

# MODELO TERRITORIAL PARA LA PRIORIZACIÓN DE INVERSIÓN EN INFRAESTRUCTURA

---

INFORME FINAL

**Observatorio de Ciudades UC**

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

## CONTENIDO

ABREVIATURAS .....	3
1. INTRODUCCIÓN .....	4
1.1. Antecedentes .....	4
1.2. Objetivos.....	4
2. REVISIÓN DE FUENTES SECUNDARIAS.....	5
2.1. Base de datos .....	5
2.2. Homologación de datos geoespaciales.....	8
2.3. Revisión de criterios para definir asentamientos rurales en Chile.....	12
3. MARCO DE REFERENCIA PARA UN MODELO CLUSTERIZADO .....	16
3.1. Fundamentos de métodos de clusterización basados en densidad.....	16
3.2. Marco conceptual del modelo territorial de agrupamiento .....	19
4. METODOLOGÍA DEL MODELO DE AGRUPAMIENTO.....	25
4.1. Metodología.....	25
4.2. Preprocesamiento – Consolidación de viviendas en aldeas .....	25
4.3. Procesamiento – Tratamiento de Edificaciones Rurales APC23.....	32
4.4. Implementación del agrupamiento de viviendas.....	35
4.5. Calibración y filtros finales .....	38
5. TRABAJO DE CAMPO.....	43
5.1. Estructura y programación del trabajo de campo .....	43
5.2. Validación de la metodología para el trabajo de campo.....	47
5.3. Hallazgos .....	49
6. MODELO DE AGRUPAMIENTO RURAL .....	51
6.1. Resultados generales .....	51
6.2. Macrozona Norte .....	59
6.3. Macrozona Centro.....	72
6.4. Macrozona Sur.....	85
6.5. Macrozona Austral .....	98
7. FOCALIZACIÓN DE LA INVERSIÓN.....	110
7.1. Estandarización de las variables .....	114
7.2. Método de cálculo .....	114
7.3. Escalas de análisis .....	115
7.4. Visualización de resultados.....	116

8.	CREACIÓN DE UN TOOLBOX PARA ARCGIS PRO.....	120
9.	CONSIDERACIONES.....	123
10.	BIBLIOGRAFÍA.....	126
11.	ANEXOS.....	130
11.1.	Anexo 1. Script de código R desarrollado para la homologación y estandarización de datos geoespaciales. ....	130
11.2.	Anexo 2. Caracterización de los nueve territorios CASA (ODEPA, 2024) .....	140
11.3.	Anexo 3. Resultados del algoritmo HDBSCAN a nivel de comuna. ....	140
11.4.	Anexo 3. Diccionario de datos. ....	150
11.5.	Anexo 4. Resultados del algoritmo HDBSCAN en clústeres a nivel de comuna (MB: Muy Bajo, B: Bajo, M: Medio, A: Alto, MA: Muy Alto, C: Consistente y I: Inconsistente).....	154

## ABREVIATURAS

APC: Actualización Cartográfica Precensal

BCN: Biblioteca del Congreso Nacional

CIREN: Centro de información de Recursos Naturales

CONADI: Corporación de Desarrollo Indígena

CUT: Código Único Territorial

DBSCAN: *Density-Based Spatial Clustering of Applications with Noise*

DGA: Dirección General de Aguas

DIRPLAN: Dirección de Planeamiento

HDBSCAN: Hierarchical Density-Based Spatial Clustering of Applications with Noise

INE: Instituto Nacional de Estadística

MINVU: Ministerio de Vivienda y Urbanismo

ODEPA: Oficina de Estudios y Políticas Agrarias

OPTICS: *Ordering Points To Identify the Clustering Structure*

OSM: Open Street Maps

SECTRA: Secretaría de Planificación de Transporte

SISS: Superintendencia de Servicios Sanitarios

SUBDERE: Subsecretaría de Desarrollo Regional y Administrativo

# 1. INTRODUCCIÓN

## 1.1. Antecedentes

Las zonas rurales del país enfrentan un déficit de infraestructura básica —como agua potable, saneamiento, conectividad vial y digital— derivado de factores históricos y estructurales, incluyendo una planificación centrada en lo urbano, procesos recientes de migración ciudad–campo, y una limitada capacidad de respuesta institucional a la diversidad territorial (Orellana *et al.*, 2021). Estas brechas no solo afectan la calidad de vida y las oportunidades de desarrollo de sus habitantes, sino que también profundizan las desigualdades territoriales.

Frente a esto, el Ministerio de Obras Públicas (MOP), a través de su Dirección de Planeamiento (DIRPLAN), ha impulsado una estrategia de planificación territorial con foco en las zonas rurales, reconociendo su heterogeneidad geográfica, social y ambiental. En este marco, la priorización eficiente y equitativa de las inversiones se vuelve clave para avanzar hacia un desarrollo territorial más justo y resiliente.

El presente estudio busca contribuir a ese objetivo mediante el desarrollo de un modelo territorial que permita identificar y priorizar zonas rurales donde la inversión pública en infraestructura sea más urgente y estratégica. Para ello, se utilizarán metodologías avanzadas de análisis espacial y socioeconómico, incluyendo algoritmos de clusterización (HDBSCAN) y criterios de sostenibilidad, equidad y vulnerabilidad climática. Asimismo, se contempla la creación de una base de datos geoespacial bajo estándares IDE-MOP, el desarrollo de una herramienta automatizada en ArcGIS Pro y una plataforma de visualización interactiva para apoyar la toma de decisiones del MOP y otros organismos públicos.

## 1.2. Objetivos

### *Objetivo General*

Desarrollar un modelo territorial que integre información espacial, socioeconómica y ambiental para priorizar inversiones de infraestructura en zonas rurales, orientado a reducir brechas de conectividad y mejorar el acceso a servicios básicos.

### *Objetivos Específicos*

- a) Diseñar y aplicar una metodología de agrupamiento territorial que permita identificar niveles de concentración de viviendas rurales.
- b) Ajustar y validar la metodología de agrupamiento, y construir una base de datos geoespacial (GDB) estructurada bajo estándares IDE-MOP.
- c) Elaborar un sistema de priorización de inversión en infraestructura, basado en criterios de sostenibilidad, equidad territorial y vulnerabilidad climática, que permita generar escenarios estratégicos de intervención.
- d) Desarrollar una herramienta toolbox automatizada e implementar un sistema de visualización geográfica para la interpretación de resultados.

## 2. REVISIÓN DE FUENTES SECUNDARIAS

Este capítulo presenta el proceso de recopilación, organización y estandarización de las fuentes secundarias utilizadas en el estudio. Se detalla la conformación del repositorio preliminar de datos geospaciales y el proceso de homologación técnica según la normativa IDE MOP, pasos fundamentales para la posterior creación de la Geodatabase con estándar de la IDE MOP.

Además, se presenta una revisión crítica de los criterios existentes para la definición de asentamientos rurales en Chile, incluyendo tanto definiciones normativas como metodologías recientes basadas en análisis espacial, lo que permite contar con un marco conceptual actualizado para la caracterización de lo rural, y la posterior categorización de los clusters resultantes del modelo de agrupamiento.

### 2.1. Base de datos

El repositorio de datos corresponde a la consolidación de todos los shapefiles utilizados en el estudio, tanto para la aplicación del algoritmo HDBSCAN como para la caracterización territorial, provenientes de diversas fuentes oficiales y secundarias. Otro aspecto relevante es que esta BBDD contiene la información secundaria recopilada y los resultados generados en el estudio, como los clústeres (polígonos), la caracterización territorial o la priorización, que se encuentran entregados y validados por la contraparte técnica del estudio.

La Tabla 1 incluye capas auxiliares (por ejemplo, límites comunales en el agrupamiento HDBSCAN), así como insumos para el análisis de datos. Esta versión final está compuesta de 43 capas, organizadas en 5 categorías temáticas: Cartografía Base, Infraestructura, Social y Habitacional, Equipamiento, y Medio Ambiente, además de las capas de análisis (de elaboración propia). Las capas están organizadas en una Geodatabase (GDB) estandarizada y consolidada, acorde a la norma IDEMOP.

La organización temática responde a los lineamientos de la norma ISO 19115-1, que establece un marco referencial para la estructuración y documentación de bases de datos geospaciales mediante categorías y subcategorías normalizadas.

Esta última versión incluye, además de la información secundaria recopilada, los resultados de los análisis realizados por el equipo.

Tabla 1: Catálogo de capas de la geodatabase IDE MOP.

Categoría	Capa	Fuente	Año	Nombre layer en GDB
Cartografía Base	Viviendas rurales 2023	INE	2023	VIVIENDAS_RURALES
Cartografía Base	Limites comunales	INE	2024	LIMITE_COMUNAL
Cartografía Base	Índice de envejecimiento	INE	2024	IND_ENVEJECIMIENTO

Categoría	Capa	Fuente	Año	Nombre layer en GDB
Cartografía Base	Índice de feminidad	INE	2024	IND_FEMINIDAD
Cartografía Base	Jefatura de hogar de personas mayores	INE	2024	IND_JEFATURA_HOGAR
Cartografía Base	Viviendas desocupadas comunal	INE	2024	VIVIENDAS_DESOCUPADAS_COMUNAL
Cartografía Base	Cantidad de habitantes comunal	INE	2024	HABITANTES_COMUNAL
Cartografía Base	Cantidad de viviendas comunal	INE	2024	CANT_VIVIENDAS_COMUNAL
Cartografía Base	Limite Urbano Censal	INE	2024	LIMITE_URBANO_CENSAL
Cartografía Base	Limite Urbano IPT	MINVU	2022	LIMITE_URBANO_IPT
Cartografía Base	Área Urbana Consolidada	MINVU	2024	AUC
Equipamiento	Establecimientos educación escolar	MINEDUC	2021	ESTABLEC_EDUCACION
Equipamiento	Establecimientos de salud	MINSAL	2024	ESTABLEC_SALUD
Infraestructura	Red vial	MOP	2024	RED_VIAL
Infraestructura	Servicios Sanitarios Rurales	MOP	2024	AGUA_POTABLE
Infraestructura	Aeródromos	MOP	2014	AERODROMOS
Infraestructura	Planta de tratamiento de aguas servidas	SISS	2022	PTAS
Infraestructura	Territorios operacionales	SISS	2025	TERRITORIOS_OPERACIONALES
Infraestructura	Embalses	MOP	2024	EMBALSES
Infraestructura	Infraestructura portuaria y costera	MOP	2023	INFR_COSTERA
Infraestructura	Puentes	MOP	2025	PUENTES
Infraestructura	Red Aeroportuaria	MOP	2025	RED_AEROPORTUARIA

Categoría	Capa	Fuente	Año	Nombre layer en GDB
Social y habitacional	Índice Socio Material y Territorial	OCUC	2022	ISMT
Social y habitacional	Áreas de Desarrollo Indígena	CONADI	2024	AREAS_DESARROLLO_INDIGENA
Social y habitacional	Campamentos	MINVU	2024	CAMPAMENTOS
Social y habitacional	Parcelas de agrado	MINVU	2024	PARCELAS_AGRADO
Medio ambiente	Red hidrográfica nacional	BCN	2017	RED_HIDROGRAFICA_NACIONAL
Medio ambiente	Cuencas BNA	DGA	2024	CUENCAS_BNA
Medio ambiente	Acuíferos SHAC	DGA	2024	ACUIFEROS_SHAC
Medio ambiente	Acuíferos y Vegas Protegidas	DGA	2024	ACUIFEROS_PROTEGIDOS; VEGAS_PROTEGIDAS
Medio ambiente	Amenaza volcánica	SERNAGEOMIN	2024	AMENAZA_VOLCANICA
Medio ambiente	Amenaza de tsunami e inundaciones	IGM	2023	COTA_30
Medio ambiente	Recurrencia de incendios forestales	CONAF	2024	INCENDIOS_FORESTALES
Medio ambiente	Afectaciones a la infraestructura	SENAPRED	2024	AFECCIONES_INFRAESTRUCTURA
Medio Ambiente	Catastro de remociones en masa	SENAPRED	2024	REMOCIONES_MASA
Análisis Territorial	Análisis territorial por comuna - Medidas de centralidad de kmeans	Elaboración Propia	2026	ANALISIS_COMUNA_MC_KM
Análisis Territorial	Análisis territorial por distrito censal - Medidas de	Elaboración Propia	2026	ANALISIS_DISTRITO_MC_KM

Categoría	Capa	Fuente	Año	Nombre layer en GDB
	centralidad de kmeans			
Análisis Territorial	Análisis territorial por provincia – Medidas de centralidad de kmeans	Elaboración Propia	2026	ANALISIS_PROVINCIA_MC_KM
Análisis Territorial	Análisis territorial por región – Medidas de centralidad de kmeans	Elaboración Propia	2026	ANALISIS_REGION_MC_KM
Clustering HDBSCAN	Clusters HDBSCAN con caracterización	Elaboración Propia	2026	CLUSTERS_HDBS_DETALLADO
Grillas Hexagonales	Grilla H3 regional	Elaboración Propia	2026	GRID_H3_REGIONAL
Grillas Hexagonales	Grilla H3 nacional	Elaboración Propia	2026	GRID_H3_NACIONAL_FINO

## 2.2. Homologación de datos geoespaciales

Al provenir de distintas fuentes y representar distintas temáticas, las coberturas obtenidas para el repositorio preliminar presentan distintos (o ningún) estándares a la hora de estructurar la información. En general, se encontraron inconsistencias que se repetían a través de datasets, como el formato de los CUT (muchas veces pasado a numérico por lo que se pierde el 0 a la izquierda de las primeras 9 regiones que normaliza la variable), los nombres de variables, los nombres de archivos y, la más constante, los diferentes sistemas de coordenadas. En contraste, la IDE MOP especifica una serie de

normas para esta estructura que se aplicaron al nuevo repositorio, basadas en la norma ISO-19110 y que se encuentran descritas en la normativa<sup>1</sup> geoespacial oficial del MOP.

Producto de lo anterior se realiza un proceso de homologación de datos a través del lenguaje de programación R, proceso que entrega como resultado una serie de capas organizadas que cumplen con los estándares especificados por la IDE MOP, en un estado listo para ser agregados directamente a la Geodatabase final.

Se entenderá por homologación un proceso de manipulación de los datos para lograr una estructura y formato consistentes, que permitan una alta compatibilidad y por lo tanto una fácil interacción entre los distintos datasets en procesos posteriores, independiente de su fuente, formato o estructura de origen. En este caso, la estructura y formato al que los datasets son adaptados están definidos por la norma ISO-19110, referenciada en la normativa oficial IDE MOP. A continuación, se detallan los 4 pasos de homologación que se aplicaran a los datos geoespaciales.

### 1. Sistema de coordenadas geográficas

Si bien la normativa IDE MOP sólo habla en términos generales del sistema de referencia geodésico SIRGAS Chile (p. 14), para este proyecto se especificó el uso del sistema SIRGAS Chile 2016, EPSG 9153. En R, se utiliza el paquete **sf** para realizar la transformación.

### 2. Nombres de variables

Los nombres de las variables se adhieren a las especificaciones de la sección “Nombre de campo” de la normativa IDE MOP (p. 20), que prohíben el uso de ñ, tildes, espacios, números como primer carácter, caracteres especiales, y fechas, además de exigir el uso de mayúsculas. El proceso en R utiliza paquetes de manipulación de datos pertenecientes al **tidyverse**, como **dplyr** y **stringr**, así como **R base**. A través de la base de datos, se pone cuidado de mantener la consistencia en nombres de variables que pudieran repetirse, tales como las variables de identificación geográfica (CUT) y las toponimias, de forma de minimizar la necesidad de recurrir a un diccionario de variables.

### 3. Nombre de archivo

Los nombres de archivo siguen lo explicitado en la sección “Nombre del elemento” de la normativa IDE MOP. Dentro del proceso de homologación se respetó la lógica descrita en la normativa, con el nombre en sí mismo manteniendo el estándar descrito en el punto anterior (“Nombres de variables”) en términos de caracteres permitidos y formato. Estos nombres normalizados se aplicaron al nombre de cada layer dentro de la Geodatabase, y la correspondencia con el dato está descrita en la Tabla 1. Al considerarse como layers “estáticas” (es decir, sin actualización sistemática y

<sup>1</sup> Normativa para Base de Datos IDEMOP. Unidad de Gestión de la Información Territorial - IDEMOP. Julio 2022.

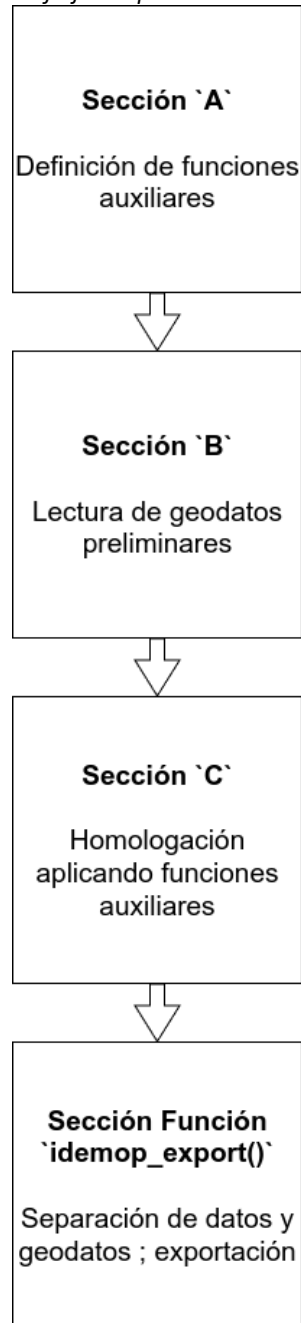
periódica), no aplica la lógica de separación de datos tabulares y datos espaciales, por lo que los layers se mantienen unidos.

#### 4. Toponimias

De la misma forma que en el caso de los nombres de variables y de archivos, la normativa IDE MOP exige el uso de un formato específico, que a su vez limita el universo de caracteres permitidos. La aplicación de esta norma resulta en una variable de texto absolutamente normalizada al estándar mencionado, pero que requiere un paso adicional para generar cartografía adecuada.

En el diagrama que se presenta a continuación (Figura 1), se puede visualizar el flujo de trabajo del proceso de homologación. En los recuadros flotantes, se especifica a qué sección del script incluido en el Anexo 1 corresponde cada parte del proceso.

Figura 1: Diagrama de flujo del proceso de homologación.



A continuación, se describen las secciones representadas en el diagrama:

1. Sección `A`: Es la sección dedicada a definir funciones auxiliares para el resto del script. Tareas repetitivas (como p.ej. obtener el CUT de región desde el CUT de comuna, la normalización de un *string* a las normas IDEMOP, o la separación de datos y geodatos) son resumidas en un *snippet* de código que acepta uno o más parámetros para poder iterar sobre los distintos datasets. De esta forma, se

- evitan redundancias en el código. Todas las reglas de homologación definidas en esta sección provienen del documento de Normativa IDEMOP.
2. Sección `B`: Corresponde a la lectura de los geodatos preliminares (descritos en el punto 2.1) al ambiente de programación. Sin este paso, el script no puede acceder a los datasets para modelar la información.
  3. Sección `C`: Se trata del núcleo del proceso de homologación. El script identifica los datasets que deben ser estandarizados y aplica sobre ellos las funciones definidas en la Sección `A`, además de procedimientos propios a la especificidad de cada dataset. Los procesos que se repiten de un dataset a otro se aplican a través de las funciones, mientras que las particularidades de cada dataset se tratan individualmente. Esta sección resulta en los datasets ya estandarizados a la normativa, aún cargados al ambiente de desarrollo.
  4. Sección de aplicación de función `idemop_export()`: Entre las funciones definidas en la Sección A, se encuentra la función `idemop_export()`, que engloba toda la parte final del script. Incluye la reparación y el formateo de geometrías para compatibilizar con el driver de GDB. Modularizar este paso permite aplicarlo de forma transversal a todos los layers.

El resultado de la ejecución completa del script es una Geodatabase (GDB) compilada con todos los layers listados en la Tabla 1, con el objetivo de ser incorporada de forma óptima a la IDE MOP.

### 2.3. Revisión de criterios para definir asentamientos rurales en Chile

En Chile existen diversas definiciones para establecer cuándo una zona se considera rural, las cuales se han caracterizado por tener un enfoque principalmente operativo, vinculado al programa o política en cuestión. Con el objetivo de recopilar lineamientos que permitan una categorización integral de las viviendas rurales, se realizó una revisión de las distintas clasificaciones existentes en el país, considerando tanto la normativa vigente como políticas públicas y metodologías desarrolladas recientemente.

Las principales definiciones de “lo rural” provienen de tres fuentes: la Ley General de Urbanismo y Construcciones (LGUC), que establece criterios normativos para distinguir entre zonas urbanas y rurales; el Instituto Nacional de Estadísticas (INE), cuya definición —aunque no oficial— ha sido ampliamente utilizada con fines operativos y estadísticos; y la Política Nacional de Desarrollo Rural (PNDR), que propone una concepción más integral del desarrollo rural, reconociendo su diversidad territorial y funcional (ODEPA, 2022).

Además de estas definiciones, en los últimos años se han desarrollado metodologías que permiten una caracterización más detallada del territorio a escala comunal y subcomunal, incorporando herramientas de análisis espacial y técnicas estadísticas. Entre estas destacan tres propuestas relevantes:

- La propuesta de SUBDERE (2020), contenida en el estudio *“Asentamientos humanos rurales en Chile, una clasificación comunal: Una aproximación desde el análisis espacial de la concentración y dispersión de la población”*. Esta clasificación Utiliza métodos de clusterización –como K-means y agrupamientos jerárquicos– para clasificar las comunas considerando variables asociadas a movilidad y conectividad interna. La propuesta reconoce la heterogeneidad territorial rural y ofrece una categorización basada en la accesibilidad y movilidad dentro de cada comuna.
- La *“Metodología para determinar las áreas funcionales de Chile”* (INE et al., 2020), elaborada por INE, MINVU, SUBDERE y SECTRA, propone una clasificación territorial basada en la funcionalidad urbana (AFU). Utiliza variables como conectividad, acceso a servicios y flujos de movilidad cotidiana para identificar centros urbanos y sus zonas de influencia. Esta propuesta reconoce relaciones urbano-rurales reales y supera los límites administrativos tradicionales. Es clave para comprender la integración funcional del territorio y delimitar zonas rurales en transición.
- La propuesta de ODEPA (2025), desarrollada en el estudio *“La gradiente urbano-rural en Chile: Una propuesta de clasificación del territorio a escala sub-comunal con relevancia para el ciclo de políticas públicas en Chile”*. La metodología establece una clasificación continua del territorio basada en un gradiente urbano-rural aplicado a nivel de distrito censal. Utiliza variables como densidad poblacional y tiempo de desplazamiento a centros urbanos, permitiendo caracterizar con mayor precisión zonas rurales, periurbanas y urbanas. Una de sus fortalezas es la definición de nueve categorías territoriales, lo que refleja la diversidad y gradualidad del territorio más allá de lo dicotómico urbano–rural.

Estas metodologías complementan las definiciones normativas existentes y entregan herramientas actualizadas para abordar la heterogeneidad del mundo rural chileno, especialmente en el marco del diseño e implementación de políticas públicas orientadas a reducir brechas territoriales. Lo anterior, se resume en la Tabla 2.

Tabla 2: Listado comparativo de propuestas y definiciones de lo rural.

Uso/Objetivo	Normativa Institución	Definición	Criterios	Escala	Dato geoespacial
Ordenamiento Territorial/ Zonificación	LGUC/ OGUC	Territorio ubicado fuera del límite urbano (Art. 1.1.2., OGUC)	Línea imaginaria que diferencia el área urbana del territorio rural	Sub-comunal	Zonificación IPT; Límite Urbano

Uso/Objetivo	Normativa Institución	Definición	Criterios	Escala	Dato geoespacial
Información	INE	Asentamiento con población menor o igual a 1.000 habitantes, o entre 1.001 y 2.000 habitantes, donde más del 50% de habitantes se dedica a actividades primarias	Tamaño poblacional, población que declara haber trabajado y características geográficas	Sub-comunal	Cartografía censal rural
Desarrollo Territorial	PNDR	Territorio cuya densidad poblacional es menor a 150 (habs/km <sup>2</sup> ), con una población máxima 50.000 habitantes, y la unidad básica de organización es la comuna	Densidad poblacional	Comunal	Indicadores SICVIR
Metodologías adicionales	SUBDERE 2020	Clasificación a nivel de complejidad territorial, a partir de un modelo basado en 4 variables	Movilidad y accesibilidad interna comunal, respecto al centro comunal	Comunal	Comunas clasificadas
	MINVU, SECTRA & INE 2020	Todo lo que no se clasifica como área funcional urbana.	Funcionalidad urbana	Sub-comunal	Áreas funcionales
	ODEPA 2025	Áreas a lo largo del continuo urbano-rural según sus características de densidad y distancia a una escala de distrito	Densidad de población (hab/km <sup>2</sup> ) y tiempo de desplazamiento (hrs) al principal centro urbano	Distrito	Distritos clasificados según las 9 categorías propuestas

Fuente: Elaboración propia, a partir de ODEPA (2025).

A partir de esta revisión, solo dos metodologías serán utilizadas como insumo para la caracterización territorial del Modelo (ver apartado 7), rescatando elementos específicos que permiten dar continuidad conceptual y evitar la creación de nuevas categorías. A continuación, se detalla de que forman serán consideradas:

- SUBDERE (2020). De esta metodología se rescatan dos elementos que se aplicaran en el Capítulo 7: (i) el uso de índices de aislamiento como factor determinante en la tipología del asentamiento; y (ii) el uso del método *K-means* como herramienta de clusterización, útil para estandarizar la categorización entre clústeres rurales considerando su diversidad territorial.

- ODEPA (2024). A partir de la cobertura geoespacial levantada con las nueve tipologías, se utilizó como un input en el Modelo Territorial de Focalización (Capítulo 7), en donde cada celda de la grilla hexagonal H3 está asociada a una categoría de la clasificación CASA. De este modo, cada unidad territorial resultante de la Focalización tendrá asociada una categoría de la gradiente urbano-rural. En el Capítulo 7 se detalla metodológicamente como se incorporan esta información.

En síntesis, la incorporación de estas metodologías permite dar continuidad al trabajo previamente desarrollado por las instituciones, evitando la creación de nuevas categorías y asegurando coherencia con lineamientos oficiales. Al mismo tiempo, se privilegia el uso de métodos y unidades territoriales que son compatibles y comparables en el tiempo, lo que garantiza la pertinencia y robustez de la caracterización territorial.

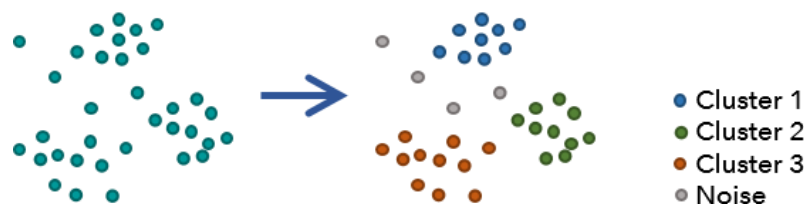
### 3. MARCO DE REFERENCIA PARA UN MODELO CLUSTERIZADO

Los métodos de agrupamiento (clustering o clusterización) son herramientas claves para identificar estructuras y patrones en datos complejos. Por este motivo, a continuación, se presentan los principales fundamentos conceptuales y teóricos de tres algoritmos basados en densidad (DBSCAN, HDBSCAN y OPTICS), analizando su aplicación internacional-nacional, funcionamiento, ventajas y limitaciones. Esto, para finalmente establecer un marco comparativo para justificar la elección del método más adecuado para el estudio, describiendo en detalle su funcionamiento y aplicación.

