se acentúa debido al valor de MinPts adoptado, que privilegia la detección de micro-densificaciones. La consistencia de los clusters es elevada en todas las regiones, superando el 90% en promedio, lo que valida la robustez del modelo y la estabilidad de los agrupamientos generados. No obstante, regiones con topografías más extremas o poblamiento más disperso —como Magallanes o zonas altoandinas— presentan un porcentaje ligeramente mayor de clusters no consistentes, reflejando patrones más irregulares.

El análisis nacional también muestra diferencias marcadas entre regiones con alta concentración de viviendas en pocos núcleos —como Antofagasta o Atacama— y regiones con alta cantidad de viviendas rurales, pero distribución más dispersa —como La Araucanía, Biobío o Los Lagos—. Este contraste confirma que la eficacia del modelo no depende del volumen total de viviendas, sino del patrón espacial de ocupación. En regiones con topografía extrema o disponibilidad limitada de recursos, los asentamientos tienden a organizarse en enclaves compactos, facilitando la clusterización. En regiones agrícolas o forestales, en cambio, la distribución más homogénea y discontinua genera mayor proporción de ruido y clusters pequeños.

En síntesis, el resultado nacional refleja un país con importantes diferencias territoriales en su poblamiento rural: mientras algunos territorios reproducen patrones concentrados y bien definidos, en otros se distinguen interfaces irregulares, y otros presentan una ruralidad más dispersa y heterogénea. El modelo logra capturar estas variaciones, ofreciendo una base útil para el análisis y la planificación de inversiones en infraestructura rural.

### Macrozona Norte

En la Macrozona Norte, el modelo muestra uno de sus desempeños más altos a nivel nacional, reflejando la tendencia histórica de esta zona a formar asentamientos rurales compactos, condicionados por factores ambientales, productivos y culturales. Regiones como Antofagasta, Atacama y Tarapacá alcanzan índices de clusterización superiores al 74%, evidenciando que la dispersión es menor y que las viviendas suelen agruparse en localidades definidas, ya sea en torno a valles, enclaves mineros o comunidades indígenas altoandinas.

En Antofagasta, la combinación de faenas mineras, localidades tradicionales y presencia de asentamientos indígenas genera agrupamientos densos y de contornos bien delimitados, especialmente en sectores asociados a Calama, María Elena y comunidades del altiplano. En Atacama, la ruralidad se estructura en torno a corredores naturales como los valles de Copiapó y Huasco, y al borde costero, donde existen asentamientos que, pese a su heterogeneidad, conforman clusters amplios y fácilmente detectables por el modelo. En Tarapacá, las comunas del interior —como Colchane, Camiña o Huara— muestran una alta capacidad de agrupamiento, incluso con

poblaciones pequeñas, debido a su carácter de micro-asentamientos concentrados en quebradas y altiplano.

En contraste, Arica y Parinacota y Coquimbo presentan índices ligeramente menores, reflejando una ruralidad más dispersa. En Coquimbo, particularmente, la configuración territorial de valles transversales y zonas de secano interior genera una ocupación discontinua, con viviendas distribuidas en quebradas, terrazas y laderas, dificultando la formación de clusters bajo criterios estrictos de densidad. En la Cordillera Andina del Norte Chico, las barreras naturales —cerros, quebradas profundas, divisorias de aguas— limitan la conectividad territorial, lo que explica que viviendas cercanas geoméricamente no pertenezcan a un mismo tejido rural. Este tipo de territorio es especialmente complejo para modelos basados exclusivamente en distancia, como HDBSCAN.

Las medidas metodológicas incorporadas —como ejecutar el algoritmo a escala comunal y reducir MinPts a 20— permitieron atenuar parte de estas limitaciones, mejorando la representatividad de los clusters en entornos complejos o dispersos. Aun así, el resultado refleja fielmente la estructura real del poblamiento: los territorios más aislados y fragmentados concentran mayor proporción de ruido, mientras que las zonas con asentamientos tradicionales más definidos muestran agrupamientos robustos.

En términos generales, la Macrozona Norte presenta un patrón dual: por un lado, altos niveles de clusterización asociados a enclaves mineros, valles y comunidades rurales compactas; por otro, territorios de montaña o secano donde la dispersión es mayor y el modelo clasifica una fracción relevante de viviendas como ruido permanente. Este comportamiento confirma que el poblamiento rural del norte se organiza en núcleos pequeños, pero bien definidos, lo que facilita su identificación para análisis territoriales y priorización de infraestructura.