#### 3.1. Fundamentos de métodos de clusterización basados en densidad

En el marco de la clusterización espacial, se han implementado distintos métodos para comprender las dinámicas rurales y apoyar el desarrollo sostenible. Un avance significativo ha sido el uso de algoritmos de aglomeración basados en densidad (Figura 2), como es el DBSCAN y sus variantes: HDBSCAN y OPTICS (Schubert et al., 2017).

Figura 2: Esquema de clusterización por densidad.



Fuente: ESRI, 2025a.

A nivel internacional, estos algoritmos han sido implementados mayoritariamente en China, como es el caso de Liu et al. (2007), quienes destacaron su utilidad en la planificación de infraestructuras rurales, optimizando la identificación de áreas prioritarias para servicios básicos como el agua potable, o Zhou et al. (2017), quienes lo aplicaron para mapear áreas urbanas, superando las limitaciones de métodos tradicionales. Además, destaca el caso del Ministerio de Transportes y Comunicaciones en Perú que implementó DBSCAN para identificar y priorizar intervenciones en la red vial del país (Illacanchi, 2021).

Asimismo, los algoritmos se han aplicado para optimizar el acceso a servicios, como el transporte público entre zonas urbanas y rurales (Li et al., 2022), la ubicación estratégica de centros de vacunación (Said et al., 2022), la planificación de infraestructura energética (Namaganda & Mutale, 2020; Rahimi et al., 2025), optimización de rutas de transporte público (Reda et al., 2024; Ng et al., 2025) y en la identificación de patrones de concentración de hostales en áreas urbanas y rurales (Wang & Ma, 2024).

A nivel nacional, no existe una implementación amplia de los algoritmos, encontrando estudios destinados principalmente a análisis de transporte y movilidad (Ramírez, 2022; Castillo et al., 2024). No obstante, también se ha implementado en otras áreas, como la detección de eventos sísmicos (Nicolis et al., 2024; García et al., 2024), análisis económicos (Banco Central Chile, 2022), entre otros.

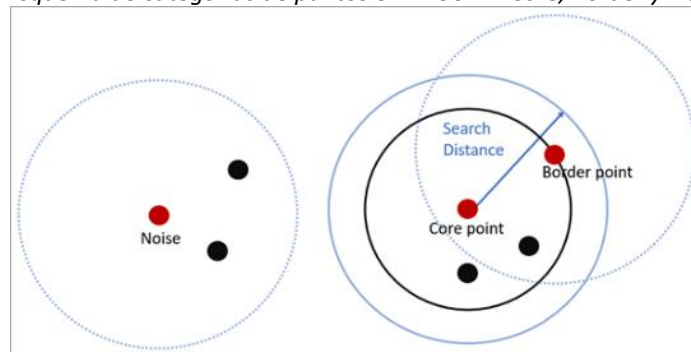
Es relevante mencionar que, si bien en estos estudios el uso de los algoritmos no implicó predefinir el número ni forma de los clusters, estos sí exigen establecer los parámetros: distancia (*Eps*) y densidad (*MinPts*). Esta exigencia puede ser una limitación metodológica, pero también ofrece la posibilidad de adaptar el análisis a criterios o normativas específicas.

Específicamente, los algoritmos de la familia DBSCAN (HDBSCAN y OPTICS) son métodos de agrupamiento avanzados que permiten analizar grandes volúmenes de datos geospaciales (Tabla 3). A diferencia de otros métodos tradicionales, se basan en densidad, no requieren de hipótesis predefinidas (por ejemplo, el número o forma de los clusters) y permiten que ciertas observaciones queden sin agrupar, pudiendo ser interpretados como outliers (Anselin, 2020). Estas características les otorgan mayor flexibilidad y gran capacidad para identificar patrones emergentes, especialmente cuando los datos son dispersos, mejorando la representatividad y precisión.

Tabla 3: Comparación entre los métodos DBSCAN, HDBSCAN y OPTICS.

Características	Algoritmos de clusterización basados en densidad		
	DBSCAN	HDBSCAN	OPTICS
Parámetros	<i>MinPts</i> y <i>Eps</i>	<i>MinPts</i>	<i>MinPts</i> y <i>Eps</i>
Distancias jerárquicas	No	Sí	Implícita
Sensibilidad a parámetros	Alta	Baja-Media	Media
Adaptabilidad a densidades variadas	Media	Alta	Baja
Métricas complementarias	No	Sí	No

Figura 3: Esquema de categorías de puntos en DBSCAN: Core, Border y Noise points.



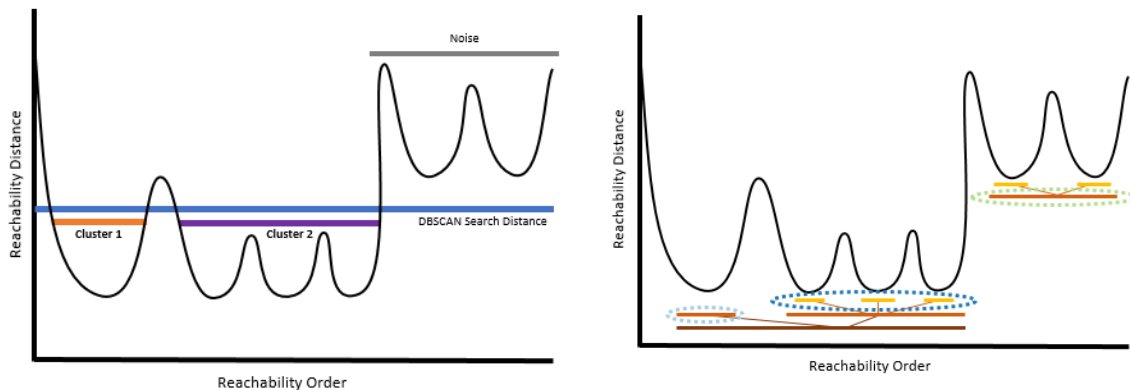
Fuente: ESRI, 2025b.

El DBSCAN permite identificar núcleos de alta concentración de puntos en un espacio geográfico (Ester et al., 1996; Sander et al., 1998; Gan & Tao, 2017), considerando heurísticos de optimización global y procedimiento de aprendizaje de máquina para

mejorar los resultados (Schubert et al., 2017). El método considera dos parámetros: radio de vecindad (*Eps*) y cantidad mínima de puntos (*MinPts*). Además, agrupa los puntos en tres categorías (Figura 3) que corresponden a: *Core* (son incluidos dentro de un cluster), *Border* (pueden estar o no incluidos en un cluster) y *Noise* (no forman parte de ningún cluster). Entre sus limitantes, se encuentra la sensibilidad a los parámetros de entrada, siendo la definición del valor de *Eps* el punto más crítico (Anselin, 2020). Ante esto, se desarrollaron variaciones para subsanar esta problemática, destacando los métodos HDBSCAN y OPTICS, que ofrecen mejoras en la flexibilidad y precisión.

El algoritmo HDBSCAN es jerárquico y permite trabajar con múltiples niveles de densidad (Figura 4), resultando adecuado para representar estructuras espaciales complejas y heterogéneas (Campello et al., 2013; Campello et al., 2015; McInnes & Healy, 2017). A diferencia de DBSCAN, construye una jerarquía de agrupamientos y selecciona automáticamente la partición más estable según un criterio de persistencia (*cluster stability*), eliminando la necesidad de definir un umbral fijo de *Eps* y utilizando únicamente el parámetro *MinPts*. Además, entrega métricas complementarias que favorece la interpretación de los resultados: el identificador de cluster, la distancia mínima al vecino más cercano (*core distance*), medida de robustez del cluster (*stability score*), probabilidad de pertenencia al cluster (*membership probability*) y probabilidad de que el punto sea considerado atípico (*outlier score* o índice de GLOSH). En el contexto del estudio, permite identificar estructuras de clusters jerárquicos, volviéndose útil para detectar distintos niveles de concentración poblacional.

Figura 4: Ilustración de comparativa del funcionamiento del DBSCAN (izquierda) y HDBSCAN (derecha)

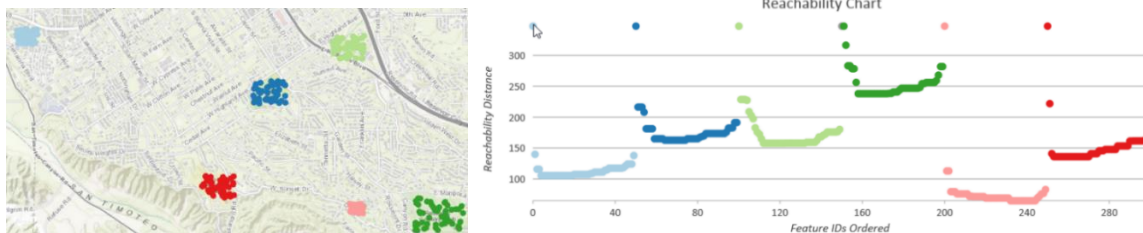


Fuente: ESRI, 2025b.

Por último, OPTICS mejora la identificación de agrupaciones en territorios con diferentes densidades, proporcionando más detalle y adaptabilidad a las condiciones específicas de cada área (Figura 5). Sin embargo, según Campello et al. (2013) y McInnes et al. (2017), no logró cerrar el problema central del DBSCAN de cómo detectar, jerarquizar y extraer automáticamente agrupamientos con densidades variables y relaciones de anidamiento reales. Por ende, se excluye del presente estudio por: la falta

de extracción automática (outputs) de clusters (que implica una interpretación caso a caso); lo implícito de la estructura jerárquica y su difícil interpretación; y porque no ofrece un criterio cuantitativo interno para decidir cuáles son los clusters más relevantes en términos de densidad, estabilidad y persistencia.

Figura 5: Ejemplo de resultados del método OPTICS.



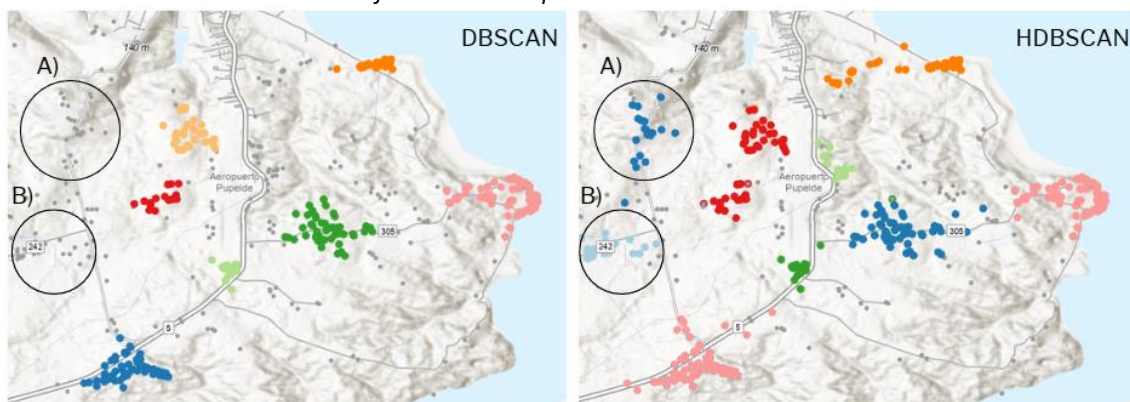
Fuente: ESRI, 2025b.

### 3.2. Marco conceptual del modelo territorial de agrupamiento

En el contexto del presente estudio, se concluye que los algoritmos DBSCAN y HDBSCAN son los óptimos para las pruebas iniciales del estudio, dada las limitaciones de OPTICS. Estos métodos permiten agrupar viviendas rurales de manera eficiente según su proximidad y densidad, adaptándose a la diversidad geográfica y demográfica de las mismas, ya que su distribución no sigue patrones predecibles. En este sentido, se consideran claves los parámetros de mínimo de puntos (*MinPts*) y distancia (*Eps*).

El parámetro *Eps* representa un umbral de distancia máxima, donde se declara un radio fijo alrededor de cada vivienda para el caso del DBSCAN, mientras que para el HDBSCAN es calculado dinámicamente. Por su parte, *minPts* debe definirse en ambos y corresponde al mínimo de viviendas dentro del radio definido por *Eps* para formar un cluster. En la Figura 6 se muestra un ejemplo comparativo de ambos métodos aplicados a viviendas rurales cercanas al Aeródromo Pupelde en Chiloé, región de Los Lagos.

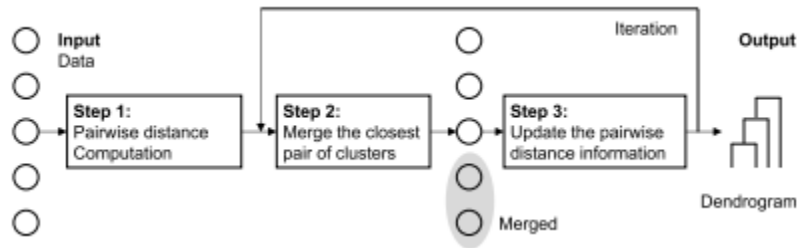
Figura 6: Ejemplo de resultados con DBSCAN (izquierda) y HDBSCAN (derecha), A) y B) son agrupaciones diferentes al comparar ambos métodos.



En ambos métodos se define un mínimo de viviendas de 15 y en el caso del DBSCAN, una distancia de búsqueda de 250 m. El ejemplo demuestra que los resultados de DBSCAN son más restrictivos respecto a los clusters obtenidos ya que, al ser la distancia fija de 250 m, son menores las agrupaciones de viviendas que cumplen con ambos requisitos. En cambio, el HDBSCAN permite identificar nuevos clusters (A y B) que con DBSCAN eran catalogados como ruido (*noise*), ya que la distancia de búsqueda es dinámica y depende de la jerarquización alcanzada.

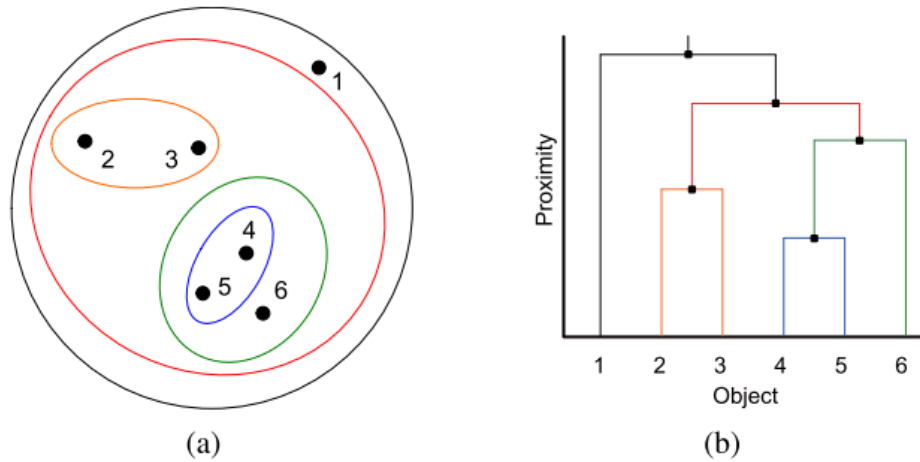
Específicamente, como se mencionó anteriormente, el HDBSCAN presenta la particularidad de ser jerárquico, esto se ejemplifica en la Figura 7. En general, identifica estructuras jerárquicas que aparecen en los datos de entrada, a partir de múltiples iteraciones, donde cada una debe calcular y actualizar las distancias por pares entre todos los clústeres intermedios. El método se basa en cinco pasos: 1) Calcular la distancia por pares, 2) Fusionar el par de grupos más cercanos, 3) Actualizar la información de distancia por pares, 4) Iteración del paso dos y 5) Obtención del dendograma final (Jeon et al., 2017).

Figura 7: Etapas de un método de clusterización jerárquico. Fuente: Jeon et al. (2017).



El dendograma (Figura 8) es uno de los productos principales que entrega este algoritmo y representa la información sobre los clústeres individuales. Cada elemento del dendograma corresponde a un clúster individual y un nodo interno representa la unión de dos clústeres (uno con mayor umbral y el otro con un menor umbral). La longitud de un nodo indica la distancia o disimilitud a la que se produce una unión. En otras palabras, el dendograma muestra un umbral de proximidad alto que disminuye gradualmente y, a medida que disminuye, se generan subdivisiones, denotando una jerarquía de componentes conectados (desde completamente conectados hasta desconectados), con diferentes niveles de umbral (Jeon et al., 2017).

Figura 8: Ejemplo de funcionamiento de HDBSCAN, a) jerarquización hipotética de clústeres en seis puntos, y b) dendograma correspondiente donde cada unión es representada por un punto. Fuente: Jeon et al. (2017).



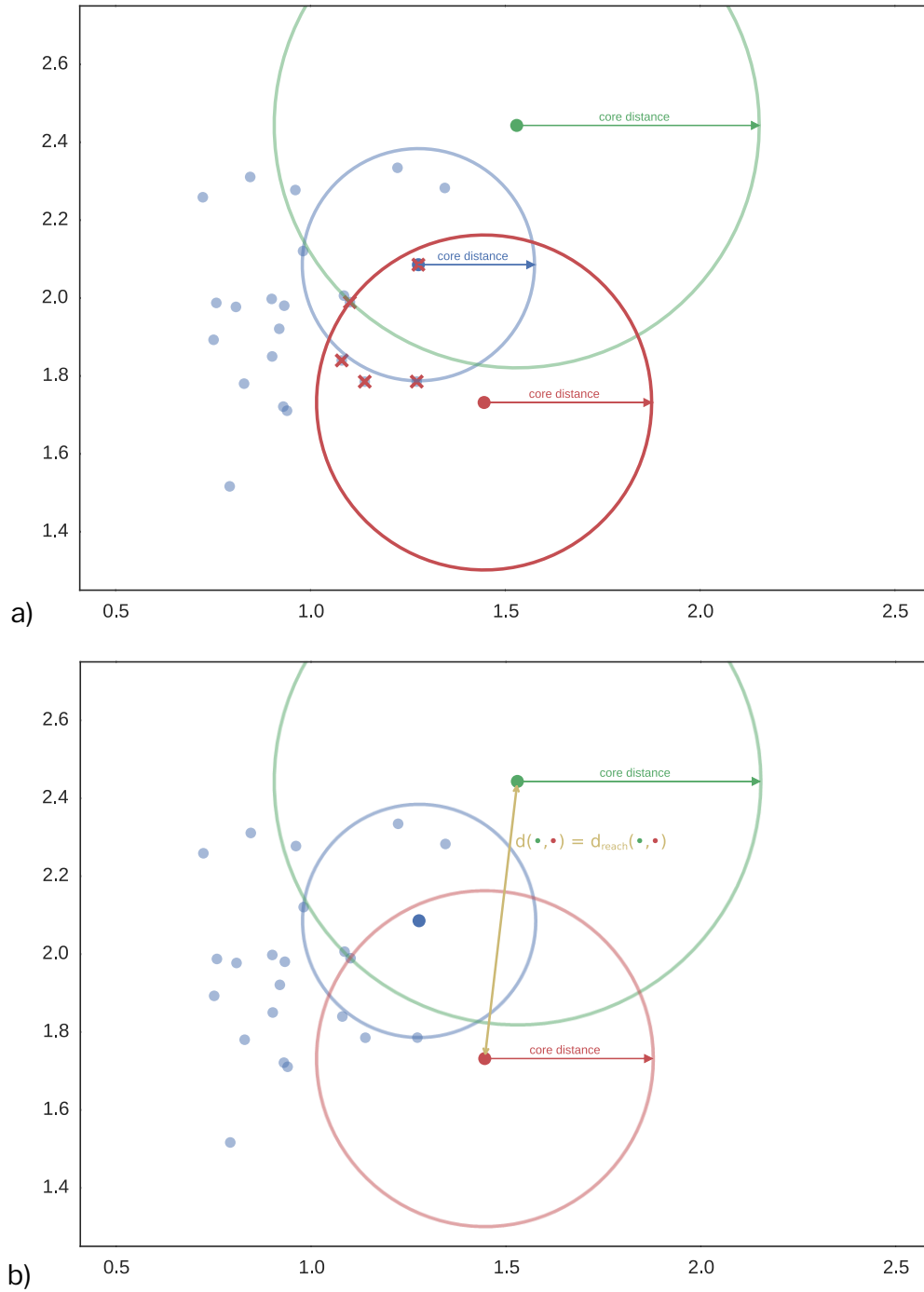
Específicamente en la Figura 8, se logra observar la concordancia entre el dendograma (b) y la agrupación espacializada (a). El primer elemento del dendograma (color negro), corresponde a la agrupación de mayor tamaño que incorpora los 6 puntos. Luego, al disminuir el umbral (proximidad), se obtiene una subdivisión del primer clúster, resultando en la agrupación de cinco puntos, sin considerar el punto n°1 (color rojo), al no cumplir con el umbral de distancia. El proceso continúa, hasta no encontrar más subdivisiones posibles según los umbrales. En este sentido, la última agrupación identificada corresponde al último clúster identificado con el menor umbral posible, resultando en la unión de los puntos n°4 y 5 (color azul).

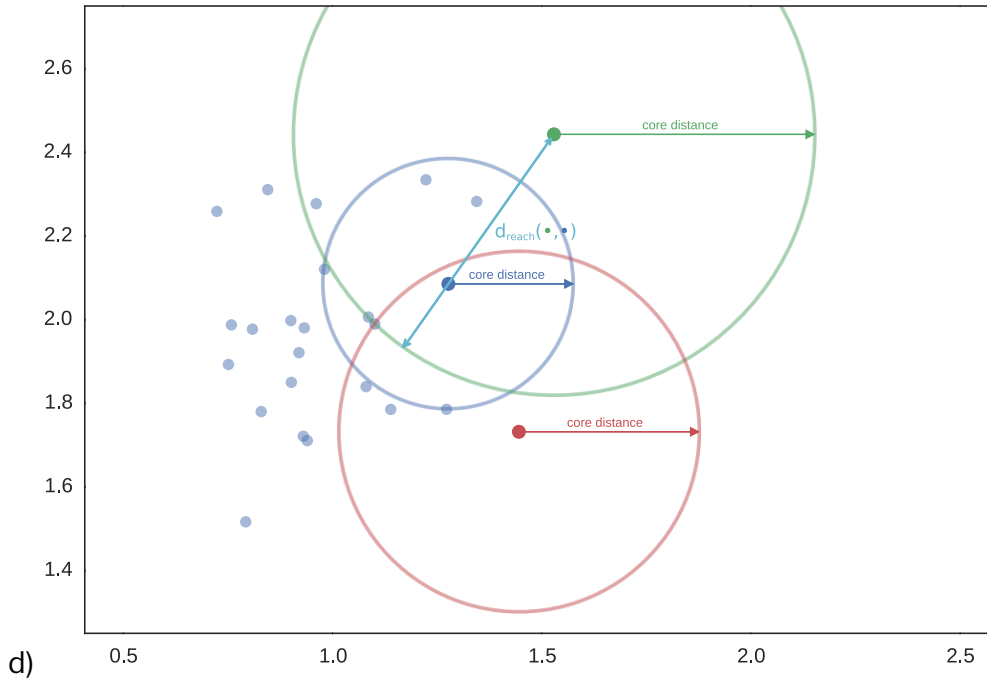
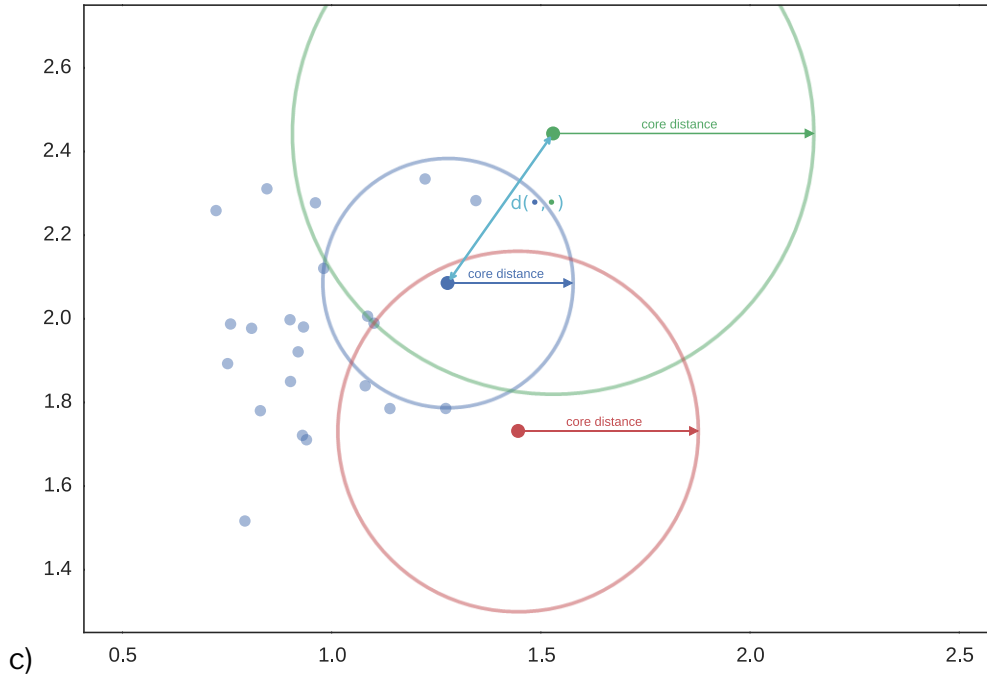
Una métrica que explica estas subdivisiones corresponde la distancia de alcance mutua (Mutual Reachability Distance), la cual indica el menor umbral necesario para conectar dos puntos Figura 9. En el caso de los puntos con alta densidad (distancia central baja) se mantienen a la misma distancia entre sí, pero los puntos más dispersos, con baja densidad (distancia central alta), se alejan, para estar al menos a la distancia central de cualquier otro punto (Figura 8, b). Resultando fundamental este indicador para separar puntos con baja densidad.

Figura 9: Cálculo de la distancia de alcance mutua. Fuente: McInnes et al. (2017).

$$d_{\text{mreach-}k}(a, b) = \max\{\text{core}_k(a), \text{core}_k(b), d(a, b)\}$$

Figura 10: Ejemplo de las diferentes distancias centrales entre tres puntos. Cada figura representa: a) Distancias centrales (Core distance) de tres puntos, b) Distancia de alcance mutua (Mutual Reachability Distance) entre punto verde y rojo, y c) y d) La distancia de alcance mutua entre punto verde y azul. Fuente: McInnes et al. (2017).





En la Figura 10 (a) se observan tres puntos con diferentes distancias centrales (*Core distance*): la mayor corresponde al punto verde y la menor al azul. En el primer caso, al unir el punto rojo con el verde, la distancia de alcance mutua supera la distancia central en ambos núcleos, atravesando ambos círculos (Figura 10, b). En cambio, en el segundo caso, de unir el punto verde con el azul, la distancia central del punto verde supera a la distancia entre ellos (Figura 10, c), resultando en una distancia de alcance mutua mayor, siendo el equivalente del radio del verde (Figura 10, d).

Entre los otros resultados que se obtienen al aplicar el método HDBSCAN, destacan cinco indicadores que complementan los resultados descritos en la Tabla 4. Si bien el método permite obtener otras métricas, estas destacan por ser útiles para filtrar y analizar los clústeres obtenidos. Además, se incorporan a la tabla de atributos de la capa resultante de los puntos agrupados, facilitando su uso en análisis geoespaciales posteriores.

Tabla 4: Indicadores resultantes del algoritmo HDBSCAN.

Indicadores HDBSCAN	Descripción
<i>Cluster</i>	Identificador del cluster
<i>Core distance</i>	Distancia mínima al vecino más cercano
<i>Stability score</i>	Puntaje de calidad del cluster. Proxy de robustez.
<i>Membership_prob</i>	Probabilidad de pertenencia al cluster asignado (0 a 1)
<i>Outlier_scores</i>	Probabilidad de ser un valor outlier (atípico) (0 a 1)

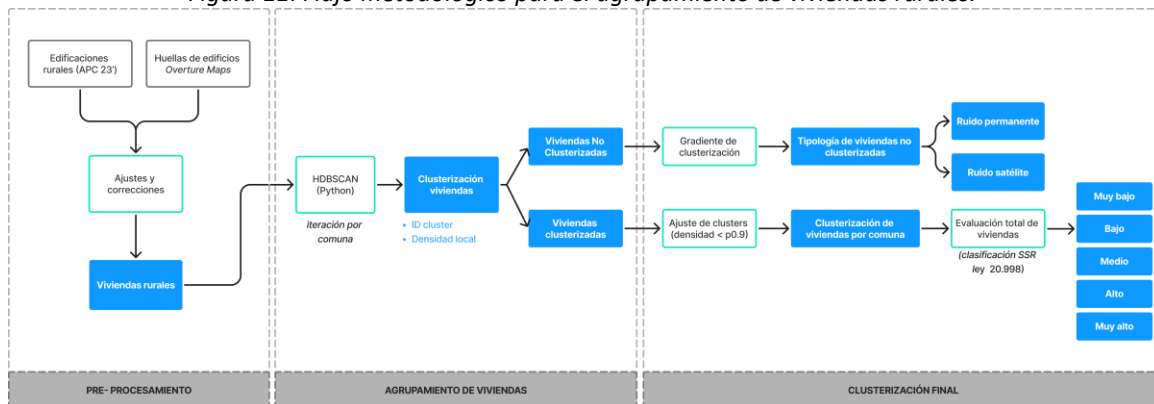
## 4. METODOLOGÍA DEL MODELO DE AGRUPAMIENTO

En esta sección se describen los pasos metodológicos para la elaboración del Modelo de Agrupamiento de viviendas rurales (Figura 11), además de la elaboración del Modelo territorial para la focalización de infraestructura (Figura 12). En ambos casos se entrega un contexto general de los procedimientos, actividades y productos que se desarrollaron.

### 4.1. Metodología

Para llevar a cabo la clusterización de viviendas rurales mediante el algoritmo HDBSCAN se siguió un flujo metodológico definido en la Figura 11, el cual está dividido en tres grandes etapas: Preprocesamiento de datos, Agrupamiento de viviendas y Clusterización final.

Figura 11: Flujo metodológico para el agrupamiento de viviendas rurales.



### 4.2. Preprocesamiento – Consolidación de viviendas en aldeas

El cálculo del agrupamiento de viviendas rurales se realizó a partir de los Puntos de Edificación Rural de la Actualización Cartográfica Precensal 2023 (APC23). Cada punto representa una edificación clasificada como vivienda particular, vivienda colectiva, otro uso o edificación, esta última categoría se utiliza para aquellas estructuras cuyo uso no pudo definirse con certeza.

Una limitación de este insumo es que en zonas rurales concentradas y semiconcentradas las viviendas no se registran como puntos, sino como Aldeas INE (geometría de polígono). Estas entidades corresponden a núcleos rurales de entre 301 y 2.000 habitantes, generalmente con amanzanamiento y continuidad de viviendas en torno a una vía estructurante. Sus polígonos se desagregan en Manzanas Aldeas, que representan las unidades básicas del asentamiento. Este criterio de representación provoca la omisión de las viviendas incluidas en aldeas dentro del Modelo de Agrupamiento, lo que genera un sesgo en el análisis. Como se observa en Figura 13 los puntos de edificación aparecen dispersos alrededor del polígono de la aldea, cuando en realidad forman parte de un mismo patrón de poblamiento.

Figura 12: Flujo metodológico del Modelo Territorial para la priorización de infraestructura.

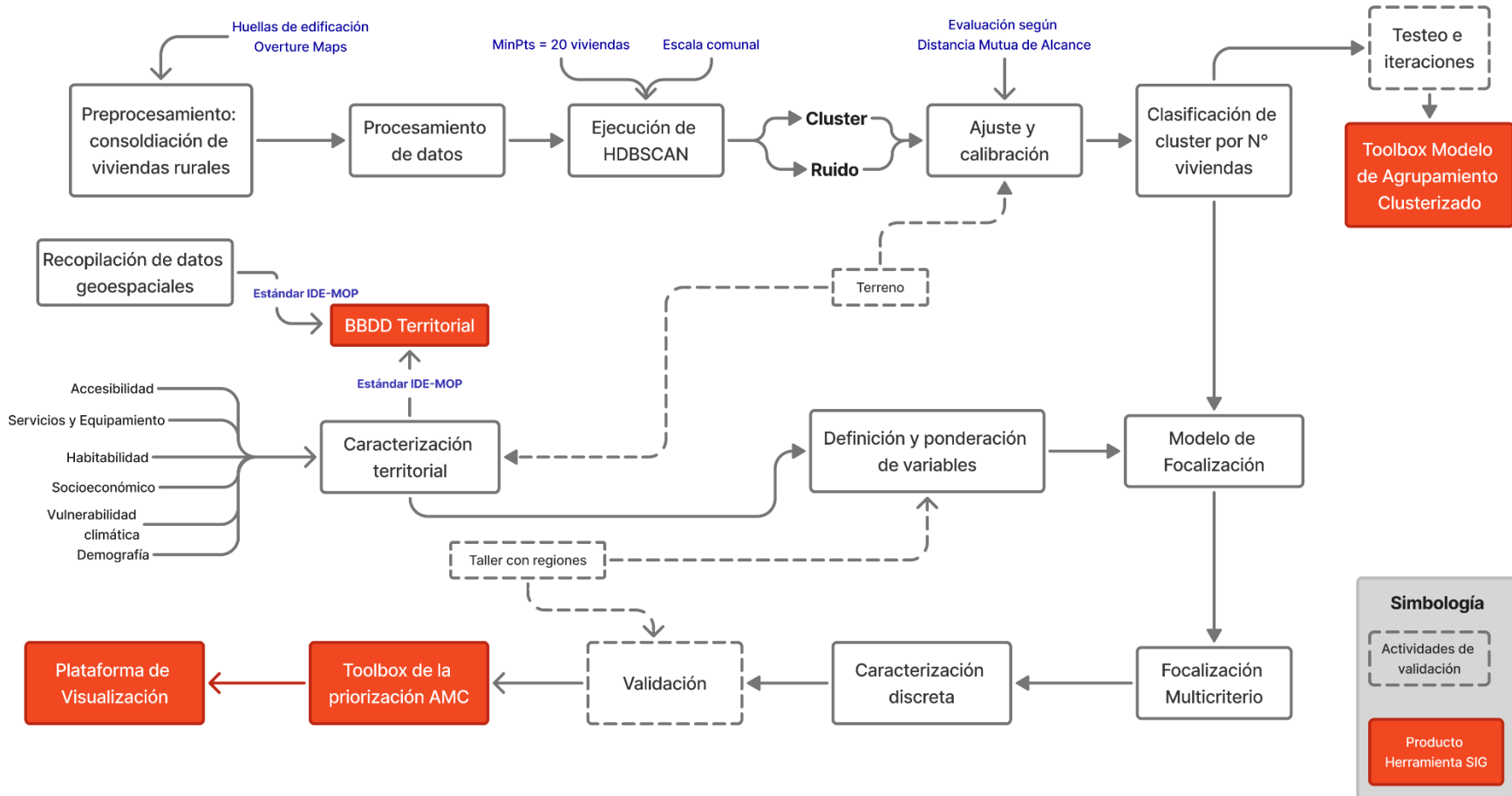


Figura 13: Visualización de puntos edificación rural y el polígono de Aldea.



La problemática también fue levantada<sup>2</sup> por analistas territoriales regionales de DIRPLAN, quienes plantearon que esta diferencia de registro entre puntos y polígonos condiciona la representatividad y precisión del Modelo de Agrupamiento.

Con el fin de abordar esta problemática, se trabajó con el conjunto de datos de Overture Maps, que compila y prioriza huellas de edificios de diversas fuentes, en donde predominan las provenientes de OpenStreetMap, ESRI Community Maps, Google Open Buildings, y Microsoft. El conjunto de datos de las huellas de edificio se obtuvo a través de Python siguiendo la documentación provista por la fuente.

A continuación, en la Tabla 5 se desagrega la información consolidada.

<sup>2</sup> La problemática fue advertida en el trabajo de terreno y en las observaciones realizadas por las regiones

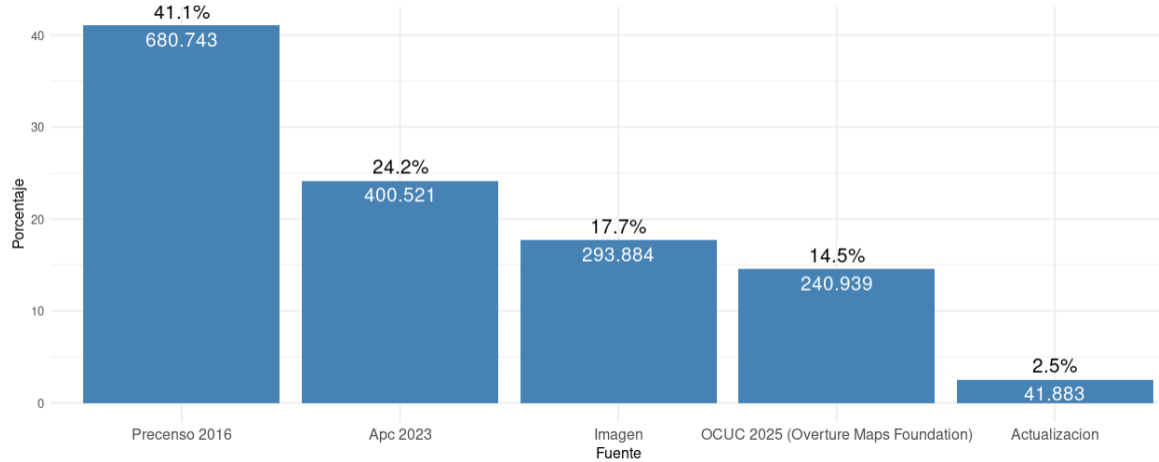
Tabla 5: Detalle de los puntos y fuentes de origen.

Nombre	Aldeas (polígono)	Puntos Overture	Viviendas Rurales catastradas en Aldeas	Viviendas para eliminar	Total de viviendas (puntos) de Aldeas
Arica y Parinacota	10	11.698	3.448	49	3.399
Tarapacá	22	14.576	9.431	614	8.817
Antofagasta	19	12.579	7.214	969	6.245
Atacama	31	27.369	15.967	294	15.673
Coquimbo	94	57.954	29.501	1.541	27.960
Valparaíso	100	64.043	30.110	1.179	28.931
Metropolitana	98	45.780	24.058	1.078	22.980
O'Higgins	145	80.886	38.410	481	37.929
Maule	145	74.816	36.297	961	35.336
Nuble	35	20.564	9.845	115	9.730
Biobío	50	25.280	13.068	456	12.612
Araucanía	38	19.973	10.729	157	10.572
Los Ríos	24	16.312	7.597	32	7.565
Los Lagos	26	16.241	8.037	161	7.876
Aysén	14	8.389	4.281	100	4.181
Magallanes	7	2.036	1.168	35	1.133
<b>Total país</b>	<b>858</b>	<b>498.496</b>	<b>249.161</b>	<b>8.222</b>	<b>240.939</b>

De este modo, con la incorporación de las viviendas rurales catastradas en las Aldeas INE (polígono), la columna “Fuente” de la capa de edificaciones rurales queda con cinco categorías de fuente: las que provenían de la capa de origen (Precenso 2016, APC 2023, Imagen y Actualización) y la nueva llamada “OCUC 2025 (Overture Maps)” (Figura 12). De este modo, la cobertura queda disponible para que pueda ser utilizada para los fines necesarios de cada región. La Figura 14 presenta la distribución total de los puntos de edificaciones rurales según la fuente de origen de los datos.

Durante la revisión de los datos obtenidos, se identificó una diferencia entre la cantidad de viviendas catastradas en las Aldeas INE y el número de puntos generados a partir de las huellas, los cuales estaban duplicados. Esta inconsistencia pudo detectarse gracias a la información disponible sobre el número de viviendas por Aldea Rural según el APC 2023. Finalmente, al comparar las huellas extraídas con las viviendas duplicadas (errores), fue posible ajustar los datos y obtener el total de viviendas rurales corregidas a partir del proceso de extracción mediante Overturemaps.

Figura 14: Distribución de edificaciones rurales en todas las regiones según su fuente en porcentaje (se incorpora por fuente el número total de viviendas).

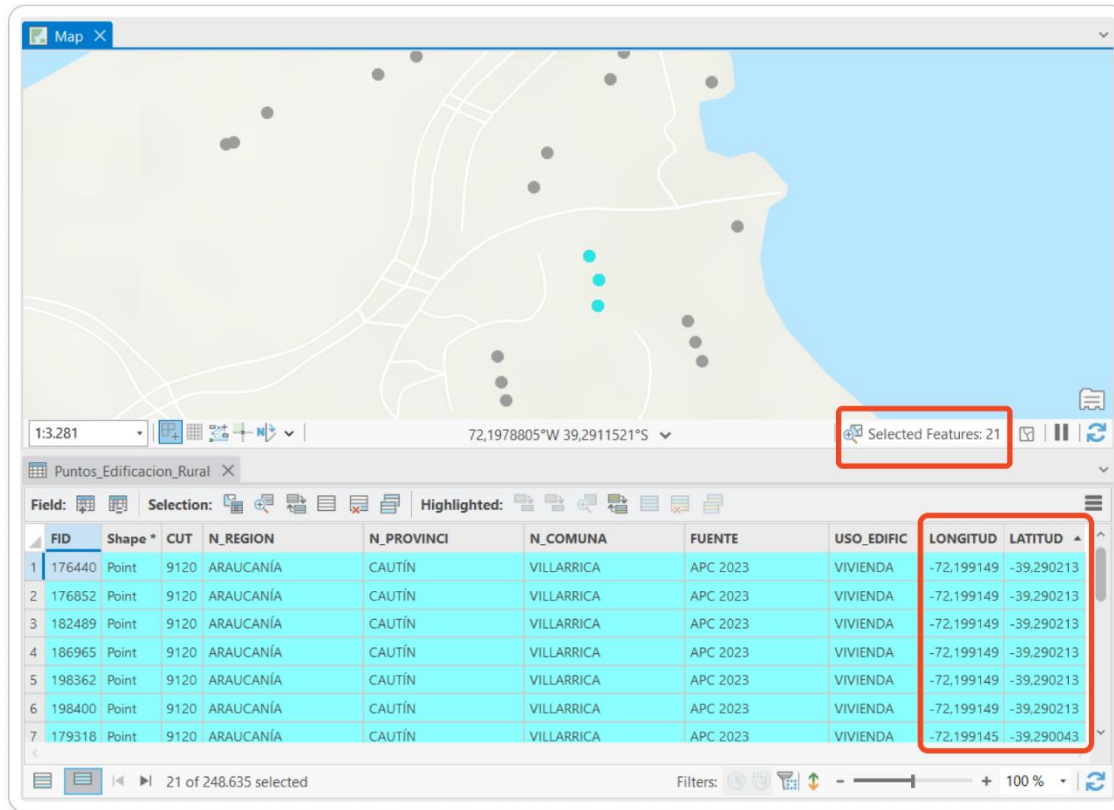


El dataset resultante se filtró conservando únicamente los puntos contenidos dentro de los polígonos de las aldeas, mediante un proceso de superposición espacial en R. En la mayoría de los casos, el conteo de puntos superó al número de viviendas reportado por el polígono de la aldea. Para corregir esta diferencia, se aplicó un muestreo aleatorio, seleccionando un subconjunto de puntos equivalente al número de viviendas registradas en la base INE.

Una vez incorporados los puntos de viviendas rurales provenientes de las Aldeas INE, fue necesario realizar un segundo ajuste a la base de datos. Durante los análisis de clusterización con HDBSCAN (MinPts = 20), se detectó un error en los puntos de viviendas rurales, relacionado a la presencia de registros duplicados (superpuestos) originados en la capa de Edificaciones Rurales APC23 del INE. Aunque el error corresponde a 2.325 edificaciones duplicadas —equivalente al 0,16% del total de registros—, su impacto es relevante al aplicar métodos basados en distancias entre puntos, ya que provoca la formación de clústeres erróneos, distancias mal calculadas y, en consecuencia, una clasificación poco precisa de los agrupamientos rurales.

Lo anterior se ejemplifica en la Figura 15, que muestra un caso de la región de La Araucanía: en el shapefile de Edificaciones Rurales APC23, al seleccionar tres puntos en ArcGIS, la tabla de atributos despliega 21 registros, todos con coordenadas de latitud y longitud idénticas.

Figura 15: Ejemplo de duplicidad de edificaciones rurales APC 2023 INE, región de la Araucanía.



A partir de esta inconsistencia detectada en los datos de la APC23 provenientes del INE, fue necesario revisar el resto de las regiones para asegurar que no existieran edificaciones duplicadas. Para ello, mediante RStudio, se identificaron los registros duplicados por región utilizando dos métodos complementarios: (i) detección de registros shape con coordenadas idénticas a nivel de atributos y, como verificación, (ii) identificación geoespacial de puntos superpuestos. Los resultados muestran que el error fue inferior al 1% en todas las regiones; en dos de ellas no se registraron errores, mientras que la RM, Tarapacá y La Araucanía presentaron los mayores porcentajes de viviendas duplicadas, lo que se resume en la Tabla 6.

Tabla 6: Síntesis de edificaciones duplicadas (error) y cantidad final de edificaciones a utilizar (corregidas).

Región	Edificaciones Rurales APC23 - INE	Edificaciones Rurales Duplicadas APC23	Proporción del error de duplicidad	Edificaciones Rurales APC23 - Corregidas
Arica y Parinacota	12.554	0	0%	12.554
Tarapacá	12.803	33	0,26%	12.770
Antofagasta	11.713	0	0%	11.713
Atacama	24.418	41	0,17%	24.377
Coquimbo	85.289	54	0,06%	85.235
Valparaíso	107.819	782	0,73%	107.037
Metropolitana	125.422	1.014	0,81%	124.408
O'Higgins	104.861	158	0,15%	104.703
Maule	146.362	7	0%	146.355
Ñuble	106.100	6	0,01%	106.094
Biobío	144.343	52	0,04%	144.291
Araucanía	248.635	542	0,22%	248.093
Los Ríos	72.180	37	0,05%	72.143
Los Lagos	179.516	271	0,15%	179.245
Aysén	21.262	14	0,07%	21.248
Magallanes	16.772	7	0,04%	16.765
<b>Total país</b>	<b>1.420.049</b>	<b>3.018</b>	-	<b>1.417.031</b>

Una vez corregidas las Edificaciones Rurales APC23, se consolidó la base de datos de viviendas que serán analizadas, para lo cual se sumaron las viviendas obtenidas de las Aldeas Rurales INE (240.939) y las Edificaciones Rurales APC23 corregidas (1.417.031). El detalle del total de viviendas rurales final se presenta en la Tabla 7 a nivel regional.

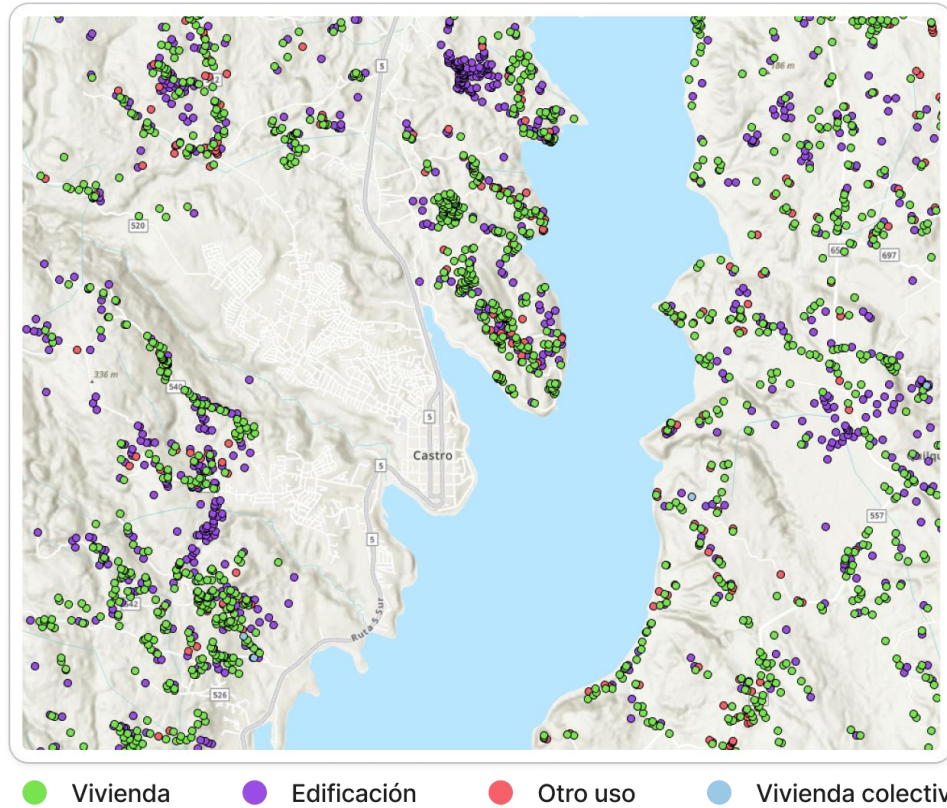
Tabla 7: Consolidación de base final de viviendas rurales 2023.

Región	Edificaciones Rurales APC23 - Corregidas	Viviendas rurales obtenidas de huellas Aldeas Overture - ajustadas	Total de viviendas rurales (versión final)
Arica y Parinacota	12.554	3.399	15.953
Tarapacá	12.770	8.817	21.587
Antofagasta	11.713	6.245	17.958
Atacama	24.377	15.673	40.050
Coquimbo	85.235	27.960	113.195
Valparaíso	107.037	28.931	135.968
Metropolitana	124.408	22.980	147.388
O'Higgins	104.703	37.929	142.632
Maule	146.355	35.336	181.691
Ñuble	106.094	9.730	115.824
Biobío	144.291	12.612	156.903
Araucanía	248.093	10.572	258.665
Los Ríos	72.143	7.565	79.708
Los Lagos	179.245	7.876	187.121
Aysén	21.248	4.181	25.429
Magallanes	16.765	1.133	17.898
<b>Total país</b>	<b>1.417.031</b>	<b>240.939</b>	<b>1.657.970</b>

### 4.3. Procesamiento – Tratamiento de Edificaciones Rurales APC23

Una vez que se consolidó la capa de edificaciones rurales incluyendo las viviendas existentes en Aldeas y corregidas las edificaciones duplicadas, se procedió a ajustar las edificaciones originales provenientes de la actualización precensal 2023. Esta capa de geometría de tipo punto contiene la localización georreferenciada de lo que el INE considera como edificaciones rurales, las que se categorizan en cuatro tipos: edificación, otro uso, vivienda y vivienda colectiva. A modo de ejemplo, la Figura 16 muestra la distribución de las edificaciones rurales por tipo en el sector de Ancud, Región de Los Lagos.

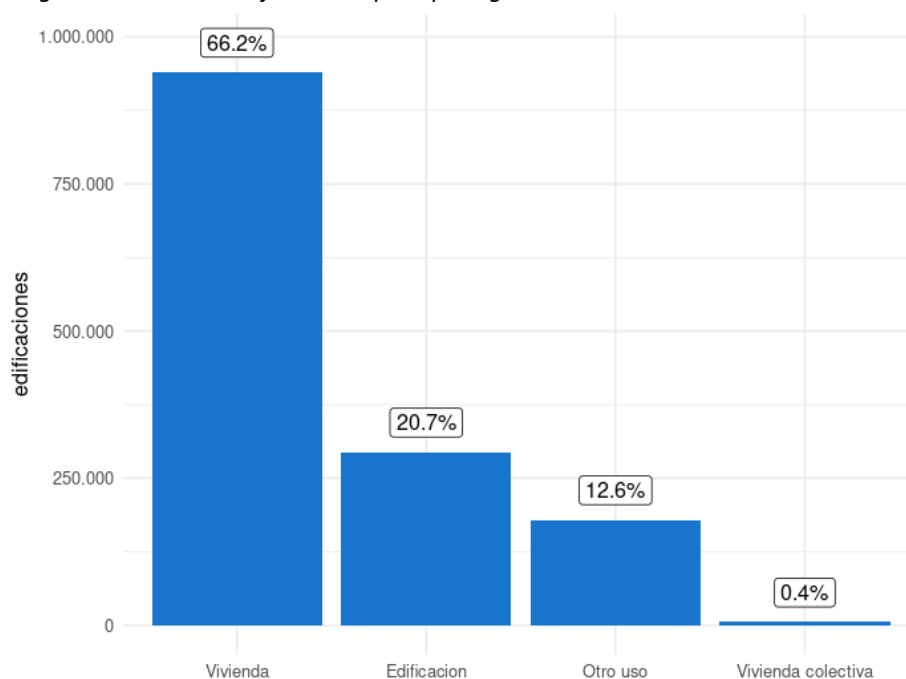
Figura 16: Vista de los tipos o categorías de edificaciones rurales de la región de Los Lagos, sector de Castro.



Al contabilizar el total de viviendas por tipo a nivel nacional (Figura 17), si bien los puntos que corresponden a viviendas rurales propiamente tal representan el 66,6%, se optó por incluir las categorías de Edificación y Otro uso para realizar el análisis de cluster con HDBSCAN.<sup>3</sup> Esta decisión responde a la incompatibilidad de esta capa de información con su homóloga del Precenso 2016, lo que en la práctica significa que el universo total de viviendas disminuyera en este periodo de tiempo.

<sup>3</sup> El análisis de cluster con HDBSCAN se llevó a cabo utilizando el 100% de los puntos provenientes de la actualización pre-censal 2023, sin hacer distinción entre el tipo de edificación.

Figura 17: Total de edificaciones por tipo según Actualización Pre Censal 2023 INE.



Una vez descargadas las capas por región de las edificaciones rurales (desde ahora en adelante llamadas viviendas rurales) se procedió a realizar una serie de ajustes para su uso posterior. Estas correcciones fueron necesarias de realizar para que el algoritmo HDBSCAN pudiera clusterizar las viviendas rurales de forma correcta, y evitar así cualquier problema en el procesamiento de la información.