### Macrozona Centro

En la Macrozona Centro —Valparaíso, Metropolitana, O'Higgins y Maule— se observa un patrón de poblamiento rural más continuo que en el norte, asociado a la morfología abierta del valle central y a la menor presencia de barreras físicas abruptas. Este contexto favorece una distribución más homogénea de las viviendas rurales y explica, en parte, que el nivel de agrupamiento alcance un 70,7%, ligeramente inferior a las macrozonas con poblamiento más concentrado. En total, el modelo identificó 4.223 clústeres que agrupan 429.543 viviendas rurales, con comportamientos territoriales diferenciados entre regiones: mientras Valparaíso, Metropolitana y O'Higgins presentan índices similares, en torno al 70–75%, el Maule muestra un valor algo menor debido a la mayor dispersión explicada por áreas agrícolas extensivas.

A escala provincial y comunal se mantienen estas tendencias. Las provincias ubicadas en el valle central, como Talagante o Maipo, presentan niveles altos de clusterización dado que las viviendas se organizan en continuidades asociadas a parcelas, predios

agrícolas o bordes de caminos, características de periferias metropolitanas con interfaces urbano-rurales difusas. En contraste, sectores con mayor ruralidad tradicional, como Cauquenes o Curicó, muestran una mayor proporción de viviendas aisladas. Las viviendas ruidó satélite son significativamente más frecuentes en las zonas más densas —particularmente en Talagante— lo que indica que parte del poblamiento posee continuidad territorial, pero no alcanza la densidad mínima para incorporarse al clúster bajo el criterio actual.

En cuanto al tamaño de los clústeres, predominan ampliamente los de magnitud “muy baja”, coherente con la definición metodológica basada en MinPts = 20. Sin embargo, en comunas con zonas rurales fuertemente pobladas o expansión periurbana —como Melipilla o Colina— emergen agrupamientos mucho mayores, reflejando la transición entre áreas rurales tradicionales y áreas de crecimiento urbano. La consistencia interna de los clústeres es mayoritariamente alta en toda la macrozona, superando el 90% en casi todas las provincias, lo que valida la estabilidad geométrica del agrupamiento. Solo algunas zonas de alta densidad, como Talagante, muestran porcentajes mayores de inconsistencia, relacionados con patrones más irregulares o extensiones alargadas de viviendas.

En conjunto, la Macrozona Centro exhibe un poblamiento rural continuo y diversificado, donde coexisten pequeños núcleos tradicionales con áreas periurbanas en consolidación. Esta heterogeneidad permite que el modelo identifique tanto densificaciones consolidadas como agrupamientos lineales vinculados a caminos y parcelaciones, configurando a la zona como un caso relevante para evaluar políticas de infraestructura rural en territorios de interfaz o transición urbano-rural. Desde esta perspectiva, resulta más pertinente abordarla mediante un enfoque de gradiente urbano-rural que a partir de categorías dicotómicas tradicionales.

### Macrozona Sur

La Macrozona Sur —Ñuble, Biobío, La Araucanía y Los Ríos— presenta uno de los patrones de ruralidad más complejos del país. La presencia de barreras naturales, como ríos, humedales, lagos y cordones montañosos, junto con un mayor poblamiento en áreas rurales, fragmenta el territorio y da lugar a una ruralidad dispersa, conformada por múltiples núcleos de distinto tamaño. Este contexto explica que el índice de clusterización alcance un 62,5%, menor que en el norte o en zonas extremas, donde el poblamiento tiende a concentrarse en enclaves más definidos. En total, se agruparon 382.225 viviendas rurales distribuidas en miles de núcleos de diferente magnitud, reflejando la diversidad funcional y geográfica de la macrozona.

Las diferencias intrarregionales son significativas: mientras provincias como Arauco o Concepción muestran mayores niveles de agrupamiento debido a la presencia de asentamientos concentrados vinculados a actividades productivas o a zonas de borde urbano, territorios como Ranco presentan menor clusterización debido a una ruralidad

muy dispersa y extendida. A nivel comunal, destacan casos robustos como Padre Las Casas o Villarrica, donde la expansión en torno a los núcleos urbanos genera densificaciones amplias y detectables por el modelo, muchas asociadas al turismo, a la expansión periurbana y a la parcelación de suelos rurales.

La proporción de ruido permanente se mantiene en torno al 70% en toda la macrozona, lo que refleja un patrón estructural de dispersión que no depende de la calibración del parámetro MinPts. Aun así, existe un volumen relevante de ruido satélite, cercano al 30%, que evidencia bordes de agrupamiento y transiciones territoriales donde el poblamiento no alcanza a consolidarse densamente, pero mantiene cercanía espacial con los clústeres principales.