El primer cambio que se realizó fue una reproyección cartográfica de las viviendas rurales a un sistema de coordenadas proyectadas, específicamente al UTM. Debido a que HDBSCAN calcula una matriz de distancias para identificar las agrupaciones de viviendas, es necesario que las coordenadas no estén en un sistema de coordenadas geográficas para que las unidades de medidas de grados sexagesimales no interfieran en el cálculo de estas distancias.

De las 16 capas de viviendas rurales aquellas pertenecientes a las regiones del Biobío, Los Ríos, Los Lagos, Aysén y Magallanes fueron reproyectadas en el huso 18 Sur, mientras que el resto de las regiones fueron transformadas al huso 19 Sur.

Otro de los ajustes que se realizó fue la obtención del par de coordenadas XY para las viviendas rurales, las cuales se almacenaron en campos o columnas independientes dentro de las capas de cada región. En concreto, esto significó que la coordenada UTM Este se guardó en una columna llamada "POINT\_X", mientras que la coordenada UTM Norte se almacenó en otra columna llamada "POINT\_Y" (Figura 18).

Dentro de esta etapa del preprocesamiento este ajuste adquiere especial importancia ya que serán estos campos o columnas los que serán utilizados como variables de agrupamiento por el algoritmo HDBSCAN. En consecuencia, cualquier error en el cálculo y creación de estas columnas afectará directamente al resultado que se obtenga de la clusterización de viviendas

Figura 18: Vista de la tabla de atributos de la capa de viviendas rurales de la región de los Lagos, con énfasis en las columnas POINT\_X, POINT\_Y, POINT\_Z y POINT\_D.

	N_PROVINCIA	N_COMUNA	N_DISTRITO	TIPO_DISTRITO	USO_EDIFICACION	POINT_X	POINT_Y	POINT_Z	POINT_D
1	LLANQUIHUE	PUERTO MONTT	PANITAO	MIXTO	VIVIENDA	662483,1406	5399054,9963	96	82,521243
2	LLANQUIHUE	PUERTO MONTT	LA PALOMA	MIXTO	VIVIENDA	674324,6027	5411169,5697	107	31,302176
3	LLANQUIHUE	PUERTO MONTT	ALERCE	MIXTO	VIVIENDA	678431,825	5416015,3796	147	30,741441
4	LLANQUIHUE	PUERTO MONTT	MIRASOL	MIXTO	OTRO USO	665355,2289	5404676,9371	99	43,293471
5	LLANQUIHUE	PUERTO MONTT	ALERCE	MIXTO	VIVIENDA	679009,877	5415638,8419	160	15,381631
6	LLANQUIHUE	PUERTO MONTT	ALERCE	MIXTO	VIVIENDA	679253,8279	5410821,1093	149	11,577142
7	LLANQUIHUE	PUERTO MONTT	PELLUCO	MIXTO	VIVIENDA	677991,6305	5405086,7712	39	29,608168
8	LLANQUIHUE	PUERTO MONTT	ALERCE	MIXTO	VIVIENDA	678756,7584	5413727,7771	143	23,28733
9	LLANQUIHUE	PUERTO MONTT	ALERCE	MIXTO	OTRO USO	679052,4145	5417502,154	149	27,907094
10	LLANQUIHUE	PUERTO MONTT	PELLUCO	MIXTO	VIVIENDA	679521,6124	5407019,7301	114	61,884291

Junto con las coordenadas UTM X e Y también se agregaron a las capas de viviendas rurales dos variables complementarias, las que corresponden a la elevación o altitud de cada vivienda, y la distancia de cada vivienda a la red vial más cercana. El propósito original de ambas variables (almacenadas en columnas llamadas “POINT\_Z” y “POINT\_D” respectivamente) era ser utilizadas junto a las coordenadas XY como variables de agrupación, de forma tal que los clusters agrupen viviendas con elevaciones y distancias a caminos similares.

Sin embargo, luego de realizar varias pruebas se concluyó que estas variables complementarias no generaban resultados que fueran significativos ya que existían casos en que la densidad de viviendas no era el criterio principal de agrupamiento. De todas formas, se optó por no eliminar estas variables de las capas de viviendas para poder utilizarlas al momento de caracterizar el comportamiento de cada agrupación.

#### 4.4. Implementación del agrupamiento de viviendas

Una vez ajustadas y corregidas las capas de viviendas rurales, se procedió a ejecutar el algoritmo HDBSCAN para identificar los agrupamientos en cada comuna. Tal como se señaló en el trabajo en terreno, uno de los principales hallazgos es que valores menores del parámetro MinPts permiten una mejor detección de pequeñas densificaciones. En este contexto, un ajuste clave en esta entrega fue la reducción del parámetro MinPts de

50 a 20 viviendas, con el fin de mejorar la representatividad de los patrones de poblamiento observados.

Otro ajuste relevante del Modelo de Agrupamiento corresponde a la escala territorial de aplicación del algoritmo HDBSCAN. En esta entrega, el universo de análisis se definió a nivel comunal y no regional, lo que implica que cada vivienda se agrupa considerando únicamente a las demás de su comuna. De este modo, el algoritmo se ejecuta de manera independiente en las 312<sup>4</sup> comunas del país, generando un dendrograma y clusters para cada unidad administrativa. Este cambio mejora la pertenencia territorial de los resultados, al reflejar con mayor precisión la dinámica local de los asentamientos. Además, el uso de la escala comunal tiene la ventaja de incorporar, de manera implícita, las barreras naturales que suelen coincidir con los límites comunales —como ríos, quebradas o líneas de altas cumbres—, lo que reduce las distorsiones que se presentaban en la versión regional del modelo.

Considerando que la complejidad del algoritmo HDBSCAN es bastante mayor en comparación a DBSCAN y a OPTICS, principalmente por utilizar varias distancias de búsqueda, se optó por ejecutar este modelo utilizando herramientas de programación de Python. Esto permitió iterar el uso de este algoritmo comuna por comuna optimizando el tiempo de ejecución de este en comparación con las herramientas de caja de ArcGIS Pro. Independiente de esto, y debido a la compatibilidad entre Python y ArcGIS Pro, es importante mencionar que el Toolbox final que se entregará como resultado de este proyecto podrá sin problemas implementar el script de Python desarrollado en esta etapa. En síntesis, para clusterizar las viviendas rurales por región con HDBSCAN se configuraron cuatro parámetros principales de este algoritmo, los que se resumen en la Tabla 8.

---

<sup>4</sup> Corresponden al total de comunas que cuentan con viviendas rurales digitalizadas. Las 33 comunas restantes son solamente urbanas, por lo que no cuentan con viviendas rurales.

Tabla 8: Parámetros de ajuste para el uso de HDBSCAN.

Parámetro	Descripción	Valor
<b>Input</b>	Capa georreferenciada de puntos con viviendas rurales a clusterizar	Capa de viviendas rurales
<b>Escala geográfica</b>	Define el área en que el HDBSCAN calcula las distancias entre viviendas, determinando si la agrupación se analiza a nivel regional, provincial, comunal u otra unidad territorial.	Comunal
<b>Variables agrupamiento</b>	Campos dentro de la tabla de atributos del parámetro Input respecto a los cuales se crearán los clusters sobre el parámetro Input	POINT_X y POINT_Y
<b>Mínimo de puntos (MinPts)</b>	Número mínimo de viviendas por cluster	20

De estos parámetros hay dos que son claves para que obtener resultados que sean robustos y confiables para los objetivos de este estudio. El primero de estos corresponde a las variables de agrupamiento que en este caso son los campos POINT\_X, y POINT\_Y de cada vivienda.

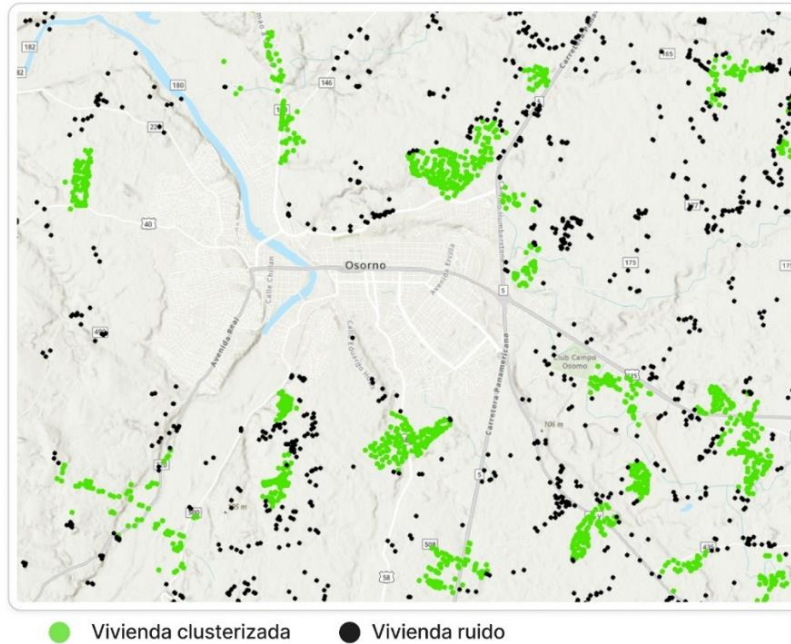
Utilizar las columnas con las coordenadas XY como variables de agrupamiento permite que HDBSCAN agrupe o genere clusters tomando como criterio principal la densidad geográfica de viviendas, es decir, lugares con mayor concentración de puntos serán catalogados como posibles candidatos a conformar un clúster.

El segundo de los parámetros clave corresponde al número mínimo de viviendas por clúster o MinPts, que en la práctica restringe el resultado de los clusters candidatos y elimina todos aquellos que tengan un total de viviendas menor al definido en este parámetro.

El resultado de la ejecución de HDBSCAN iterando para cada comuna fue posteriormente consolidado en una sola capa vectorial llamada “viviendas\_clusters\_hdbs”. Esta capa vectorial contiene todas las viviendas rurales del país que fueron utilizadas como input para HDBSCAN, incluyendo los resultados de este algoritmo tales como: el identificador de cada cluster por región, la probabilidad de pertenencia por cluster, el puntaje outlier, entre otros (Figura 19).

De forma complementaria, para mejorar la visualización e interpretación de los clusters de viviendas se creó una capa de geometría de tipo polígono, llamada “poligonos\_clusters\_hdbs”, que representa la extensión o límites aproximados de cada agrupamiento de viviendas. El detalle de los atributos que contienen ambas capas se encuentra en el apartado “Anexo 3. Diccionario de datos”.

Figura 19: Vista de viviendas clusterizadas y no clusterizadas por HDBSCAN en el sector de Osorno.



#### 4.5. Calibración y filtros finales

Junto con identificar de forma única por comuna a cada grupo o cluster de viviendas, otro resultado del algoritmo HDBSCAN es la distancia mutua de alcance la cual entrega una métrica clave para cuantificar el nivel de agrupamiento o densidad por vivienda.

La distancia mutua de alcance (o mrd) representa el grado de agrupamiento de cada vivienda respecto a los puntos que se encuentran en su proximidad, lo que entrega información respecto a la densidad de viviendas de forma local. Esta métrica está construida sobre la base de la distancia núcleo que corresponde a la distancia que tiene que recorrer un punto A para alcanzar un mínimo de puntos o MinPts.

$$d_{mr}(A, B) = \max[d_{coreA}, d_{coreB}, d_{AB}]$$

Donde:

- A: punto o vivienda A.
- B: punto o vivienda B.
- dmr: distancia de alcance mutua (en metros)
- dcoreA: distancia núcleo A
- dcoreB: distancia núcleo B
- dAB: distancia euclidiana entre punto A y punto B

De esta forma, aquellas viviendas que tengan una baja distancia mutua de alcance representan territorios con una alta densidad de puntos, es decir, necesitan recorrer una baja distancia para alcanzar un mínimo de puntos definido en el algoritmo HDBSCAN. En cambio, cuando la distancia mutua de alcance es alta, representa casos donde hay una mayor dispersión de viviendas o baja densidad de viviendas, lo que causa que sea necesario recorrer una mayor distancia para alcanzar el mínimo de puntos.

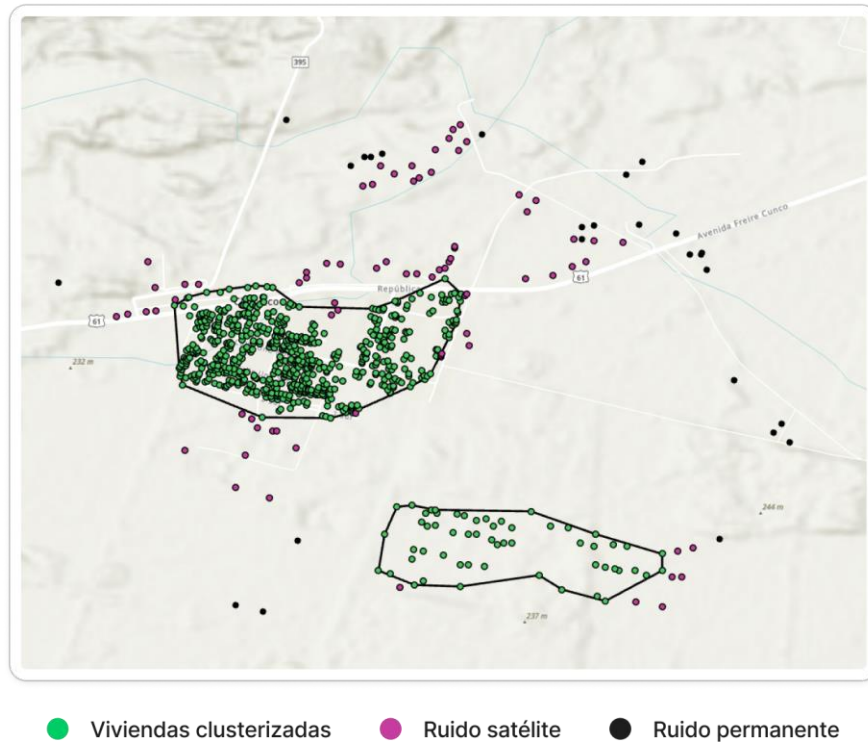
Al considerar no solo la cercanía entre viviendas, sino también la densidad del entorno en que se ubican, este indicador ayuda a filtrar agrupaciones poco representativas y ajustar los clusters hacia configuraciones más coherentes con la realidad territorial.

Producto de esto, y de acuerdo con el flujo metodológico de la Figura 11 con el objetivo de que los clusters agrupen viviendas que compartan una densidad local similar, se procedió a filtrar aquellos puntos dentro de cada cluster que tuvieran una distancia mutua de alcance que estadísticamente se considerara como atípica respecto a las demás viviendas de un mismo grupo.

El criterio que se utilizó para filtrar o remover aquellas viviendas con una densidad local con valores atípicos fue el percentil 90 de la distribución de esta variable. De esta forma, dentro de cada cluster se removieron todas aquellas viviendas que tuvieran densidades locales dentro del 10% más alto, es decir, aquellas densidades que pudieron haber ingresado a un grupo por efecto del carácter jerárquico del algoritmo HDBSCAN. Una vez que estas viviendas fueron removidas, se generaron los polígonos de los clusters dentro de cada comuna.

Por otro lado, el uso de la distancia mutua de alcance para filtrar las viviendas atípicas dentro de un cluster permitió a posterior generar una tipología de viviendas ruido, siendo posible clasificar estas en dos categorías: viviendas ruido permanente y viviendas ruido satélite (Figura 20).

Figura 20: Distribución de viviendas ruido tipo permanente y satélite.



Las viviendas ruido satélite se definen como aquellas viviendas que originalmente fueron clusterizadas por el algoritmo HDBSCAN, pero que fueron removidas debido a que presentaban una densidad local muy diferente respecto a las demás viviendas dentro del cluster al cual pertenecían. En cambio, la categoría de ruido permanente representa a aquellas viviendas que desde un principio fueron catalogadas como ruido, y que por lo tanto se mantienen en este mismo estatus después de realizar de remover las viviendas con densidades atípicas.

A modo de ejemplo la Figura 20 muestra un conjunto de viviendas rurales pertenecientes a la Aldea de Choroico en la comuna de Cunco región de la Araucanía, donde se puede observar que las viviendas ruido tipo satélite se distribuyen en las proximidades de un cluster, dando cuenta de la existencia de una gradiente o menor densidad de viviendas por fuera del área de un cluster principal.

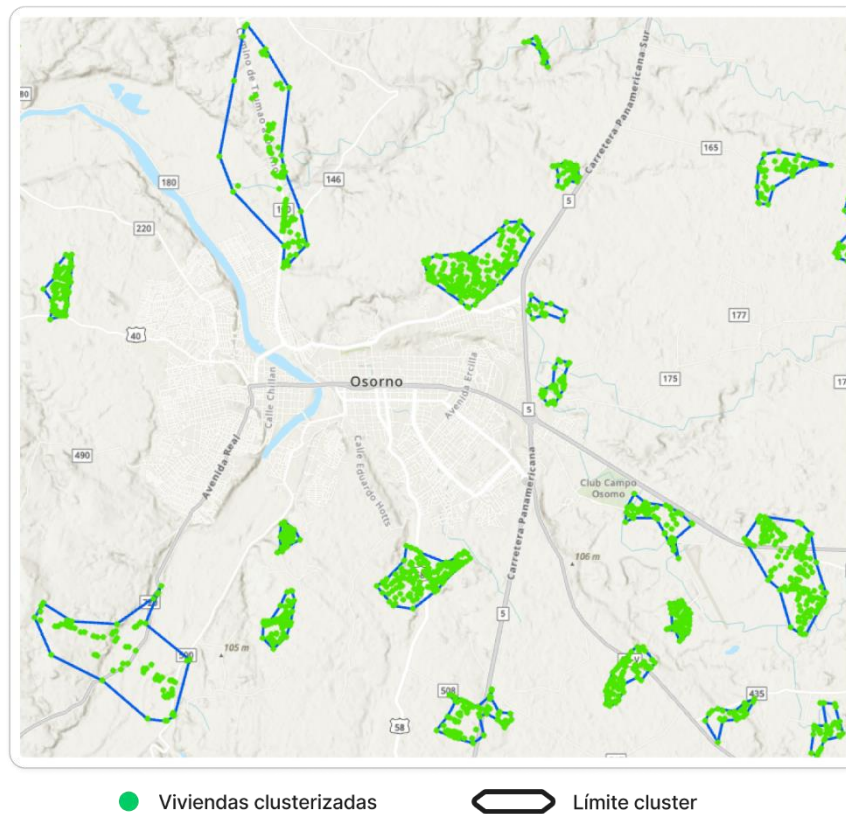
En cuanto a los polígonos que representan los límites de los clusters, gracias a la misma variable de distancia mutua de alcance se propuso una clasificación de la coherencia o consistencia de las viviendas de un cluster, existiendo la categoría de “Consistente” y “No Consistente”.

Estas categorías están creadas a partir de la distancia mutua de alcance, específicamente de cuán diferente es esta distancia (o densidad local) de las viviendas que pertenecen a un mismo cluster. De esta forma, se catalogó como “Consistente” a

aquellos clústeres que presenten viviendas con una densidad local (o distancia mutua de alcance) muy similares, mientras que los clusters catalogados como “No consistentes” corresponden a grupos que presenten viviendas con densidades local muy diferentes.

De esta manera, las categorías de consistencia de los clusters permiten comprender la morfología y conformación de cada grupo de viviendas, por lo que grupos consistentes representan a un conjunto de viviendas con conforman un polígono regular y más homogéneo, mientras que clusters irregulares o con formas en punta se clasifican en la categoría de no consistentes.

Figura 21: Vista de viviendas clusterizadas y de los límites aproximados de cada cluster



Adicionalmente, se buscó clasificar los clusters de acuerdo con su tamaño según la cantidad de viviendas. Para ello, se utilizó como base la clasificación de SSR que establece la Ley 20.988 respecto a los operadores, basado en la cantidad de arranques. En esta, se establecen tres categorías: Operador Menor, hasta 300 arranques; operador mediano, entre 301 y 600 arranques; y operador mayor, de 601 o más arranques. No obstante, si bien esta definición es una propuesta metodológica, a partir de los resultados obtenidos del modelo de agrupamiento, se evaluará según las métricas y estadísticas propias de la distribución de los datos (promedio, desviación estándar, mediana, mínimo y máximo) la división en subcategorías, en el caso que estas clasificaciones propuestas no sean representativas para los datos del modelo.

El resultado final del flujo metodológico se puede observar en la Figura 21, la cual muestra la localización de aquellas viviendas que fueron agrupadas dentro de un grupo o cluster por HDBSCAN, además de un polígono que representa el límite aproximado de cada agrupación.

El consolidado de todas las viviendas rurales con los resultados de HDBSCAN junto con los polígonos con los límites de cada cluster se almacenaron en dos capas llamadas “viviendas\_clusters\_hdbs” y “poligonos\_clusters\_hdbs” respectivamente, cuyo diccionario de datos es posible encontrarlo en el Anexo 3.

## 5. TRABAJO DE CAMPO

El modelo territorial tiene como objetivo integrar información espacial, socioeconómica y ambiental para priorizar inversiones de infraestructura en zonas rurales a nivel nacional. En este marco, el levantamiento de información en terreno adquiere un rol fundamental para garantizar la representatividad y calidad del modelo de agrupación en las distintas macrozonas de interés para el MOP.

El trabajo en terreno se estructuró en dos etapas: primero, un preterreno en la Región de Valparaíso, destinado a ajustar la metodología y validar los instrumentos de medición; y luego, una campaña de dos semanas en las macrozonas norte y sur, donde se aplicaron las herramientas y se recopilaron datos en localidades seleccionadas. A continuación, se describe el trabajo de campo realizado, la validación de los instrumentos de medición y resultados obtenidos.

### 5.1. Estructura y programación del trabajo de campo

El trabajo de terreno tuvo como finalidad validar la capacidad del Modelo de Agrupamiento HDBSCAN (MinPts=20) para representar la realidad territorial de los asentamientos rurales, evaluando su coherencia espacial frente a la distribución real de viviendas y su relación con elementos naturales, infraestructura y actividades productivas. Además, el terreno permitió levantar información clave para la caracterización de los clusters identificados, como aspectos de la materialidad de viviendas, accesibilidad a los clusters y disponibilidad de agua potable.

En primera instancia, en el trabajo de campo se realizó un preterreno en la provincia de Quillota, al ser representativa de la Macrozona Centro. El objetivo fue validar los procesos e instrumentos de medición. El preterreno se realizó en compañía de analistas regionales de la DIRPLAN, siendo guiados por el director regional de Valparaíso. Los puntos de observación se realizaron en los sectores de Pachacama (La Calera), Los Castaños (La Cruz), El Boco y Rauten (Quillota).

A partir de las observaciones obtenidas en el preterreno, se realizaron ajustes a los instrumentos de medición y fueron clave para definir la planificación, logística, actividades y metodologías a implementar en las macrozonas norte y sur. El primer terreno de esta campaña se realizó en las Provincias de Concepción, BíoBío, Itata y Punilla, al ser áreas de interés en la Macrozona Sur. En cambio, en el segundo terreno de la Macrozona Norte se visitaron las provincias de Copiapó, Huasco, Limarí y Elqui. Asimismo, en ambos terrenos el equipo técnico se reunió con analistas regionales de la Dirección de Planeamiento (DIRPLAN), en compañía de analistas de la Subdirección de Servicios Sanitarios Rurales (SSR) y de la Dirección de Vialidad.

En la Tabla 9 se ofrece detalle de las visitas realizadas dando cuenta de la comuna y localidad donde estaban ubicadas las observaciones:

Tabla 9: Ubicaciones de visitas a terreno.

Visita	Fecha	Región	Comuna	Localidad	Categoría
1	25/07/2025	Valparaíso	Quillota	La Palma	Parcela-Hijuela
2	28/07/2025	BioBío	Hualqui	Lo Vargas	Parcela-Hijuela
3	28/07/2025	BioBío	Hualqui	Llepinhue	Fundo-Estancia-Hacienda
4	28/07/2025	BioBío	Hualqui	Chanco	Parcela-Hijuela
5	29/07/2025	BioBío	Yumbel	Cambrales	Parcela-Hijuela
6	29/07/2025	BioBío	Yumbel	Los puentes de Tapihue	Parcela-Hijuela
7	31/07/2025	Ñuble	Treguaco	Valle Alegre	Parcela-Hijuela
8	31/07/2025	Ñuble	Treguaco	Paniagua	Indeterminada
9	31/07/2025	Ñuble	Portezuelo	Antequeroo	Indeterminada
10	31/07/2025	Ñuble	Ñiquén	Tiuquilemu	Parcela-Hijuela
11	04/08/2025	Atacama	Caldera	Caldera	Comunidad Indígena
12	05/08/2025	Atacama	Caldera	Barranquilla	Asentamiento Pesquero
13	05/08/2025	Atacama	Huasco	El Pino	Indeterminada
14	05/08/2025	Atacama	Freirina	Nicolasa	Caserío
15	05/08/2025	Atacama	Vallenar	Cavanca	Parcela-Hijuela
16	05/08/2025	Atacama	Vallenar	El Jilguero	Caserío
17	07/08/2025	Coquimbo	Coquimbo	Pan de azúcar	Parcela de agrado
18	07/08/2025	Coquimbo	Coquimbo	Las Barrancas	Parcela-Hijuela
19	07/08/2025	Coquimbo	Coquimbo	Guanaqueros	Parcela de agrado
20	07/08/2025	Coquimbo	Ovalle	Comunidad Agrícola Histórica Higuieritas Unidas	Parcela de agrado
21	07/08/2025	Coquimbo	Monte Patria	Comunidad Agrícola Histórica Monte Patria	Caserío
22	07/08/2025	Coquimbo	Ovalle	Talca	Indeterminada
23	08/08/2025	Coquimbo	La Serena	Punta de Piedra	Caserío

El detalle de la planificación de las actividades del trabajo de campo se presenta en la Figura 22:

Figura 22: Planificación de las actividades realizadas en terrenos.

Macrozona	Región	Actividad/Provincia	Julio					Agosto					
			25	28	29	30	31	1	4	5	6	7	8
			V	L	M	M	J	V	L	M	M	J	V
Centro	Valparaíso	Reunión con DIRPLAN Regional											
		Pre-terreno Provincia de Quillota											
Sur	Biobío	Reunión con DIRPLAN Regional											
		Terreno Provincia de Concepción											
		Terreno Provincia de Biobío											
	Ñuble	Reunión con DIRPLAN Regional											
		Terreno Provincia de Itata											
		Terreno Provincia de Punilla											
Norte	Atacama	Reunión con DIRPLAN Regional											
		Terreno Provincia de Copiapó											
		Terreno Provincia de Huasco											
	Coquimbo	Reunión con DIRPLAN Regional											
		Terreno Provincia de Limarí											
		Terreno Provincia de Elqui											

Las actividades y enfoques fueron definidos para ser aplicados en las tres regiones. Sin embargo, si bien se consideró una misma metodología, en cada actividad se tuvo que ajustar a las particularidades regionales (Figura 23 y Figura 24).

Figura 23: Ejemplo de reuniones y coordinación a puntos de observación junto al equipo DIRPLAN regional. En ellas se puede observar: a) Reunión con DIRPLAN Regional Ñuble, b) y c) Terreno Provincia de Itata y d) Reunión con DIRPLAN Regional Atacama.