En cuanto al tamaño de los clústeres, predominan ampliamente los de categoría “muy baja”, aunque se observan agrupamientos medianos y grandes en torno a corredores viales relevantes, lagos de alta atracción turística (Villarrica, Calafquén, Ranco) y valles agrícolas consolidados. Un caso particular es la provincia de Cautín, que presenta los clústeres más grandes de la macrozona, algunos superando las 10.000 viviendas rurales, lo que refleja la fuerte articulación territorial entre centros urbanos, actividad económica y vivienda.

La consistencia interna de los agrupamientos es en general alta, aunque provincias como Malleco presentan mayores niveles de inconsistencia debido a la coexistencia de áreas dispersas con densificaciones puntuales. En síntesis, la Macrozona Sur se caracteriza por un poblamiento rural diverso y fragmentado, en el que el modelo logra capturar tanto la dispersión estructural como la existencia de núcleos densos asociados a topografía, accesibilidad y actividades productivas.

### Macrozona Austral

La Macrozona Austral —Los Lagos, Aysén y Magallanes— presenta un patrón de poblamiento profundamente condicionado por la geografía. Las barreras naturales, la discontinuidad territorial y la escasa disponibilidad de suelo habitable determinan una organización rural fragmentada, aunque con núcleos relativamente compactos en áreas aptas para el asentamiento. En total, el modelo agrupó 148.456 viviendas rurales, lo que representa un 64,4% del total de la macrozona, con niveles de clusterización similares entre las tres regiones, pese a sus contrastes geográficos.

A escala regional, Los Lagos concentra la mayor cantidad absoluta de clústeres, aunque es también una de las regiones con más viviendas clasificadas como ruido, reflejando la coexistencia de asentamientos rurales densos con una dispersión significativa en áreas agropecuarias, forestales o insulares. Aysén y Magallanes presentan menores números absolutos debido a su baja población rural, pero proporciones de agrupamiento levemente superiores, explicadas por la presencia de localidades compactas y asentamientos organizados en torno a redes viales o bordes de lago.

Las diferencias provinciales muestran con claridad la diversidad territorial del extremo sur. Mientras Última Esperanza, Aysén y Chiloé presentan altos índices de clusterización, territorios como Tierra del Fuego y áreas insulares registran valores menores debido a la dispersión territorial, las grandes distancias y la fragmentación de los asentamientos. El ruido permanente alcanza valores elevados, entre 62% y 82%, lo que revela que gran parte del poblamiento rural austral está compuesto por viviendas aisladas o distantes entre sí. En cambio, el ruido satélite, aunque menor, evidencia zonas donde la transición rural es estrecha y los clústeres conviven con bordes difusos de viviendas que no alcanzan la densidad suficiente.

En relación con el tamaño de los clústeres, predominan los agrupamientos “muy bajos”, pero emergen clústeres medianos e incluso grandes en territorios asociados a actividades turísticas (como Chile Chico, Puerto Natales y Punta Arenas en el caso austral) o a valles productivos en Los Lagos y Aysén. Algunos clústeres de gran tamaño en Chiloé y Llanquihue reflejan dinámicas de ruralidad compleja, donde la disponibilidad relativa de suelo, la conectividad vial y la proximidad a centros urbanos permiten mayor densificación.

La consistencia interna de los clústeres es alta en la mayoría de las provincias, superando ampliamente el 90%, salvo en territorios extremos como Antártica Chilena o Chiloé (pequeñas islas), donde la irregularidad del continente genera mayor heterogeneidad en la densidad local. En conjunto, la Macrozona Austral muestra un poblamiento rural condicionado por la geografía y la accesibilidad, en el que el modelo logra identificar agrupaciones coherentes que reflejan tanto la concentración natural en zonas habitables como la dispersión estructural en territorios de menor aptitud.

#### **IV. Focalización de la inversión**

Los resultados del Modelo de Agrupamiento constituyen el insumo central para la etapa de focalización: los clústeres de viviendas rurales funcionan como unidades espaciales homogéneas sobre las cuales se evalúan brechas, vulnerabilidades y necesidades de infraestructura. En otras palabras, la focalización se construye directamente sobre la unidad territorial identificada en el modelo previo, permitiendo transitar desde el reconocimiento de patrones de poblamiento hacia un análisis integrado que orienta dónde y por qué intervenir.