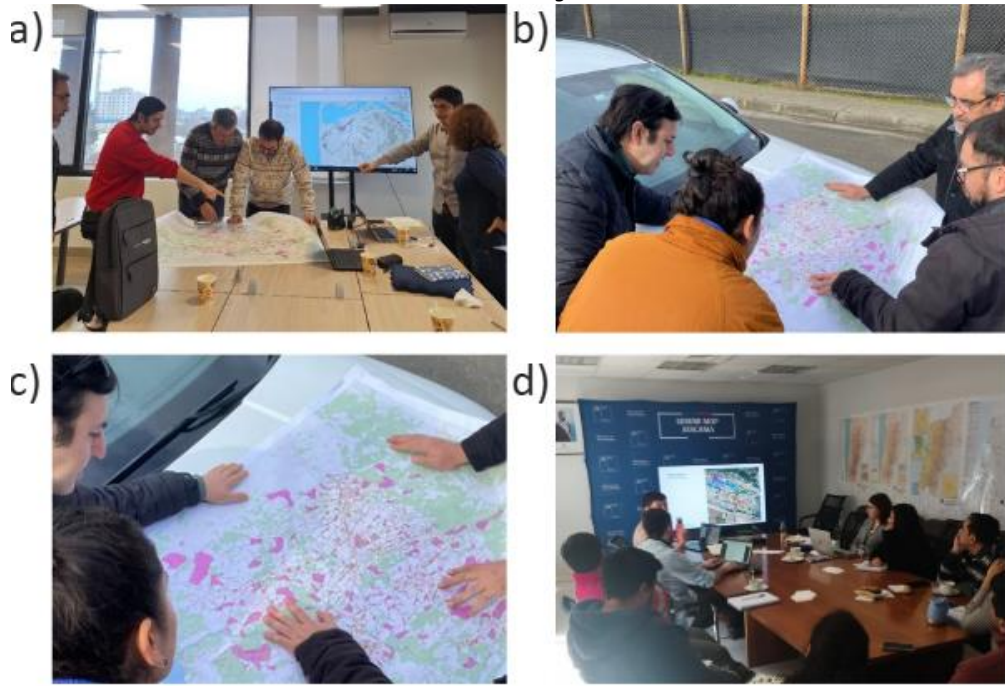
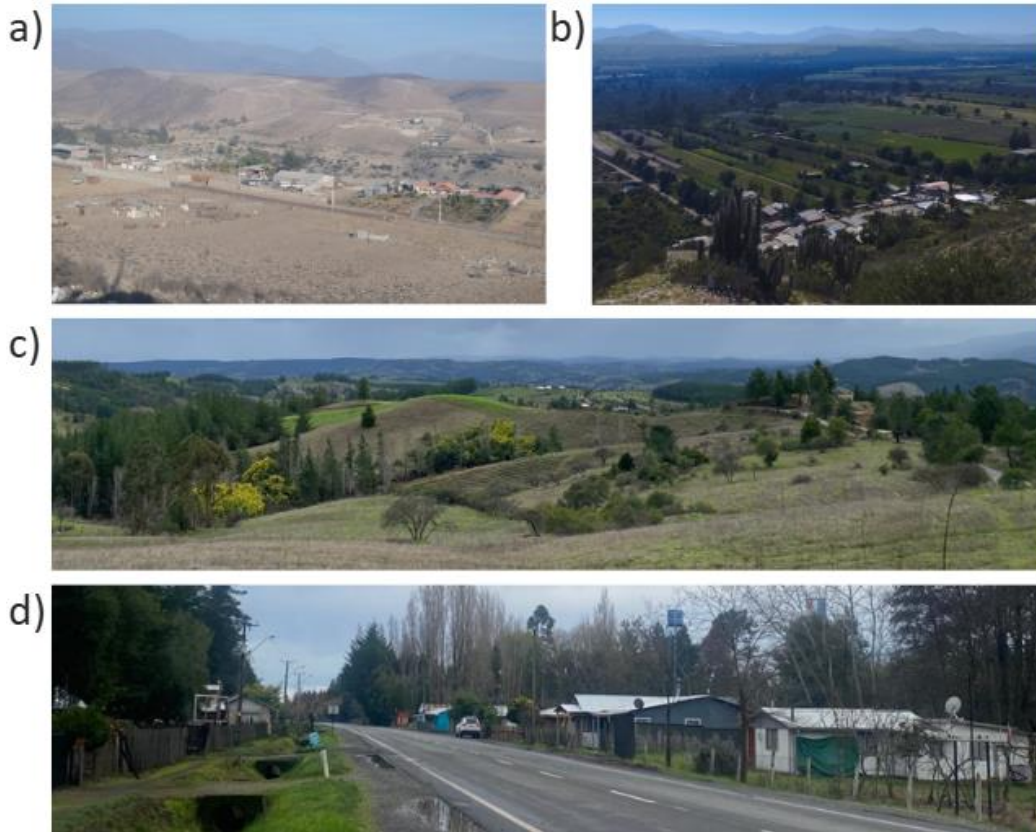


Figura 24: Puntos de observación característicos en preterreno (b), terreno Macrozona Norte (a) y Sur (c y d).



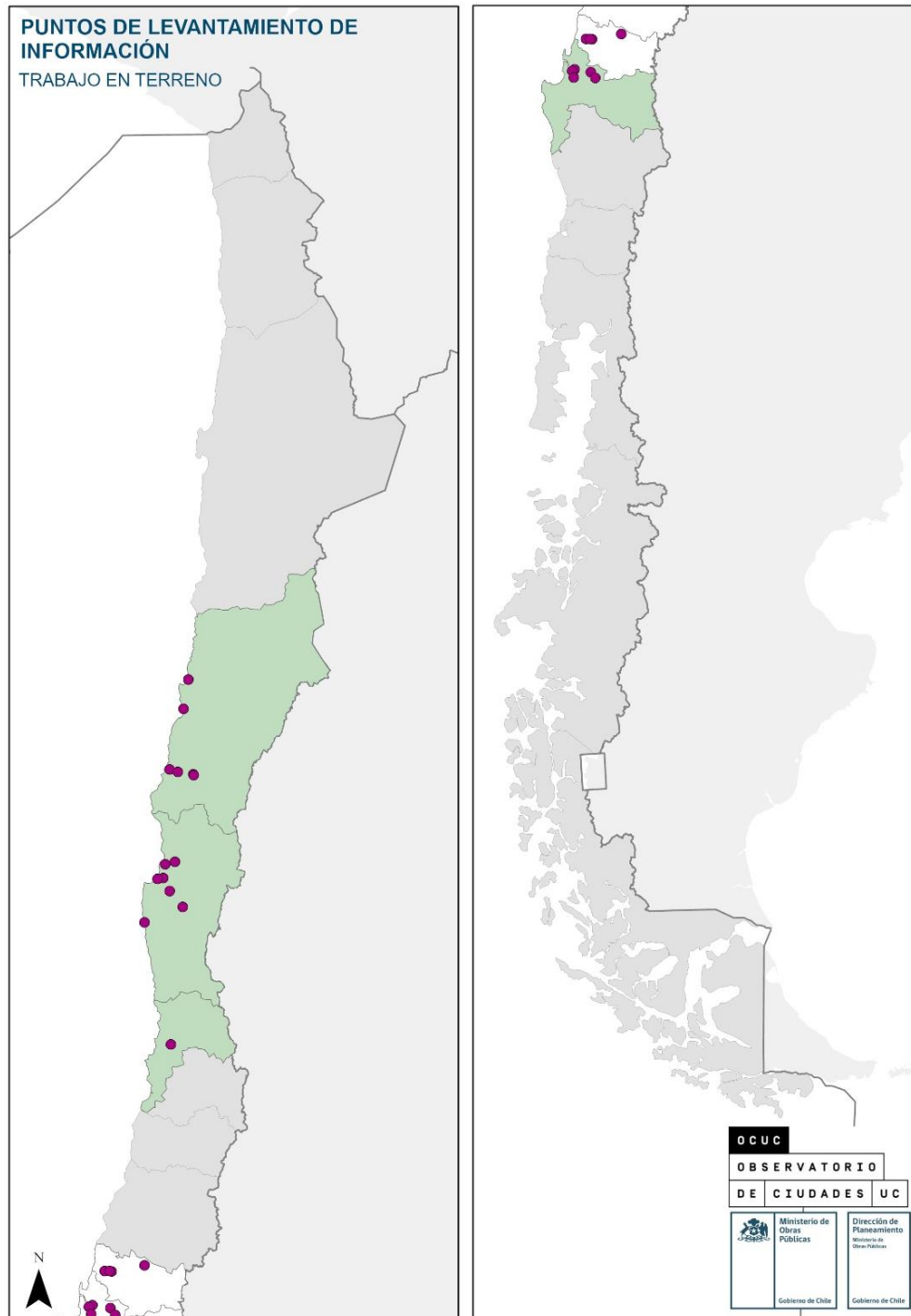
## 5.2. Validación de la metodología para el trabajo de campo

Específicamente, el levantamiento en terreno consideró la visita a 10 a 12 localidades rurales por campaña, seleccionadas para representar diferentes tipologías de asentamientos (Figura 25). La validación se estructuró en tres elementos a evaluar:

- Comparación de parámetros del modelo HDBSCAN: Evaluación de clusters generados con distintos valores de *MinPts* (30, 50 y 100), verificando su ajuste a los patrones reales de poblamiento.
- Observación de la coherencia espacial: Análisis de la continuidad y forma de los asentamientos en relación con la distribución de viviendas, presencia de barreras naturales/antrópicas (ríos, montañas, embalses, carreteras, etc.) y disponibilidad de infraestructura hídrica (abastecimiento por SSR, aljibe, sistema particular, Sanitaria, etc.)
- Levantamiento de información territorial: Registro en terreno de atributos complementarios para caracterizar los asentamientos, incluyendo accesibilidad, servicios básicos y actividades productivas predominantes.

Para la captura y sistematización de datos se utilizaron las aplicaciones ArcGIS Field Maps y Survey123, que permitieron visualizar en tiempo real los clusters generados por el modelo y asociar información georreferenciada sobre las variables observadas en terreno.

Figura 25: Puntos de levantamiento de información.



En términos generales, el trabajo en terreno permitió reforzar el objetivo general del modelo de agrupación, que es focalizar áreas de interés (considerando que la etapa de priorización es parte de un proceso posterior). De esta forma, el modelo funciona como un indicador de prefactibilidad que orienta al analista territorial en la toma de decisiones para la inversión en infraestructura.

### 5.3. Hallazgos

Las actividades, mediciones y recorridos realizados en terreno permitieron identificar patrones diferenciados de agrupamiento, evidenciando que la variación del parámetro *MinPts* incidió directamente en la definición de la tipología de asentamientos.

En general, se constató que el modelo tiende a representar adecuadamente los clusters en contextos de asentamientos dispersos, mientras que en áreas semiconcentradas se reafirmó la generación de ajustes según la realidad regional, para reflejar con mayor precisión la continuidad de las viviendas y su relación con la infraestructura y servicios.

Bajo este contexto, el levantamiento territorial permitió identificar cuatro atributos claves asociados a los clústeres o agrupamientos:

- Variabilidad en la accesibilidad y conectividad interna de los asentamientos.
- Diferencias en la cobertura de servicios básicos (agua, saneamiento, electricidad, entre otros).
- Influencia de barreras naturales (ríos, cerros, quebradas, entre otros) en la configuración espacial.
- Actividades productivas predominantes que explican la localización y morfología del poblamiento.

Específicamente, durante el preterreno realizado en la provincia de Quillota, el equipo DIRPLAN regional destacó la problemática hídrica existente en la región, comparando el escenario de la provincia con Petorca. Por lo cual, en este contexto es clave una gestión adecuada de los recursos hídricos, considerando que una proporción significativa se destina a actividades agrícolas para el riego de cultivos en los valles y cerros.

El levantamiento de información permitió identificar diversas infraestructuras, tales como postas rurales, establecimientos educacionales, sistemas de Servicio Sanitario Rural (SSR) con un fuerte componente de organización comunitaria, entre otras. Entre las problemáticas detectadas, se encuentra la dependencia a camiones aljibes en sectores alejados de los SSR, la competencia por el uso del agua, la extracción de áridos, el deterioro de caminos por derrumbes asociados a escorrentías en pendientes pronunciadas de los cerros, entre otras.

Además, en esta etapa los analistas territoriales recomendaron por primera vez al equipo técnico realizar un cruce de las viviendas rurales y las aldeas georreferenciadas por el APC 2023 (INE).

Luego, en el desarrollo del primer terreno realizado en la Macrozona Sur, se identificaron las primeras limitaciones del modelo. Entre ellas destacan: la agrupación de viviendas rurales entre barreras geográficas (caso del río Biobío); agrupaciones que incorporan

puntos distantes (generando polígonos poco representativos); clasificación de micro-densificaciones como ruido, dada la alta concentración de viviendas rurales en estas regiones (generalmente por presentar grandes superficies con cultivos o agricultura familiar, silvicultura, entre otras actividades), entre otras limitantes.

Por estas razones, los analistas territoriales destacaron el modelo por gradiente (30-50-100 viviendas), siendo considerado adecuado para poder caracterizar de mejor forma las viviendas rurales en estas regiones. Asimismo, se refuerza la idea de incorporar las aldeas georreferenciadas por el APC 2023 (INE), tal como se recomendó en el preterreno.

Por último, se discute la separación de un cluster por infraestructuras como autopistas o calles enroladas, comentando los analistas que no siempre constituye una limitante, depende del proyecto, nivel de inversión e interés de implementación.

En el caso del segundo terreno realizado en la macrozona norte, se identificaron particularidades relevantes. En primer lugar, se comentó por los analistas territoriales que las rutas principales (como rutas concesionadas, ej. Ruta 5) deben ser consideradas como barreras, a diferencia de caminos menores.

Además, a diferencia de la experiencia de la macrozona sur, los ríos no constituyen una barrera natural significativa; un ejemplo fue el caso de la infraestructura SSR Nicolasa (Freirina), cuya área de servicio atraviesa el río Huasco.

Otro aspecto relevante, similar a las micro-densificaciones encontradas en la macrozona sur, el parámetro *MinPts* no se ajustó adecuadamente con 50 viviendas, pero sí al reducir el valor (con 30 viviendas), ocurriendo especialmente en zonas no habitadas de la región de Atacama, donde la barrera geográfica era el desierto.

Por otra parte, respecto de la ocupación de terrenos por grupos de viviendas informales (ya sea de primera residencia o segunda residencia), los analistas territoriales MOP de la región, señalaron que no constituye una prioridad formal para la inversión del Ministerio.

Sin embargo, sí que es una situación considerada, y que, en última instancia, la decisión de intervenir queda supeditada a criterio de cada región. Finalmente, se verificó en terreno que numerosas edificaciones rurales eran efectivamente bodegas, galpones o industrias, volviéndose necesario realizar un filtro a la base de datos definida por el APC 2023.

## 6. MODELO DE AGRUPAMIENTO RURAL

El modelo de agrupamiento de viviendas a partir del algoritmo HDBSCAN se ejecutó utilizando un parámetro de mínimo de puntos (MinPts) igual a 20, aplicado sobre 1.657.970 viviendas rurales. Esta capa de información geoespacial proviene de la consolidación de las edificaciones rurales APC23 (corregidas) y las edificaciones rurales obtenidas desde la plataforma Overture Maps y ajustadas según el número de viviendas de Aldeas Rurales del INE (ver apartado 4.2).

Con el objetivo de organizar y estructurar el análisis, las 16 regiones del país fueron agrupadas en cuatro macrozonas, siguiendo la clasificación propuesta por DIRPLAN-MOP, adaptada para mejorar la representatividad territorial de cada zona. Las macrozonas abarcan las siguientes regiones:

- Macrozona Norte: Arica y Parinacota; Tarapacá; Antofagasta; Atacama y Coquimbo.
- Macrozona Centro: Valparaíso; Metropolitana de Santiago; Libertador General Bernardo O'Higgins y del Maule.
- Macrozona Sur: Ñuble; Biobío; La Araucanía y Los Ríos
- Macrozona Austral: Los Lagos; Aysén del General Carlos Ibáñez del Campo y Magallanes y de la Antártica Chilena.

A continuación, se presentan los resultados del modelo a escala nacional, para luego profundizar en el análisis por macrozona, abordando la caracterización del agrupamiento a nivel provincial y comunal, identificando patrones territoriales locales y contrastando distintas unidades geográficas del país.

### 6.1. Resultados generales

A nivel nacional, de un total de 1.657.970 viviendas rurales analizadas, 1.108.779 viviendas (66,88%) fueron agrupadas en clústeres, mientras que 549.191 viviendas (33,12%) fueron clasificadas como ruido, al no cumplir con los criterios mínimos de agrupamiento (*MinPts* igual a 20 viviendas). La Tabla 10 presenta el detalle de los resultados por provincia y comuna de la ejecución del algoritmo (Anexo 2 y Anexo 4).

Tabla 10: Resumen de viviendas clusterizadas y viviendas ruido por región.

Región	Total de viviendas	Viviendas clusterizadas	Índice de clúster	Viviendas ruido	Siempre ruido	Ruido satélite
<b>Arica y Parinacota</b>	15.953	10.200	64%	5.753	71%	29%
<b>Tarapacá</b>	21.587	15.868	74%	5.719	62%	38%
<b>Antofagasta</b>	17.958	13.688	76%	4.270	59%	41%
<b>Atacama</b>	40.050	29.508	74%	10.542	62%	38%

<b>Coquimbo</b>	113.195	79.291	70%	33.904	69%	31%
<b>Valparaíso</b>	135.968	99.594	73%	36.374	63%	37%
<b>Metropolitana</b>	147.388	110.142	75%	37.246	61%	39%
<b>O'Higgins</b>	142.632	100.539	70%	42.093	67%	33%
<b>Maule</b>	181.691	119.268	66%	62.423	72%	28%
<b>Ñuble</b>	115.824	72.203	62%	43.621	74%	26%
<b>Biobío</b>	156.903	100.345	64%	56.558	73%	27%
<b>Araucanía</b>	258.665	160.385	62%	98.280	75%	25%
<b>Los Ríos</b>	79.708	49.292	62%	30.416	76%	24%
<b>Los Lagos</b>	187.121	120.023	64%	67.098	74%	26%
<b>Aysén</b>	25.429	16.551	65%	8.878	73%	27%
<b>Magallanes</b>	17.898	11.882	66%	6.016	68%	32%
<b>Total nacional</b>	<b>1.657.970</b>	<b>1.108.779</b>	<b>67,9%</b>	<b>549.191</b>	<b>68,7%</b>	<b>31,3%</b>

Es relevante mencionar en la Tabla 10 se incorporan nuevos indicadores, las viviendas que siempre son ruido y las viviendas ruido “satélites” (definidas metodológicamente en el apartado 4.2.3.). En este sentido, las viviendas ruido que se catalogaron como ruido permanente a nivel nacional en promedio alcanzan un 71,05%, es decir, pese a las variaciones del parámetro base del modelo *MinPts*, la mayoría de las viviendas seguían siendo lo suficientemente significativas respecto a su densidad y distancia para separarse de las agrupaciones.

En cambio, respecto a las viviendas ruido a nivel nacional catalogadas como “satélite” en promedio, un 28,95% se encuentra asociada a las agrupaciones principales en términos de en densidad y distancia. Por ende, estas últimas son viviendas potenciales de ser parte de los clusters en caso de modificar el parámetro *MinPts*, pero que continúan siendo ruido al no cumplir los criterios del modelo. Además, es relevante mencionar que un número importante de viviendas ruido “satélite” resultaron en esta categoría a partir de los filtros aplicados en el desarrollo del modelo, bajo el objetivo de obtener viviendas clusterizadas más robustas en términos de densidad y distancia.

La Tabla 11 y Anexo 4 presenta el detalle de los clusters obtenidos y se incorporan dos indicadores adicionales: el grado de agrupamiento y consistencia. De forma complementaria, la Cartografía 1 presenta la clusterización a escala regional.

Tabla 11: Resultados del algoritmo HDBSCAN a nivel de clúster.

Región	Total de clusters	Grado de agrupamiento de los clusters (%)					Consistencia de los clusters (%)	
		Muy bajo	Bajo	Medio	Alto	Muy Alto	Consistente	No consistente
Arica y Parinacota	134	52,2	31,3	14,2	0,7	1,5	93,3	6,7
Tarapacá	152	44,7	27,6	21,7	4,6	1,3	92,8	7,2
Antofagasta	109	45	19,3	22,9	11	1,8	91,7	8,3
Atacama	246	48	27,6	16,7	4,9	2,8	93,1	6,9
Coquimbo	784	44,5	27,2	22,4	5	0,9	94,3	5,7
Valparaíso	779	38,4	29,3	23,9	6,2	2,3	92,7	7,3
Metropolitana	848	39,4	29	24,4	4,8	2,4	93,8	6,2
O'Higgins	1103	43,2	30,6	22,8	2,4	1	93,5	6,5
Maule	1493	48,9	29	19,7	1,9	0,5	94,3	5,7
Ñuble	915	49,4	29,4	18,5	2,1	0,7	94,1	5,9
Biobío	1251	48,8	29,3	18,8	3	0,2	93,8	6,2
Araucanía	1734	53	27,1	16	2,5	1,3	94,1	5,9
Los Ríos	601	51,7	25,6	19,5	2,7	0,5	94	6
Los Lagos	1320	53	28	15,2	2,9	1	93,6	6,4
Aysén	196	51,5	25	19,4	3,1	1	91,8	8,2
Magallanes	142	51,4	33,1	12,7	1,4	1,4	90,8	9,2
<b>Total nacional</b>	<b>11.807</b>	<b>47,7%</b>	<b>28%</b>	<b>19,3%</b>	<b>3,7%</b>	<b>1,3%</b>	<b>93,2%</b>	<b>6,8%</b>

En general, la Tabla 11 evidencia que el mayor número de clusters se encuentra en la región de La Araucanía (1.734), a diferencia de la región de Antofagasta que es la región que presenta el menor número, siendo únicamente 109 agrupaciones.

A partir de lo mencionado en la metodología, la clasificación inicial de los clusters de acuerdo con su tamaño según la cantidad de viviendas<sup>5</sup> se realizó en tres categorías: Operador Menor, hasta 300 arranques; operador mediano, entre 301 y 600 arranques; y operador mayor, de 601 o más arranques. Sin embargo, al clasificar los resultados en esta categoría, cerca del 90% se ubicaba en la categoría de “Menor” (hasta 300 arranques). Por ende, para obtener mayor representatividad al clasificar el tamaño de los clústeres, siguiendo las estadísticas de la distribución de los datos (promedio,

<sup>5</sup> Detalle se describe en metodología. Específicamente, se utiliza de referencia la clasificación de SSR que establece la Ley 20.988 respecto a los operadores, basado en la cantidad de arranques.

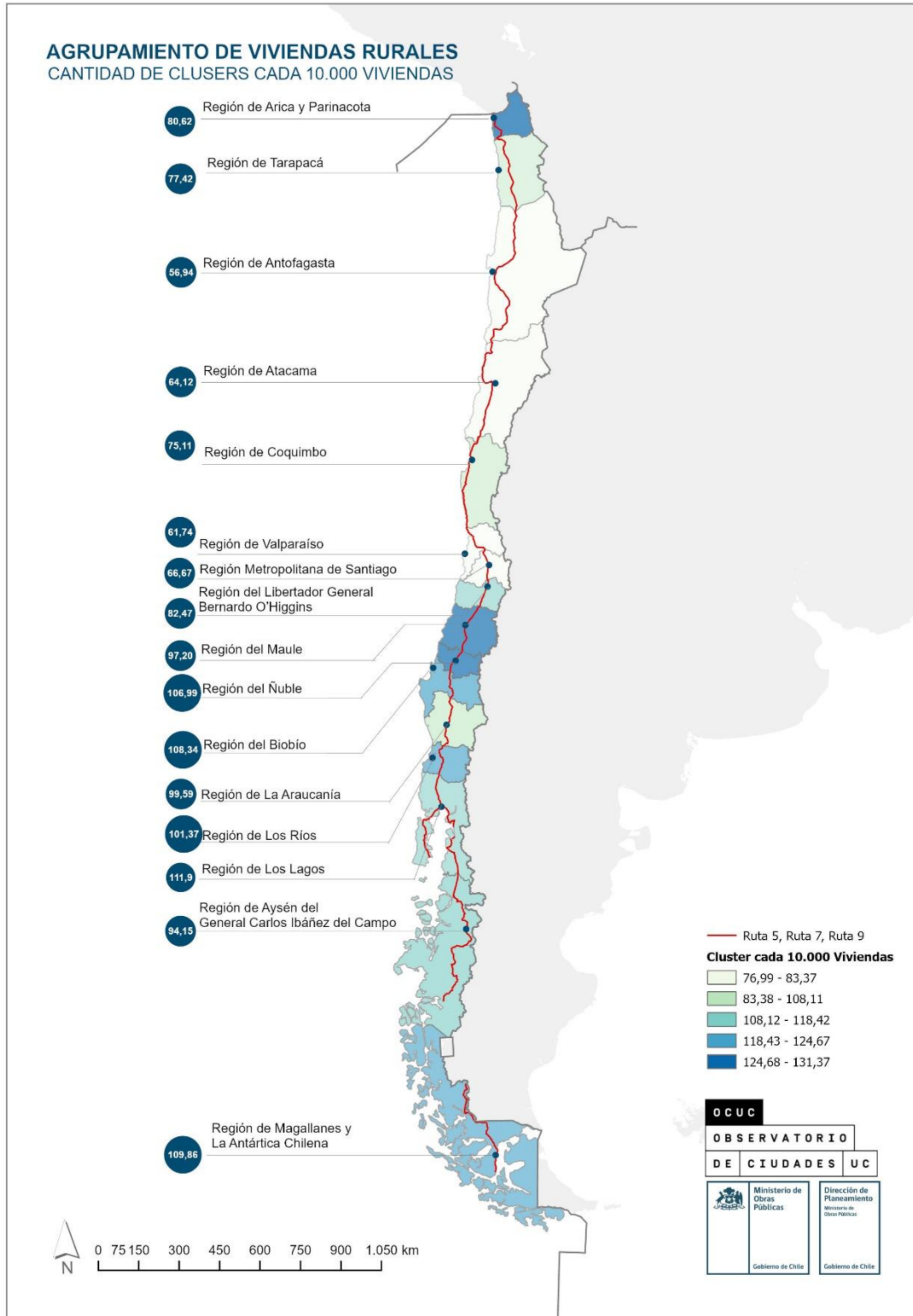
desviación estándar, mediana, mínimo y máximo), se subdividió la categoría de operador menor, estableciendo finalmente los siguientes tipos de agrupamientos:

- Agrupamiento muy bajo: menor a 50 viviendas
- Agrupamiento bajo: entre 50 y 100 viviendas
- Agrupamiento medio: entre 101 y 300 viviendas.
- Agrupamiento alto: entre 301 y 600 viviendas.
- Agrupamiento muy alto: más de 601 viviendas.

Bajo esta nueva clasificación del tamaño de los clústeres, se observa que en todas las regiones los clústeres se concentran en la categoría de agrupamiento muy bajo. Esto ocurre principalmente, porque el parámetro del algoritmo es de 20 viviendas, por lo que si este se modifica esta proporción cambia.

Por su parte, el nivel de consistencia se obtuvo a partir de la desviación estándar del parámetro de densidad local (ver metodología en punto 4.2.3.). En todas las regiones se alcanzan valores altos, siendo superiores a un 90,8% (siendo la región de Magallanes la que alcanza este valor). En cambio, los no consistentes obtienen menores valores, entre 5,7 a 9,2%, obteniéndose el primer valor en la región de Coquimbo y el último en la región de Magallanes.

Cartografía 1: Cantidad de clusters por región cada 10.000 viviendas



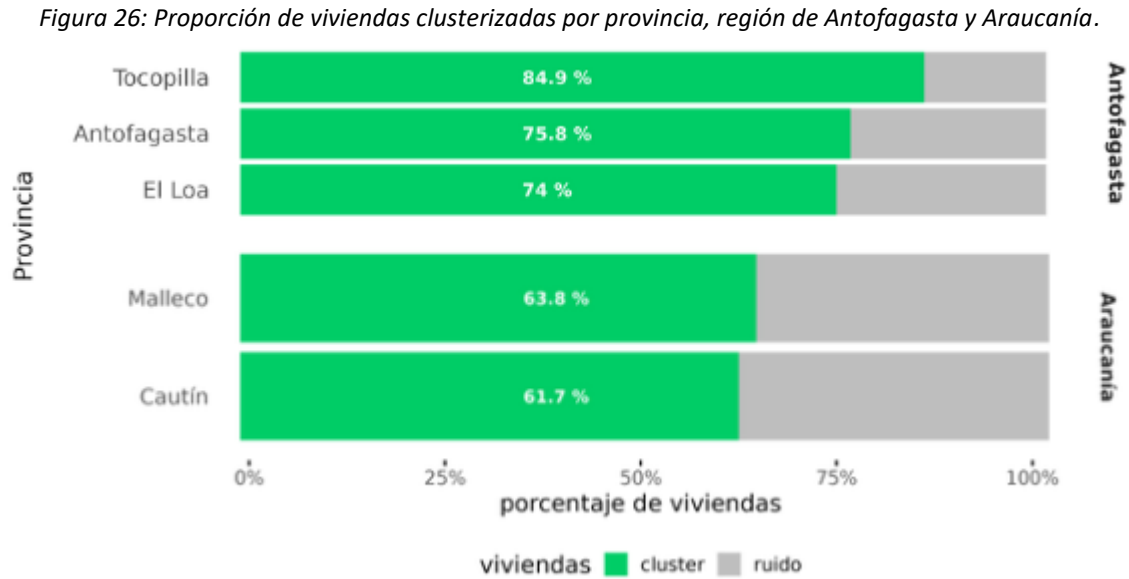
En la Cartografía 1 es posible observar que, en las regiones con condiciones geográficas extremas, como el Desierto de Atacama en el norte o la Patagonia en el sur, los asentamientos rurales tienden a concentrarse en núcleos compactos y definidos, debido a factores como la escasez de recursos, la presencia de barreras naturales y la necesidad de cercanía a servicios.

En estos contextos, el modelo detecta agrupamientos más fácilmente, ya que las viviendas presentan menores distancias relativas y mayor agrupamiento espacial. Por el contrario, en regiones con mayor dispersión de viviendas rurales, como ocurre en zonas agrícolas del centro y sur del país, es más frecuente que las viviendas queden fuera de los umbrales de agrupamiento, aumentando el porcentaje clasificado como ruido. Por ello, estas regiones del norte y extremo sur alcanzan altos niveles de clusterización relativa, a pesar de su baja cantidad de habitantes en términos absolutos.

En la misma línea, se observa que regiones con alto volumen de viviendas rurales (como La Araucanía, Los Lagos y Biobío), presentan un bajo índice de clusterización entre 62 a 64% (proporción entre el total de viviendas y las clusterizadas), aunque presentan un número elevado de clusters. Esto sugiere la presencia de una gran cantidad de núcleos rurales pequeños, fragmentados y distribuidos de manera heterogénea, lo que puede ser un insumo relevante para el diseño de infraestructura diferenciada según tipo de cluster.

Asimismo, se observa una diferencia entre el volumen total de viviendas rurales y la eficiencia del agrupamiento. Por ejemplo, La Araucanía, que concentra la mayor cantidad de viviendas rurales del país (258.665), presenta una tasa de clusterización del 62%. En contraste, Antofagasta, con apenas 17.958 viviendas, alcanza un 76% de agrupamiento. Esta diferencia evidencia que la eficacia del modelo no está determinada por la cantidad de viviendas, sino por el patrón espacial de ocupación rural. Este hallazgo subraya la importancia de incorporar un enfoque territorial diferenciado en el desarrollo del modelo de agrupamiento, considerando las particularidades morfológicas y de distribución de las viviendas en cada región.

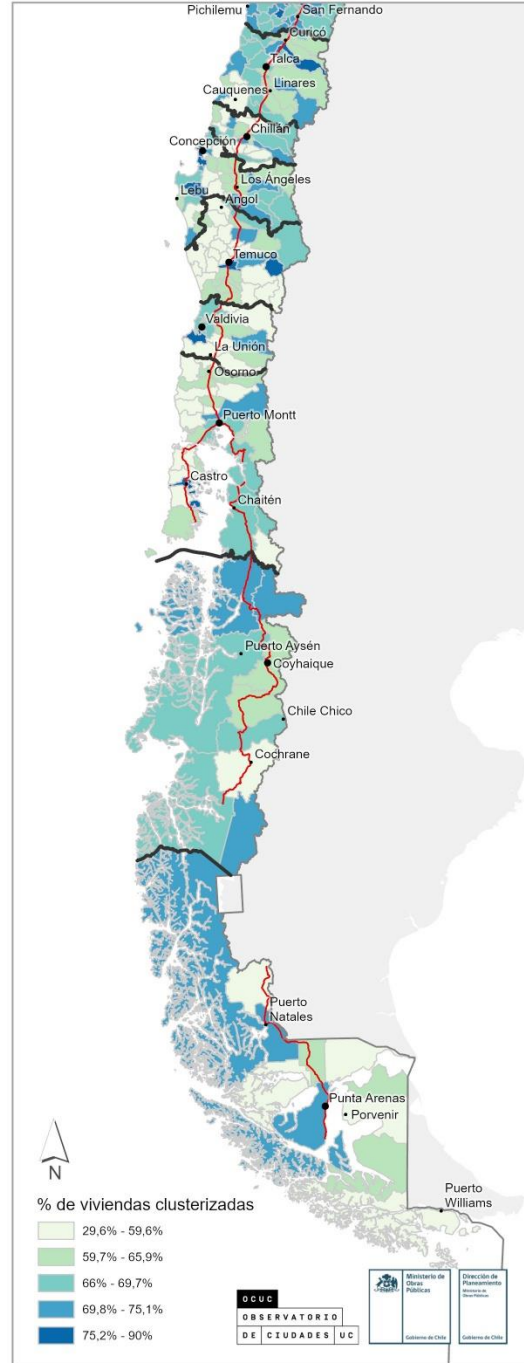
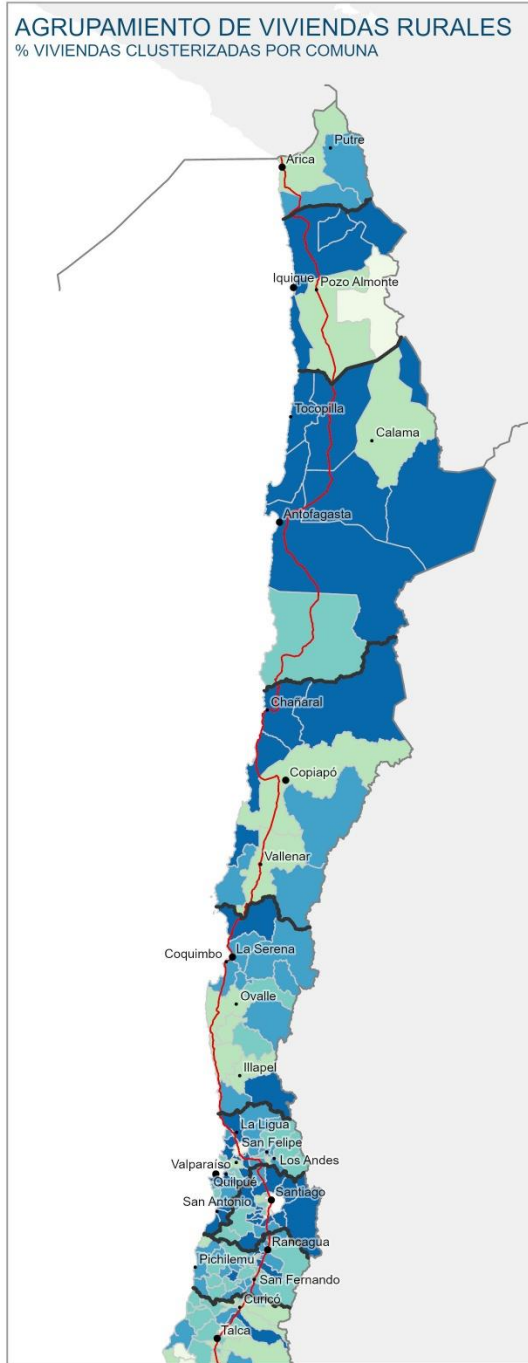
Esta diferencia se ilustra en la Figura 26, donde se compara el porcentaje de viviendas clusterizadas y no clusterizadas para cada provincia por región. En Antofagasta, provincias como El Loa y Tocopilla, a pesar de contar con pocas viviendas rurales, alcanzan niveles de agrupamiento superiores al 74%, debido a que las viviendas se concentran en pequeños enclaves bien definidos. Por el contrario, en La Araucanía, tanto en Cautín como Malleco, el algoritmo no supera el 63,8% de clusterización, ya que las viviendas rurales presentan una mayor dispersión territorial, dificultando la formación de conglomerados densos bajo los umbrales establecidos por el modelo.



A nivel comunal, la Cartografía 2 muestra los resultados del modelo de agrupamiento, evidenciando que las comunas con mayor cantidad de viviendas rurales (ubicadas principalmente en las macrozonas centro y sur) presentan índices más bajos de clusterización. En contraste, con comunas con menor número de viviendas rurales, localizadas en zonas extremas del país como el Desierto de Atacama, el Altiplano, la Patagonia y el territorio insular austral (macrozonas norte y austral), que registran mayores niveles de agrupamiento.

Por último, para esta segunda entrega del modelo de agrupamiento, se adjuntan los resultados de la clusterización en formato Geodatabase (.gdb), la cual contiene dos *feature classes*. La primera corresponde a los puntos de viviendas rurales agrupadas mediante el algoritmo HDBSCAN, con el nombre “viviendas\_agrupamiento\_hdbs”. La segunda capa, denominada “cluster\_viviendas\_hdbs”, representa los 11.807 clusters identificados, esta vez en formato de polígono, permitiendo visualizar la extensión territorial de cada agrupamiento. En el Anexo 3 se encuentra el diccionario de metadatos con la descripción de cada una de las variables contenidas en ambas capas.

Cartografía 2: Proporción de viviendas clusterizadas por comuna.



## 6.2. Macrozona Norte

A nivel regional, Antofagasta (76%), Atacama y Tarapacá (ambas con 74%) registran los mayores porcentajes de viviendas clusterizadas. En estos casos, se trata de regiones con baja densidad de población rural y patrones de asentamiento altamente concentrados en enclaves. En Antofagasta, este comportamiento se explica por dos factores principales: primero, por su carácter minero, que ha impulsado el desarrollo de localidades rurales vinculadas a faenas extractivas y parques industriales (como el sector Industrial La Negra o en María Elena); y segundo, por la presencia de asentamientos indígenas—como Peine, Socaire o Talabre— que forman clusters compactos en el desierto altoandino. En el caso de Atacama, el patrón de agrupamiento responde a distribución de los asentamientos rurales, donde las viviendas se concentran en torno a grandes ciudades y sus respectivos valles, como es el caso de Copiapó (valle del río Copiapó) y Vallenar (valle del río Huasco). En la región de Atacama, también destacan clusters localizados en el borde costero de la región y en la zona andina en los valles de las comunas de Tierra Amarilla y de Alto del Carmen.

En la región de Tarapacá, el algoritmo muestra un buen desempeño al agrupar más del 80% de las viviendas ubicadas en zonas remotas, como las comunas de Camiña (83%), Colchane (85,1%) y Huara (82%). Esto evidencia que, al utilizar un parámetro de MinPts bajo (20 viviendas), es posible detectar microdensificaciones.

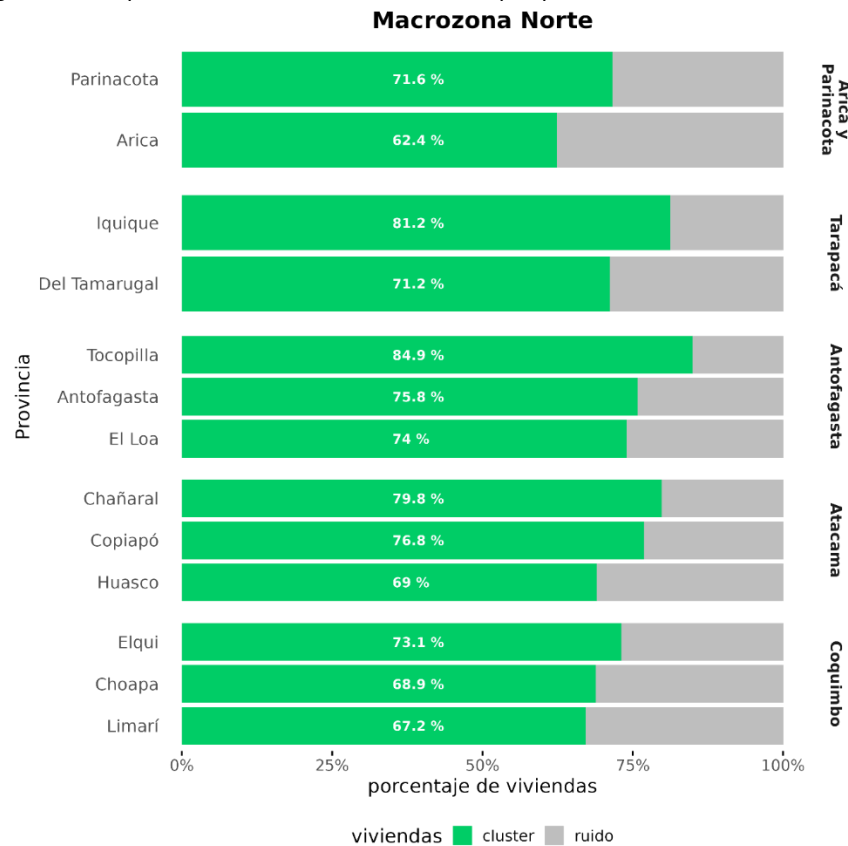
En contraste, las regiones de Arica y Parinacota (64%) y Coquimbo (70%) presentan una pequeña diferencia en la cantidad de viviendas agrupadas niveles de agrupamiento, reflejando una ruralidad más dispersa. La región de Coquimbo, visitada en terreno se caracteriza por el desarrollo de actividades agropecuarias tradicionales desarrolladas en valles transversales —como Elqui y Limarí— y en zonas de secano interior. Esta configuración genera una ocupación discontinua y de baja densidad, con viviendas ubicadas en quebradas, montañas o terrazas agrícolas, lo que dificulta que el modelo identifique conglomerados suficientemente densos bajo sus parámetros, aumentando la proporción de viviendas ruido.

Este fenómeno se acentúa en la zona de la Cordillera Andina del norte chico, que abarca comunas como Salamanca, Monte Patria, Río Hurtado, Paiguano, Vicuña y Alto del Carmen. En estos territorios de alta montaña, el relieve irregular y la presencia de múltiples barreras naturales —como cerros, quebradas profundas y divisorias de aguas— afectan directamente el rendimiento del algoritmo. El cual, al operar únicamente con criterios de densidad espacial, no reconoce elementos del paisaje como obstáculos a la conectividad territorial. Esto implica que viviendas ubicadas en laderas opuestas de una misma cordillera o separadas por quebradas sin conexión directa, pueden ser agrupadas en un mismo cluster por el modelo, pese a no tener una relación territorial real. En otras palabras, el algoritmo interpreta la cercanía geométrica sin considerar las limitaciones topográficas y de accesibilidad propias de estos entornos.

Frente a esta limitación, en la presente versión del Modelo de Agrupamiento se incorporaron dos ajustes metodológicos orientados a mejorar su robustez. Primero, el algoritmo se ejecutó a escala comunal, lo que permite que el modelo considere de manera implícita las barreras naturales que suelen coincidir con los límites administrativos (ríos, quebradas, líneas de cumbre, entre otros). Segundo, se redujo el parámetro MinPts a 20 viviendas, lo que hace posible identificar pequeños agrupamientos presentes en territorios más dispersos, como aquellos ubicados en zonas de alta montaña.

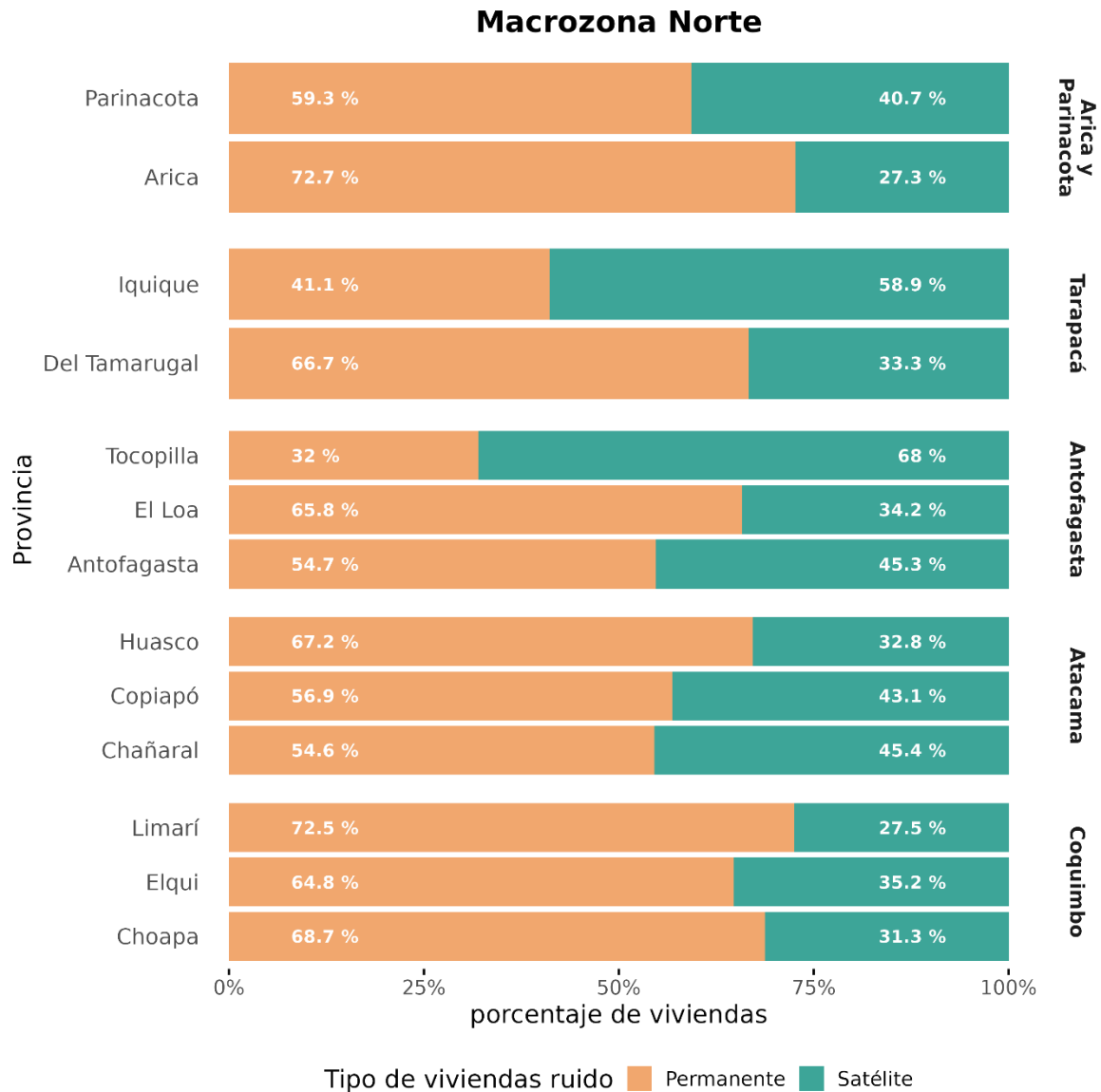
Por otra parte, a nivel provincial, la Figura 27 presenta la proporción de viviendas clusterizadas, en donde se observa que las provincias de Iquique y Tocopilla alcanzan el nivel más alto de agrupamiento. Al analizar los clusters de Iquique, se observa que el alto agrupamiento esta influenciado por un clúster muy grande al sur de la comuna de Iquique y Alto Hospicio que corresponde al campamento “Mirando las Estrellas”, el cual agrupa 1.643 viviendas. Por su parte, la provincia de Tocopilla alcanza valores altos porque en la comuna de Tocopilla se agrupan pequeños grupos de viviendas en sectores costeros, mientras que en la comuna de María Elena hay un solo cluster que agrupa 211 viviendas en el sector de Quillagua.

Figura 27: Proporción de viviendas clusterizadas por provincia en la Macrozona Norte



Como se mencionó anteriormente, con el fin de hacer una diferenciación entre las viviendas ruido -a partir de un filtro del percentil 90-, la Figura 28 presenta la proporción de los dos tipos de viviendas ruido: las que son ruido permanente y las que son de tipo satélite. Las de tipo satélite son viviendas no clusterizadas pero que se ubican en la periferia del clúster, sienten una especie de “satélite” a su alrededor. En cambio, las viviendas ruido permanente son las que posterior al filtro, se mantienen siendo ruido.

Figura 28: Proporción de viviendas ruido según su tipo en la Macrozona Norte.

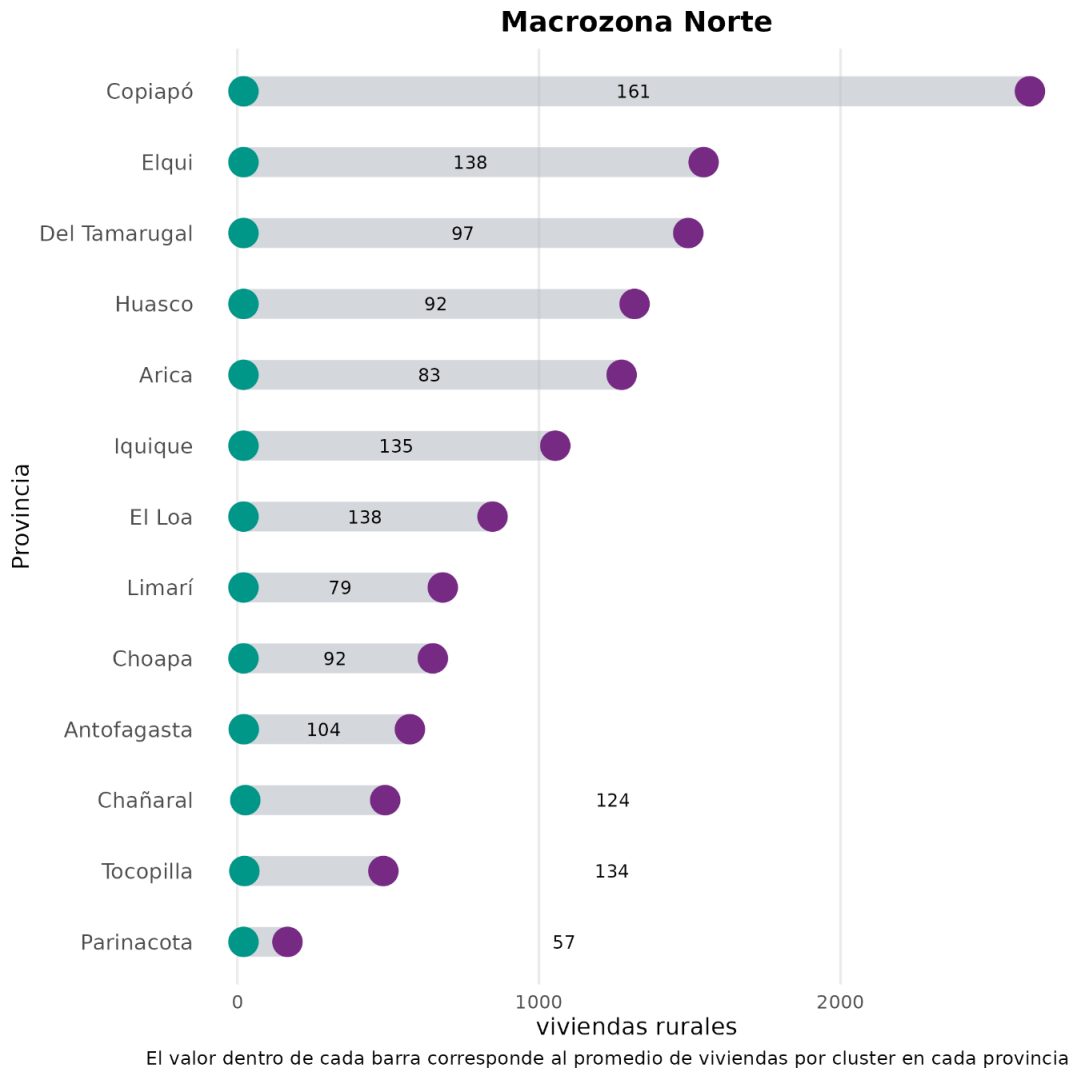


Esta clasificación resulta especialmente útil para identificar viviendas clasificadas como ruido, pero que se localizan en la periferia de los clusters. La Figura 29 muestra que las provincias de Iquique y Tocopilla concentran la mayor cantidad de este tipo de viviendas “ruido satélite”, lo que se relaciona directamente con sus altos niveles de clusterización. En otras palabras, en territorios donde existen fuertes tendencias al agrupamiento, pero

la densidad no alcanza el umbral exigido por el algoritmo HDBSCAN, es más probable encontrar un número significativo de viviendas clasificadas como ruido satélite.

De forma complementaria, la Figura 29 presenta el rango de viviendas por cluster a nivel provincial, mostrando los valores mínimo, máximo y promedio de viviendas rurales agrupadas en cada cluster. Tal como era de esperar, el valor mínimo en la mayoría de los casos corresponde al umbral de densidad establecido por el modelo (20 viviendas), mientras que el valor máximo varía significativamente entre provincias. Destacan particularmente tres clusters de gran tamaño: uno ubicado en la provincia de Arica, que agrupa 4.352 viviendas correspondientes a el sector de Villa Frontera; otro en la provincia de Copiapó, con un clúster al norte de Caldera correspondiente a los sectores de Los Pulpitos y Los Pulpos que concentra 3.931 viviendas rurales; y el ultimo en la provincia de Elqui correspondiente al sector Cruz de Caña (2.891 viviendas), al sur de la ciudad de Coquimbo.

Figura 29: Rango de viviendas (mínimo y máximo) de clusters por provincia en la Macrozona Norte



Un aspecto relevante que se infiere de estos resultados es que los clusters con mayor volumen de viviendas tienden a localizarse en zonas rurales cercanas a áreas urbanas consolidadas. Este patrón es visible en provincias como Arica (ciudad de Arica), El Loa (Calama), Copiapó (ciudad de Copiapó) y Elqui (conurbación La Serena–Coquimbo). En estos casos, la expansión urbana hacia áreas rurales o periurbanas, junto con una mayor conectividad territorial e infraestructura, favorece la formación de núcleos rurales de alta densidad, los cuales son detectados con mayor facilidad por el modelo.