El modelo territorial de focalización desarrollado en este estudio busca caracterizar de manera sistemática las condiciones estructurales de los territorios rurales y orientar la inversión pública desde un enfoque de equidad territorial. Para ello, integra dimensiones ambientales, demográficas, socioeconómicas y de accesibilidad, de modo que las decisiones de inversión consideren tanto la evidencia cuantitativa como el conocimiento cualitativo del territorio. En este marco, focalizar significa identificar los territorios que presentan mayores brechas, comprender las razones de esa

vulnerabilidad y orientar, *ex post*, la asignación de recursos hacia aquellos lugares donde la infraestructura puede generar mayor impacto social.

Durante el estudio se evaluaron distintos enfoques metodológicos para priorizar inversión rural. Entre ellos, el análisis multicriterio destaca por su capacidad para sintetizar información compleja mediante ponderaciones, aunque puede dificultar la explicación directa de los factores que influyen en la priorización. Las técnicas de clusterización, por el contrario, permiten caracterizar grupos territoriales con atributos similares, pero no determinan prioridades por sí mismas. Considerando estas fortalezas y limitaciones, se adoptó una estrategia modular que combina ambos enfoques de forma complementaria.

El modelo se estructura en cuatro fases. La primera evalúa la equidad territorial mediante análisis multicriterio, utilizando variables demográficas y socioeconómicas que permiten medir diferencias estructurales entre territorios. La segunda identifica brechas de infraestructura y equipamiento, evaluando la presencia, ausencia o distancia a servicios esenciales. La tercera incorpora una focalización auxiliar basada en *k-means* para identificar grupos poblacionales sensibles, como adultos mayores, mujeres jefas de hogar o menores de 14 años. Finalmente, la cuarta fase contextualiza el resultado incorporando amenazas, riesgos y atributos de habitabilidad, tales como campamentos, parcelas de agrado o remociones en masa.

En este enfoque, equidad territorial alude a la similitud o desigualdad de las condiciones estructurales de vida entre unidades espaciales —agrupamientos de viviendas, distritos, comunas, entre otros—. No se limita a variables de ingreso o pobreza, sino a un conjunto de atributos demográficos y socioeconómicos intrínsecos que posicionan a cada territorio en mejores o peores condiciones de partida para acceder a servicios e infraestructura pública. Este indicador, por sí solo, no establece la priorización, sino que alimenta un esquema metodológico que encadena todos los factores del análisis multicriterio y permite identificar territorios más favorables o rezagados.

Las variables consideradas fueron estandarizadas bajo criterios homogéneos, normalizadas entre 0 y 1 y ajustadas mediante transformaciones en los casos necesarios. Para evitar depender de unidades administrativas, el modelo opera sobre una malla regular H3 hexagonal (resolución L8 equivalente a 74 hectáreas aproximadamente), que permite integrar información territorial diversa en un soporte espacial único, manteniendo al mismo tiempo trazabilidad hacia escalas comunales, provinciales y regionales.

Todo el flujo metodológico fue implementado mediante Model Builder y empaquetado en una herramienta Toolbox, lo que constituye uno de los principales productos del estudio. Esta herramienta permite a analistas y equipos regionales reproducir, actualizar y adaptar el modelo: cambiar parámetros, incorporar nuevas variables, replicarlo para territorios específicos o integrarlo en ciclos regulares de planificación territorial. Es por

ello, que el capítulo no presenta resultados ni sus interpretaciones, sino el marco metodológico replicable y escalable que permitirá a los equipos del MOP aplicar el modelo según sus propias necesidades, territorios prioritarios y actualizaciones de información.

Finalmente, para garantizar la operatividad de los resultados, se desarrolló una plataforma WebGIS, tipo *experience* de ArcGIS (ESRI), que integra visualización multiescalar, paneles estadísticos e información metodológica. Diseñada bajo la lógica “ver-filtrar-explicar”, permite que cualquier usuario técnico explore los clústeres, revise el índice sintético y comprenda qué brechas específicas explican la prioridad territorial, articulando el conjunto del modelo con una interfaz clara y trazable.

En conjunto, este enfoque entrega al MOP una herramienta metodológica sólida, actualizable y transparente para orientar la inversión pública rural con criterios de equidad, pertinencia territorial y evidencia integrada.

## V. Consideraciones

Los modelos desarrollados en este estudio -el Modelo de Agrupamiento HDBSCAN y el Modelo Territorial de Focalización- constituyen un avance significativo para la planificación de infraestructura rural desde un enfoque de equidad territorial. Ambos modelos permiten comprender el poblamiento rural, identificar brechas estructurales y orientar decisiones de inversión con mayor precisión que los enfoques tradicionales. No obstante, para asegurar su vigencia y utilidad sostenida, es necesario consolidar ciertas prácticas técnicas, metodológicas e institucionales que fortalezcan su continuidad y permitan su evolución futura.