Con el fin de comprender los resultados de la clusterización a distintas escalas, la Tabla 12 presenta un resumen de los índices previamente mencionados, incluyendo la subclasificación de las viviendas clasificadas como ruido y el índice de clusterización comunal.

En esta se observa que comunas aisladas ubicadas en el altiplano, e incluso en la frontera, alcanzan valores muy altos de clusterización; como es el caso de Ollagüe (88%), Colchane (85%), Camiña y Huara (ambas con 83%). También, se presentan altos índices de clusterización en zonas costeras, donde predominan pequeños asentamientos pesqueros (caso Tocopilla con un 86%) y localidades con asentamientos irregulares en el borde costero de Atacama, en la comuna de Caldera (87%).

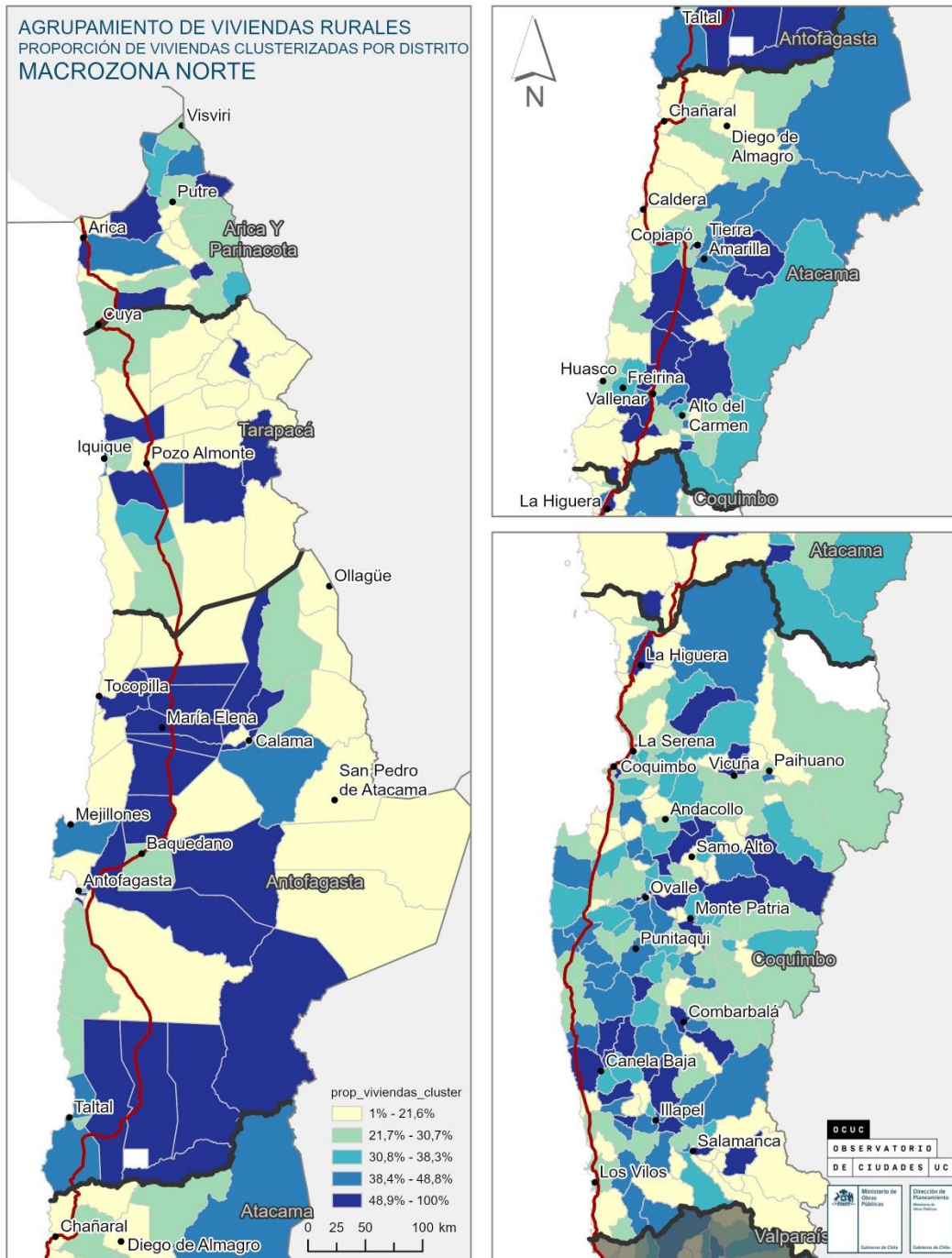
El promedio de clusterización comunal en la macrozona norte alcanza un 74%. Sin embargo, al analizar las áreas urbanas consolidadas —exceptuando Iquique y Antofagasta— todas presentan índices inferiores a este promedio. Esto se explica, probablemente, porque el algoritmo tiende a identificar clústeres periféricos a la mancha urbana, caracterizados por su irregularidad y heterogeneidad, lo que reduce el índice global y evita que toda la periferia sea considerada como un único agrupamiento.

Tabla 12: Indicadores de clusterización a nivel comuna, macrozona norte.

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Total ruido	Viviendas Ruido permanente	Viviendas Ruido Satélites	Índice de clusterización(%)
Arica y Parinacota	Arica	Arica	11.950	7.306	4.644	74	26	61
		Camarones	1.296	956	340	55	45	74
	Parinacota	General Lagos	953	624	329	72	28	66
		Putre	1.754	1.314	440	50	50	75
Tarapacá	Del Tamarugal	Camiña	1.322	1.099	223	42	58	83
		Colchane	2.308	1.964	344	28	72	85
		Huara	3.296	2.725	571	39	61	83
		Pica	1.073	593	480	77	23	55
		Pozo Almonte	8.608	5.441	3.167	76	24	63
	Iquique	Alto Hospicio	1.670	1.417	253	32	68	85
		Iquique	3.310	2.629	681	45	56	79
		Antofagasta	Antofagasta	Antofagasta	1.959	1.561	398	41
Mejillones	1.123			852	271	64	37	76
Sierra Gorda	877			690	187	58	42	79
Taltal	1.258			849	409	61	39	68
El Loa	Calama		6.071	4.003	2.068	76	24	66
	Ollagüe		588	517	71	16	85	88
	San Pedro De Atacama		3.250	2.812	438	28	72	87
Tocopilla	María Elena	267	211	56	30	70	79	

		Tocopilla	2.565	2.193	372	32	68	86	
Atacama	Chañaral	Chañaral	3.019	2.452	567	50	50	81	
		Diego De Almagro	1.183	900	283	64	36	76	
		Caldera	9.633	8.362	1.271	23	77	87	
	Copiapó	Copiapó	6.477	4.142	2.335	75	25	64	
		Tierra Amarilla	2.140	1.517	623	59	41	71	
		Alto Del Carmen	4.031	2.837	1.194	65	35	70	
	Huasco	Freirina	2.598	1.917	681	61	39	74	
		Huasco	3.310	2.626	684	41	59	79	
Vallenar		7.659	4.755	2.904	76	24	62		
Coquimbo	Choapa	Canela	5.296	3.435	1.861	71	29	65	
		Illapel	7.092	4.353	2.739	76	24	61	
		Los Vilos	4.225	3.051	1.174	65	35	72	
		Salamanca	6.432	5.028	1.404	55	45	78	
	Elqui	Andacollo	1.152	849	303	67	33	74	
		Coquimbo	17.609	12.632	4.977	66	34	72	
		La Higuera	3.051	2.485	566	45	55	81	
		La Serena	16.743	12.058	4.685	67	33	72	
		Paiguano	3.869	2.868	1.001	61	39	74	
		Vicuña	6.122	4.581	1.541	62	38	75	
	Limarí	Combarbalá	6.481	4.454	2.027	70	30	69	
		Monte Patria	9.916	6.980	2.936	69	31	70	
		Ovalle	17.183	11.244	5.939	74	26	65	
		Punitaqui	3.994	2.481	1.513	78	22	62	
			Río Hurtado	4.030	2.792	1.238	69	31	69

Cartografía 3: Agrupamiento de viviendas rurales por distrito en la Macrozona Norte



Cartografía 3La Cartografía 3 muestra la proporción de viviendas clusterizadas por distrito censal en la macrozona norte. Los distritos se clasifican según el porcentaje de viviendas agrupadas, destacándose en gris aquellas zonas donde no se identificaron clusters bajo los criterios del modelo. Un aspecto relevante, es que los distritos de la

comuna de María Elena se visualizan en color gris, esto ocurre porque no se registran viviendas rurales clusterizadas.

A nivel comunal, se observan contrastes significativos en el desempeño del modelo. Por un lado, destacan comunas donde el agrupamiento alcanza altos niveles de eficiencia, como Sierra Gorda (79%) y Ollagüe (88%). Este resultado se explica por su baja población rural localizada en sectores altamente concentrados, lo que facilita que el algoritmo identifique clusters claros y diferencie con precisión las viviendas clasificadas como ruido, generalmente más dispersas. Este patrón contribuye a que la Región de Antofagasta presente en su conjunto elevados niveles de clusterización, reflejando la tendencia de la población rural de esta zona a concentrarse en enclaves asociados a faenas mineras y pequeños núcleos tradicionales.

En la Región de Atacama, comunas como Caldera (87%) y Chañaral (81%) también presentan altos índices de clusterización, asociados principalmente a asentamientos localizados en el borde costero, muchos de ellos con características de ocupación irregular. Estos asentamientos se configuran en torno a la proximidad de las áreas urbanas, lo que genera clusters densos y bien definidos, coherentes con la dinámica territorial de ocupación del litoral observada en terreno.

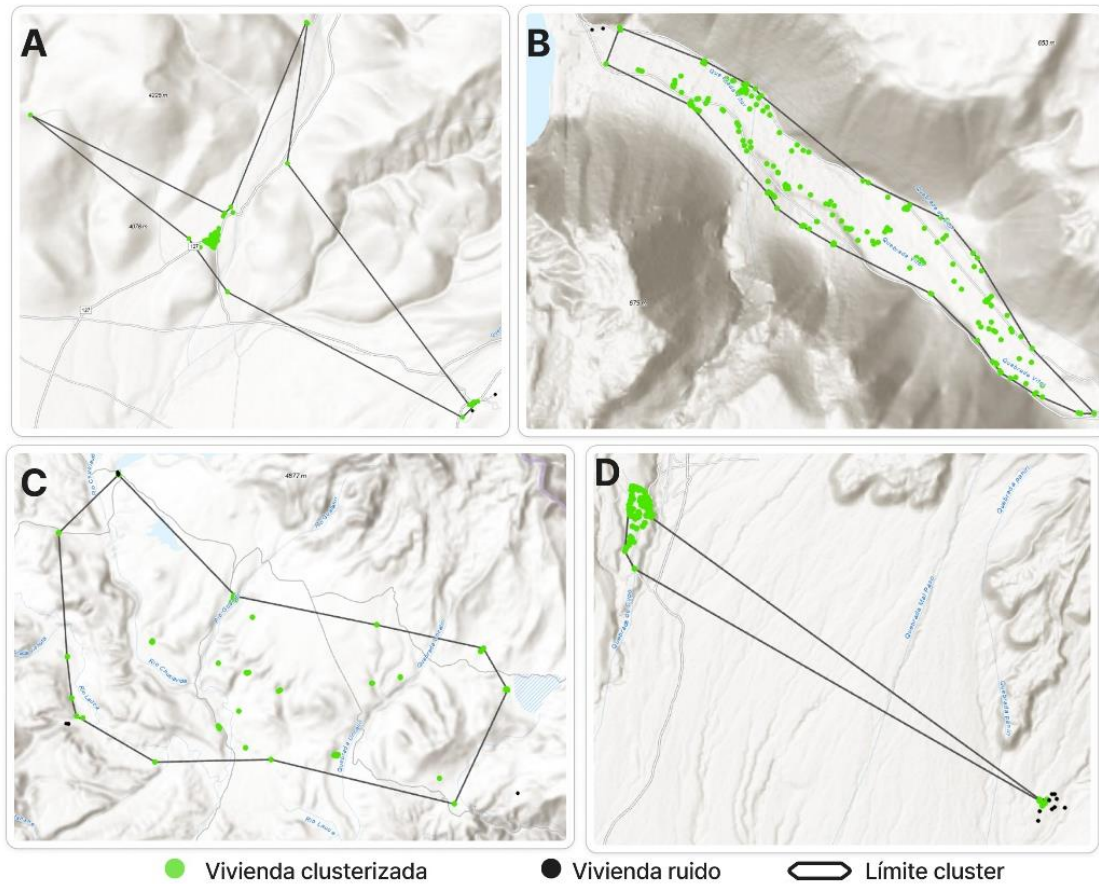
En contraste, los menores índices de clusterización se registran en las comunas de Pozo Almonte (63%) y Pica (55%). En Pozo Almonte, si bien existe una importante cantidad de viviendas rurales, gran parte de ellas se organiza en enclaves concentrados —como La Tirana, Santa Rosa, Santa Emilia Sur, La Huayca y Los Andes— que presentan periferias dispersas clasificadas mayoritariamente como ruido permanente. Estos asentamientos se emplazan a lo largo de la ruta A-665, que conecta con Pica. Esta última comuna, en particular, presenta el índice de clusterización más bajo de la macrozona norte, con apenas nueve agrupamientos identificados, localizados principalmente en la periferia de su núcleo urbano. Además, su condición de comuna altiplánica limítrofe con Bolivia favorece la presencia de asentamientos altamente dispersos, asociados a viviendas aisladas que en su mayoría son clasificadas como ruido permanente, lo que reduce significativamente la capacidad de clusterizar del modelo en este territorio.

Por otra parte, con el fin de identificar clusters robustos y representativos del agrupamiento, se aplicó un método (ver apartado metodológico 4.5) que determina la consistencia de los clusters en función de la desviación estándar del parámetro de densidad local entregado por el algoritmo HDBSCAN.

La Figura 30 ilustra cuatro ejemplos de clusters no consistentes en la macrozona norte. Estos casos se originan principalmente por dos razones: (i) el algoritmo agrupa un polo concentrado de viviendas, pero incluye también viviendas distantes debido a su carácter jerárquico, clusters identificados en terreno como de “tipo flecha” (imágenes A y D); y (ii) el agrupamiento se produce en áreas donde las viviendas cumplen con el

parámetro mínimo (MinPts = 20), pero presentan una distribución dispersa o lineal, sin un patrón claro de concentración, como ocurre en algunos valles (imagen B) o en zonas altiplánicas con viviendas aisladas que alcanzan densidad suficiente (imagen C).

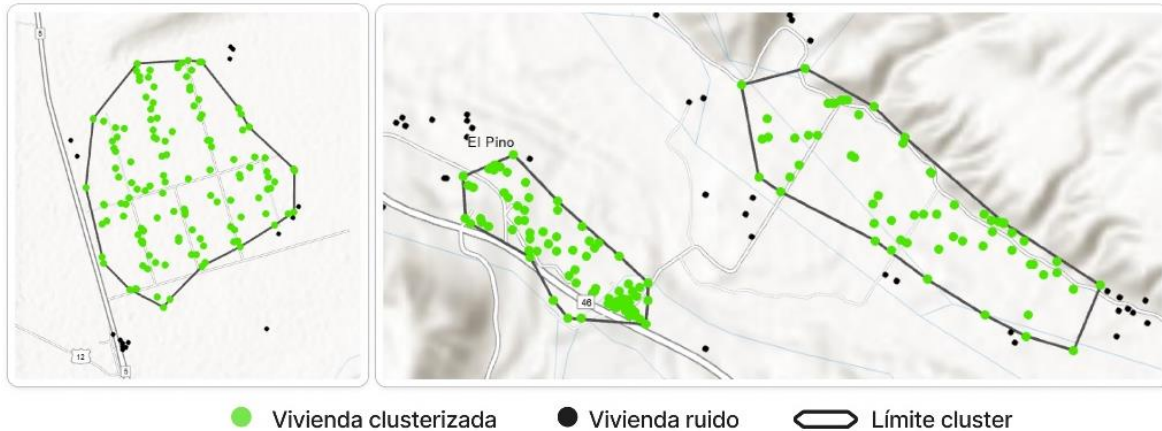
Figura 30: Ejemplos de clúster no consistentes



En contraste, la Figura 31 muestra tres ejemplos de clusters consistentes. Estos se caracterizan por un patrón regular de concentración o semiconcentración, con formas poligonales más definidas y distancias relativamente homogéneas entre viviendas.

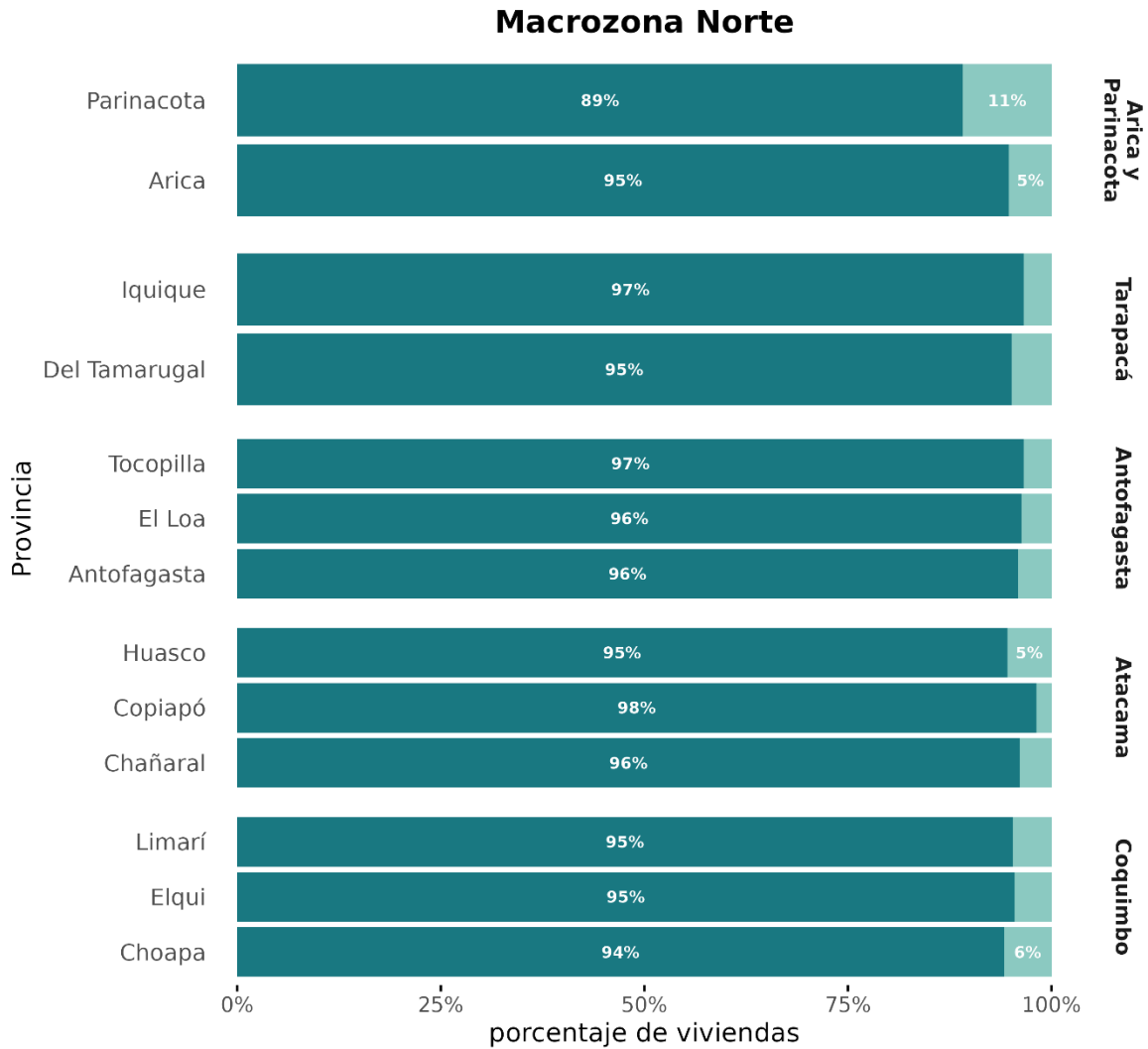
Esta distinción permite diferenciar clusters metodológicamente robustos de aquellos más débiles, entregando a los tomadores de decisiones un insumo clave para focalizar inversiones en territorios donde el agrupamiento de viviendas refleja de manera más fidedigna la dinámica de poblamiento rural.

Figura 31: ejemplos de clúster de tipo consistentes



Así pues, la Figura 32 muestra la proporción de clusters consistentes e inconsistentes en cada provincia de la macrozona. Destacan las provincias de Parinacota y Antofagasta, que presentan los porcentajes más altos de inconsistencia. En el caso de Parinacota, ello se explica por la disposición espacial de las viviendas en el altiplano, donde son escasos los agrupamientos densos y predomina una ocupación más dispersa. En la provincia de Choapa, en cambio, el porcentaje de inconsistencia (6%) responde a que, al no cumplirse el umbral mínimo de 20 viviendas, el algoritmo tiende a incorporar las más cercanas disponibles, generando clusters que abarcan viviendas distantes entre sí, tal como se observa en la Figura 32.

Figura 32: Proporción de tipos de clúster según consistencia en Macrozona Norte.



Tipo de cluster (polígonos) ■ Consistente ■ No Consistente

Las etiquetas de los segmentos menores al 5% fueron removidas para mejorar la visualización del gráfico

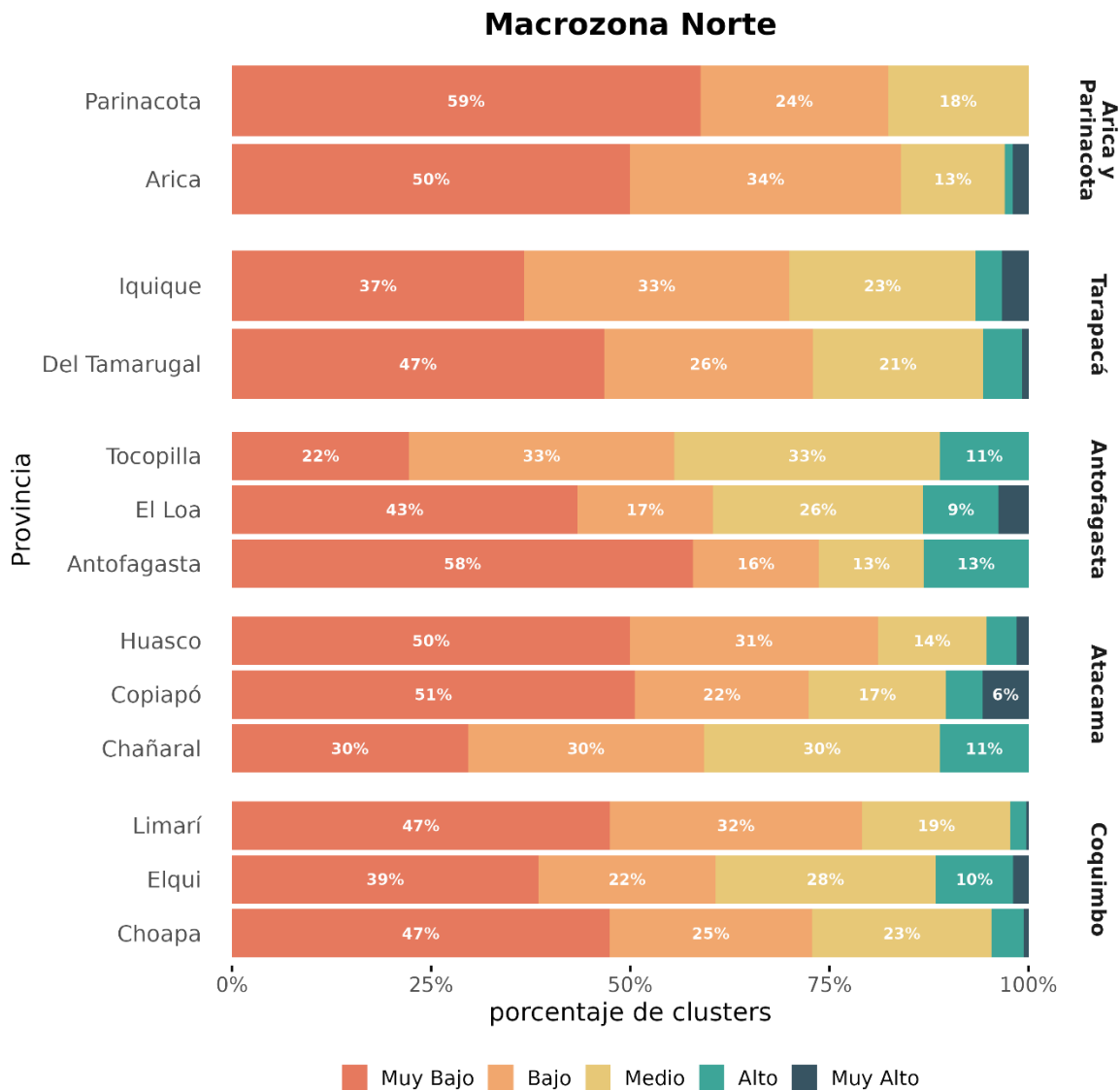
Por otra parte, tal como se mencionó anteriormente, los clusters fueron clasificados según la cantidad de viviendas que agrupan, en base a la tipología de operadores establecida en la Ley 20.998 de SSR, la cual fue adaptada para este estudio. La Figura 33 presenta la distribución de estas categorías a nivel provincial, destacándose varios aspectos relevantes.

En la provincia de Parinacota, la escasa población y la fuerte dispersión de los asentamientos se reflejan en que la mitad de sus clusters corresponden a la categoría de agrupamiento muy bajo (menos de 50 viviendas), sin registrar agrupamientos de tipo alto o muy alto. De manera similar, la provincia de Arica apenas cuenta con dos clusters de agrupamiento muy alto, localizados en la quebrada de Acha y en la periferia norte de la ciudad de Arica. En contraste, la provincia de Iquique concentra la mayor cantidad de

clusters de agrupamiento muy alto (más de 600 viviendas), asociados a tres sectores principales: el campamento Mirando las Estrellas (1.643 viviendas), Chanavayita (835 viviendas) y Los Verdes (731 viviendas).

Cabe destacar que, a nivel general, la Figura 33 evidencia un predominio de clusters de agrupamiento muy bajo (menos de 50 viviendas), lo que se relaciona directamente con la definición del parámetro MinPts de 20 viviendas aplicada en esta versión del algoritmo. Es importante señalar que la elección de este parámetro incide directamente en la proporción de las categorías, por lo que un ajuste en su valor modificaría la distribución observada.

Figura 33: Clasificación de los clústeres según su tamaño de agrupamiento en la Macrozona Norte.



Las etiquetas de los segmentos menores al 5% fueron removidas para mejorar la visualización del gráfico

Por último, la Tabla 13 muestra el detalle de los índices de clasificación de clústeres según su tamaño de agrupamiento y nivel de consistencia, organizados a escala comunal, con el propósito de facilitar la comprensión de los resultados en las distintas escalas de análisis. La comuna de María Elena no presenta resultados dado que fue la única comuna de la macrozona donde no hubo viviendas clusterizadas.

Tabla 13: Indicadores de clasificación de clusters, macrozona norte.