Un primer aspecto central es la dependencia del Modelo de Agrupamiento respecto a la capa de viviendas rurales y al parámetro de mínimo de puntos. El valor de MinPts de 20 fue consensuado entre DIRPLAN, los equipos regionales y el equipo consultor, pues permite reconocer pequeñas densificaciones y mantener coherencia territorial. Sin embargo, es importante considerar que el modelo entregará resultados distintos si este parámetro cambia. Esto no invalida los resultados del estudio: simplemente refleja que el modelo adopta 20 viviendas como unidad mínima de demanda para efectos de análisis territorial. Lo mismo ocurre con la capa de viviendas rurales utilizada como insumo: al basarse en la APC 2023 -complementada con las viviendas de Aldeas INE- la estructura del poblamiento respondía al estado del catastro al momento del análisis. Dado que no se proyectan futuras actualizaciones de la APC por parte del INE, resulta fundamental que DIRPLAN administre una capa maestra de viviendas rurales, actualizable periódicamente con nuevas fuentes (como Overture Maps u otros catastros sectoriales), garantizando que exista una versión “oficial” y consistente para ejecutar el modelo en los próximos años.

En relación con HDBSCAN, se reconoce que, al ajustar automáticamente el parámetro de distancia de búsqueda, de forma jerárquica, lo hace adecuado y robusto para analizar territorios diversos, pero al mismo tiempo requiere consideraciones metodológicas cuidadosas. La escala comunal, los filtros de densidad local, la distancia mutua de alcance, la identificación de viviendas ruido y la clasificación de consistencia de los clústeres son elementos que deben mantenerse de manera sistemática para asegurar resultados comparables. Asimismo, aunque HDBSCAN ha demostrado un desempeño robusto a escala nacional, otros algoritmos basados en DBSCAN pueden ser útiles a nivel local en casos excepcionales; por ello, es recomendable tratar este modelo como la herramienta estándar nacional, dejando la exploración de alternativas para situaciones específicas donde se requiera análisis diferenciado.

El Modelo de Focalización, por su parte, se beneficia de la estructura espacial generada por el agrupamiento, pero requiere incorporar información nueva y más diversa para robustecer su capacidad explicativa. Integrar variables actualizadas de equipamiento, accesibilidad efectiva, riesgos naturales y dimensiones socioeconómicas permitirá que el índice de equidad territorial evolucione hacia un indicador más sensible a las

transformaciones del territorio rural. En este marco, la focalización debe entenderse como un proceso dinámico, que se nutre tanto de la actualización temática como de la lectura técnica de los equipos regionales.

La integración futura del modelo al ciclo de inversión pública abre un espacio de desarrollo importante. Los resultados pueden constituir la base para la elaboración de una cartera preliminar de proyectos, articulando territorios priorizados, tamaños de agrupamiento, brechas de infraestructura y población beneficiaria. Este puente hacia la prefactibilidad permitiría traducir la priorización territorial en intervenciones concretas, alineadas con el Plan Nacional de Infraestructura Pública y los instrumentos de planificación del MOP. En esta fase, los clústeres pueden transformarse en unidades operativas para la programación plurianual, facilitando la coordinación con sectores como salud, educación, saneamiento, vialidad y gestión del agua.

Finalmente, la sostenibilidad de estos modelos requiere una gobernanza técnica estable dentro de DIRPLAN. La necesidad de que haya un profesional responsable de administrar los insumos, mantener la toolbox, ajustar parámetros y recopilar retroalimentación regional, permitirá que el modelo evolucione coherentemente y se mantenga útil más allá del ciclo del proyecto. Esta gobernanza debe complementarse con un proceso de actualización periódica -idealmente anual- que ejecute la toolbox con los insumos más recientes, monitoree cambios en los patrones de poblamiento rural y asegure la trazabilidad entre versiones.

En síntesis, el estudio deja disponible no solo dos modelos operativos, sino una arquitectura metodológica escalable, una toolbox replicable y un enfoque de equidad territorial que permite avanzar hacia una planificación más equitativa, transparente y basada en evidencia. Su continuidad dependerá de mantener actualizados los datos, resguardar los criterios metodológicos, fortalecer la gobernanza técnica y avanzar hacia su integración en la cartera de proyectos del MOP. Con estas condiciones, el modelo podrá consolidarse como una herramienta estratégica para orientar inversiones que transformen de manera efectiva la vida en los territorios rurales de Chile.