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
Arica y Parinacota	Arica	Arica	83	49,4	32,5	14,5	1,2	2,4	94	6
		Camarones	17	52,9	41,2	5,9	0	0	94,1	5,9
	Parinacota	General Lagos	12	66,7	25	8,3	0	0	91,7	8,3
		Putre	22	54,5	22,7	22,7	0	0	90,9	9,1
Tarapacá	Del Tamarugal	Camíña	12	33,3	41,7	16,7	8,3	0	91,7	8,3
		Colchane	21	47,6	33,3	9,5	9,5	0	95,2	4,8
		Huara	36	47,2	27,8	25	0	0	94,4	5,6
		Pica	10	60	30	10	0	0	90	10
	Iquique	Pozo Almonte	43	46,5	16,3	27,9	7	2,3	93	7
		Alto Hospicio	8	37,5	50	0	0	12,5	87,5	12,5
Antofagasta	Antofagasta	Iquique	22	36,4	27,3	31,8	4,5	0	90,9	9,1
		Antofagasta	14	57,1	21,4	7,1	14,3	0	92,9	7,1
		Mejillones	7	42,9	14,3	28,6	14,3	0	85,7	14,3
		Sierra Gorda	4	50	0	0	50	0	75	25
	El Loa	Taltal	13	69,2	15,4	15,4	0	0	92,3	7,7
		Calama	32	46,9	15,6	25	9,4	3,1	93,8	6,2
		Ollagüe	5	40	20	40	0	0	80	20
	Tocopilla	San Pedro De Atacama	16	37,5	18,8	25	12,5	6,2	93,8	6,2
María Elena		1	0	0	100	0	0	100	0	
Atacama	Chañaral	Tocopilla	17	23,5	35,3	29,4	11,8	0	94,1	5,9
		Chañaral	19	31,6	26,3	31,6	10,5	0	94,7	5,3
	Copiapó	Diego De Almagro	8	25	37,5	25	12,5	0	87,5	12,5
		Caldera	25	52	4	16	12	16	92	8
		Copiapó	43	46,5	32,6	18,6	0	2,3	93	7
	Huasco	Tierra Amarilla	19	57,9	21,1	15,8	5,3	0	94,7	5,3
		Alto Del Carmen	46	54,3	37	6,5	2,2	0	93,5	6,5
		Freirina	22	50	18,2	31,8	0	0	90,9	9,1
		Huasco	19	36,8	31,6	15,8	10,5	5,3	94,7	5,3
	Coquimbo	Choapa	Vallenar	45	51,1	31,1	11,1	4,4	2,2	93,3
Canela			36	55,6	16,7	25	0	2,8	94,4	5,6
Illapel			65	52,3	30,8	15,4	1,5	0	93,8	6,2
Los Vilos			29	48,3	13,8	31	6,9	0	93,1	6,9
Elqui		Salamanca	43	32,6	32,6	25,6	9,3	0	93	7
		Andacollo	9	44,4	44,4	0	11,1	0	88,9	11,1
		Coquimbo	79	38	26,6	22,8	8,9	3,8	94,9	5,1
		La Higuera	18	50	11,1	22,2	16,7	0	94,4	5,6
		La Serena	84	36,9	14,3	34,5	13,1	1,2	94	6
		Paiguano	28	42,9	21,4	32,1	0	3,6	92,9	7,1
Limarí		Vicuña	39	33,3	30,8	28,2	7,7	0	94,9	5,1
		Combarbalá	43	39,5	25,6	30,2	4,7	0	93	7
		Monte Patria	77	44,2	27,3	24,7	3,9	0	94,8	5,2
Punitaqui	Ovalle	161	51,6	33,5	13,7	0,6	0,6	95	5	
	Punitaqui	40	52,5	30	17,5	0	0	95	5	

### 6.3. Macrozona Centro

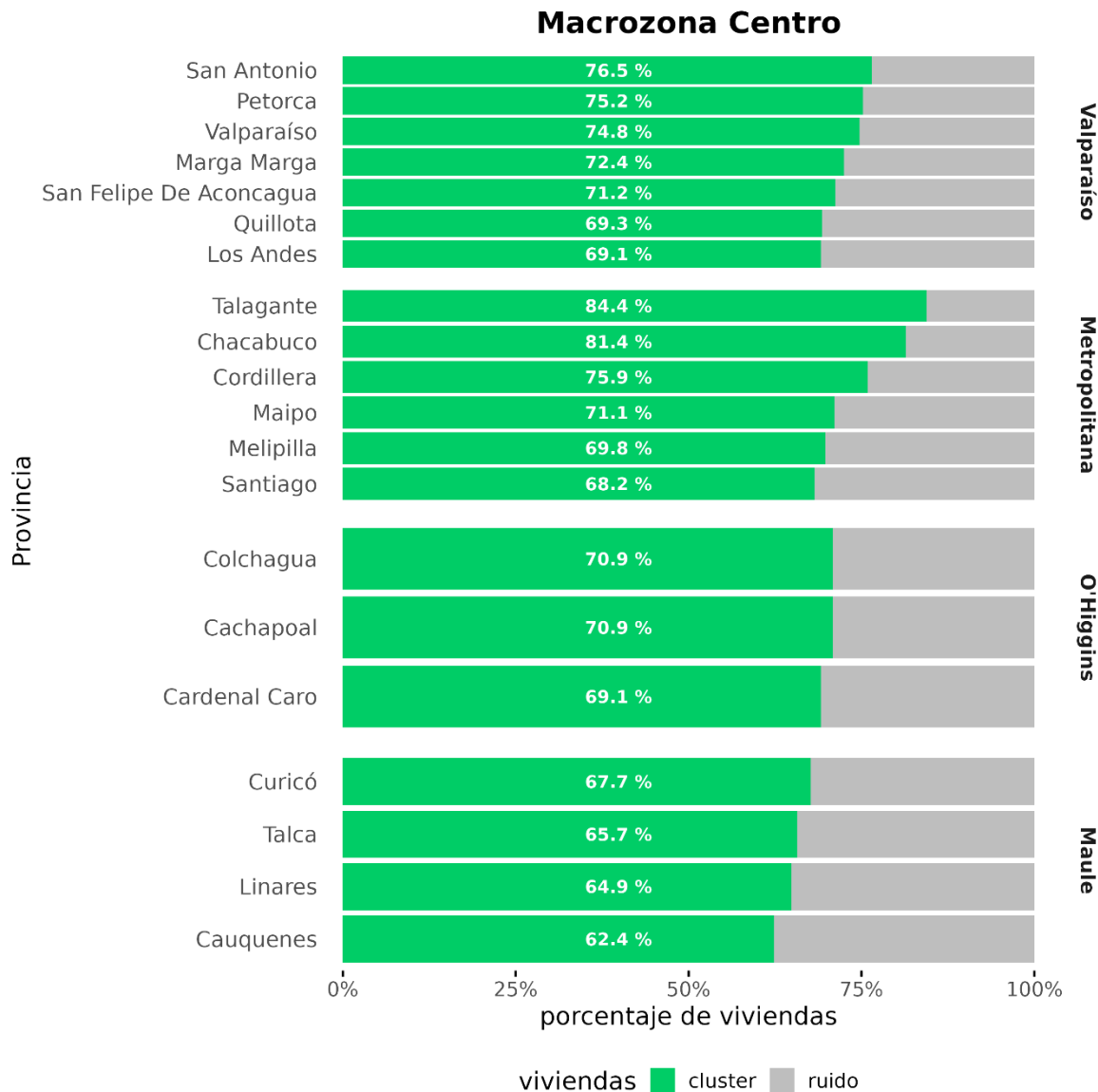
La macrozona centro del país está compuesta por las regiones de Valparaíso, Metropolitana, O'Higgins y del Maule, las que, según el Censo de Población y Vivienda del año 2024, en conjunto concentran casi el 60% de todas las viviendas a nivel nacional (INE, 2025). Esta condición releva la presión que posee este territorio frente a los diversos tipos de asentamientos humanos rurales, los cuales se distribuyen de forma menos atomizada a lo largo de los valles centrales debido a la ausencia de quebradas presentes en el extremo norte del país.

En total, en la macrozona centro se agruparon 429.543 viviendas rurales (70,69%), en 4.223 clusters o polígonos de agrupamiento. Resultando el Maule (1.493) y O'Higgins (1.103) las regiones que presentan los mayores números de clústeres, mientras que la región Metropolitana (848) y Valparaíso (779), son las que menos. Asimismo, el índice de clusterización se distribuye de forma similar en las regiones de Valparaíso (73%), Metropolitana (75%) y O'Higgins (70%), mientras que Maule es la que menos porcentaje de viviendas agrupadas posee, alcanzando un 66%.

Al desagregar estos porcentajes a escala provincia, se observan algunas diferencias intrarregionales que son importantes de destacar, las que evidencian el comportamiento propio de los clústeres en esta macrozona. En general, la proporción de viviendas clusterizadas (Figura 34) es bastante alta, desde la provincia de Talagante en la región Metropolitana que alcanza un 84,4%, hasta la provincia de Cauquenes en la región del Maule, donde disminuye a un 62,4%. Además, provincias de la región Metropolitana que se localizan en sectores del valle Central (Talagante y Maipo) tienen un porcentaje de viviendas clusterizadas muy similar (desde el 84,4% y 71,1%). En cambio, la provincia de Cordillera, además de emplazarse en su mayoría en el sector cordillerano de los Andes, tiene la característica que sus viviendas se tienden a concentrar en sectores más estrechos como el Cajón del Maipo, logrando un alto porcentaje de 75,9%. Por otro lado, a nivel de comuna, destaca la comuna de Melipilla, presentando el mayor número de viviendas clusterizadas (17.182).

Estos resultados se relacionan directamente por las características morfológicas propias de cada región. En concreto, a diferencia de la macrozona norte, que en regiones marcadas por un relieve de quebradas o valles muy estrechos (ya sean cordilleranos, altiplánicos o costeros) se evidencia un aumento del porcentaje de viviendas clusterizadas, en la macrozona centro el relieve de valle presente en la mayor parte de la superficie, permite que las viviendas se distribuyan de forma más homogénea y menos atomizada, lo que causa que el porcentaje de viviendas agrupadas tienda a ser un poco más bajo.

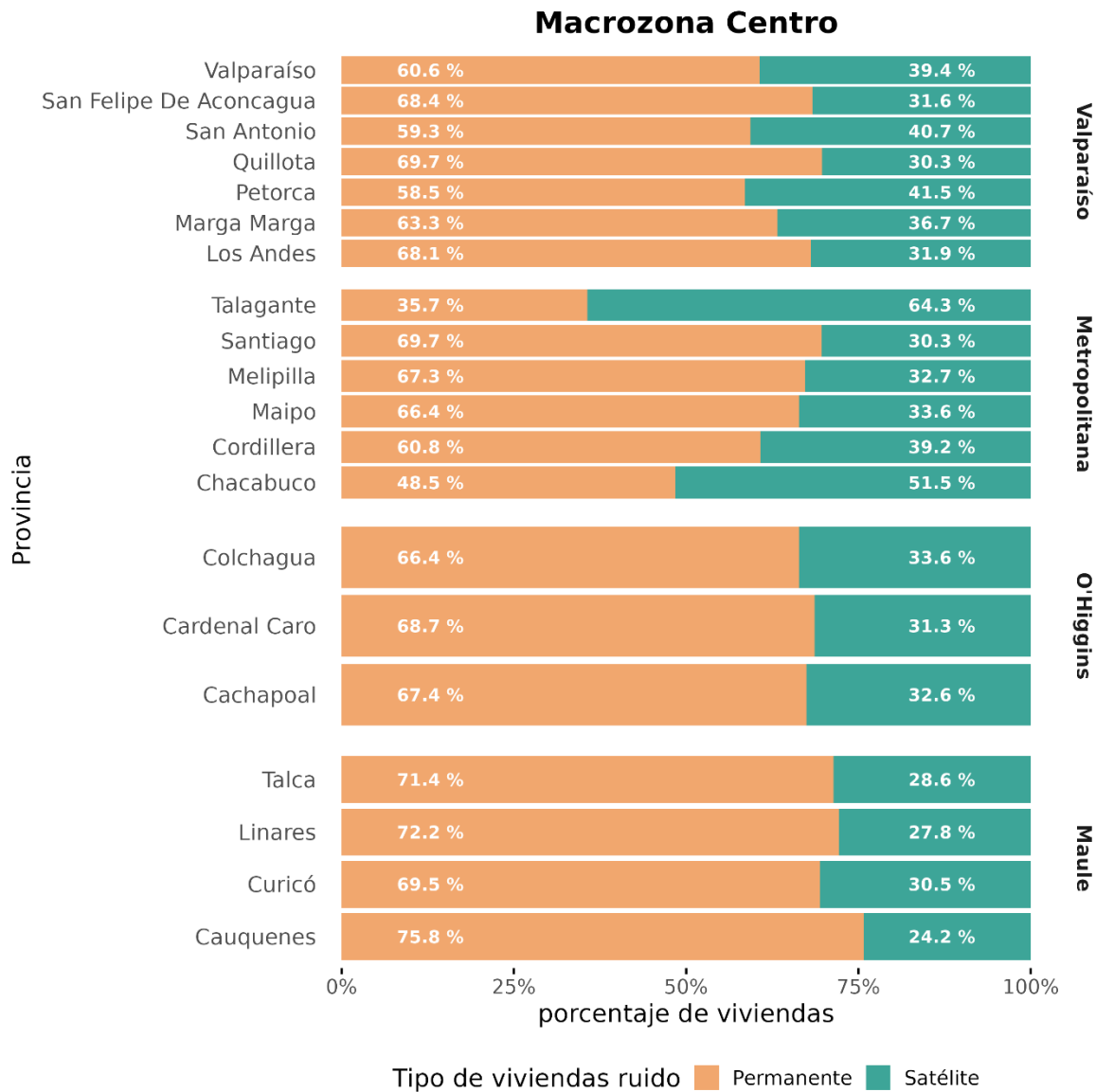
Figura 34: Proporción de viviendas clusterizadas por provincia en la Macrozona Centro.



Al visualizar los resultados obtenidos de viviendas ruido (Figura 35) en detalle, se evidencian diferencias entre las regiones. En el caso de la región Metropolitana, destaca el caso de la provincia de Talagante, que presenta un elevado porcentaje de viviendas ruido catalogadas como “satélite” dado que, por densidad y distancia, se asocian mayormente a las agrupaciones, siendo un 35,7% únicamente viviendas permanentes. En cambio, la región del Maule en la provincia de Cauquenes es la que presenta más viviendas ruido de tipo permanente, que satélites (75,8%). Estos resultados responden al alto nivel de agrupamiento obtenido en Talagante (84,4%) y por la forma en que se distribuyen las viviendas en el territorio (definida principalmente por los límites de los predios y parcelas), generando que las viviendas que no cumplen con los criterios del

modelo queden sin clusterización pero siendo sitios potenciales de agruparse en caso de modificar la métrica *MinPts*.

Figura 35: Proporción de viviendas ruido según su tipo en la Macrozona Centro.

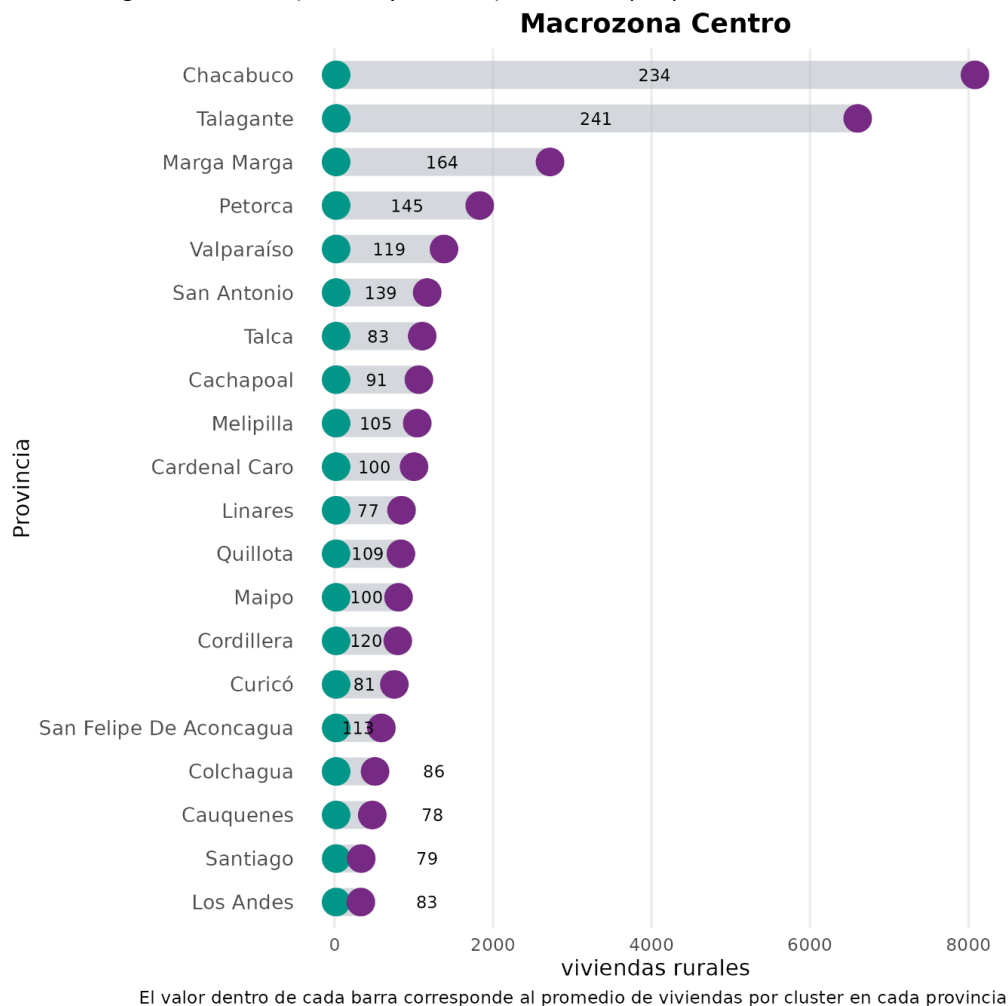


Otro aspecto relevante de analizar es el tamaño de los clústeres resultantes en términos de la cantidad total de viviendas que agrupa cada uno, los cuales pueden variar dependiendo de la macrozona en que se encuentran, y de la distribución espacial de las viviendas en el territorio. En efecto, en la Figura 36 se observa que todas las provincias poseen el mismo mínimo de viviendas clusterizadas, que coincide en la mayoría de los casos con el rango dado por el modelo (*MinPts* de 20 viviendas). Además, los clusters presentes en las provincias de Los Andes y Santiago poseen tamaños similares entre sí, donde el rango mínimo y máximo de viviendas es pequeño. Además, el promedio de

viviendas adquiere valores cercanos, entre 81 (provincia de Curicó) a 234 (provincia de Chacabuco). No obstante, el máximo difiere de forma considerable, pasando de 8.082 (provincia de Chacabuco) a 333 (provincia Los Andes) viviendas rurales agrupadas.

Por ejemplo, Cachapoal, Melipilla y Cadernal Caro son provincias cuyos clusters se asemejan entre sí en cuanto a su tamaño, donde el promedio es de 91-105 viviendas por clusters, reflejando que todos los clusters presentes en esta provincia se componen de un total de viviendas muy similar, no existiendo grupos con valores excesivamente altos. No obstante, llama la atención que el clúster con mayor tamaño es de 8.082 viviendas rurales en la provincia de Chacabuco, el cual es un valor que se escapa completamente al valor promedio. Este resultado en particular puede deberse a la forma en que se establecen las viviendas rurales en esta provincia, que son numerosas y ordenadas de forma continua a lo largo de los caminos, principalmente como parcelas o predios de agrado, sumado a la inexistencia de barreras naturales que limiten la agrupación de un mayor número de viviendas. Asimismo, a nivel de comuna, destaca Talagante con el mayor índice de clusterización de 90% en la macrozona centro.

Figura 36: Rango de viviendas (mínimo y máximo) de clusters por provincia en la Macrozona Centro.



El resumen de los índices previamente mencionados, incluyendo la subclasificación de las viviendas clasificadas como ruido y el índice de clusterización comunal se presenta en la Tabla 14. En la tabla se evidencia que es la comuna de Melipilla la que presenta la mayor cantidad de viviendas rurales en total (25.783), viviendas clusterizadas (17.182) y viviendas consideradas como ruido (8.601). Asimismo, destaca el caso de la comuna de Talagante, ya que es la única comuna que presenta el mayor porcentaje de viviendas ruido satélite (100%) e índice de clusterización (90). Por último, las comunas La Florida, Huechuraba, Puente Alto, Peñalolén, Las Condes y La Reina son las únicas que alcanzan el 100% de viviendas ruido clasificadas como “permanente”, dado que por densidad y distancia nunca se lograron clusterizar bajo los criterios del modelo de agrupamiento (MinPts = 20).

Tabla 14: Indicadores de clusterización a nivel comuna, macrozona centro.

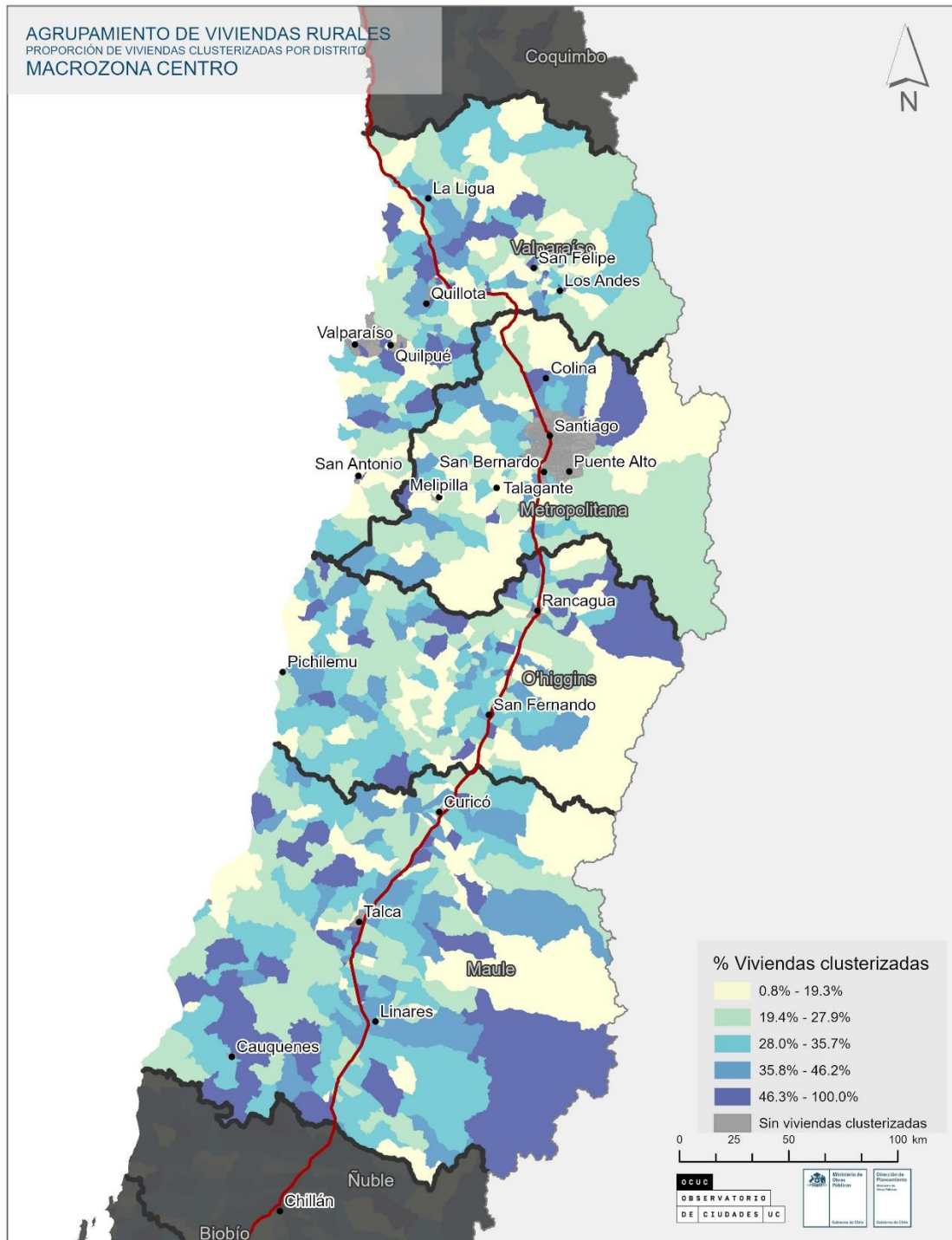
Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Total ruido	Viviendas Ruido permanente	Viviendas Ruido Satélites	Índice de clusterización	
Valparaíso	Los Andes	Calle Larga	2.004	1.475	529	67	33	74	
		Los Andes	1.919	1.291	628	72	28	67	
		Rinconada	1.176	769	407	69	31	65	
		San Esteban	4.019	2.766	1.253	66	34	69	
	Marga Marga	Limache	14.563	9.695	4.868	70	30	67	
		Olmué	4.059	3.298	761	46	54	81	
		Quilpué	5.067	4.228	839	34	66	83	
		Villa Alemana	1.497	1.026	471	69	31	69	
	Petorca	Cabildo	3.511	2.584	927	61	39	74	
		La Ligua	21.914	16.554	5.360	59	41	76	
		Papudo	1.289	921	368	60	40	72	
		Petorca	3.421	2.442	979	63	37	71	
	Quillota	Zapallar	2.859	2.319	540	43	57	81	
		Calera	1.279	1.145	134	5	96	90	
		Hijuelas	2.890	2.315	575	54	46	80	
		La Cruz	1.346	1.103	243	40	61	82	
	San Antonio	Nogales	2.325	1.245	1.080	81	19	54	
		Quillota	5.517	3.444	2.073	76	24	62	
		Algarrobo	3.775	2.560	1.215	75	25	68	
		Cartagena	1.299	987	312	57	43	76	
		El Quisco	1.245	835	410	68	32	67	
		El Tabo	1.619	1.388	231	32	68	86	
	San Felipe De Aconcagua	San Antonio	4.158	3.388	770	44	56	82	
		Santo Domingo	4.205	3.315	890	55	45	79	
		Catemu	2.993	2.127	866	69	31	71	
		Llailay	4.593	3.335	1.258	67	33	73	
		Panquehue	1.504	1.036	468	67	33	69	
		Putauendo	4.502	3.082	1.420	73	27	69	
	Valparaíso	San Felipe	2.936	2.196	740	63	37	75	
		Santa María	2.312	1.645	667	66	34	71	
		Casablanca	6.690	4.941	1.749	64	36	74	
		Concón	1.740	1.366	374	52	48	79	
		Juan Fernández	325	178	147	76	25	55	
		Puchuncaví	3.793	2.883	910	56	45	76	
	Metropolitana	Chacabuco	Quintero	4.182	2.904	1.278	68	32	69
			Valparaíso	3.442	2.808	634	47	53	82
Colina			12.115	10.647	1.468	17	83	88	

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Total ruido	Viviendas Ruido permanente	Viviendas Ruido Satélites	Índice de clusterización
	Cordillera	Lampa	11.827	9.313	2.514	56	44	79
		Tiltil	4.239	2.981	1.258	69	31	70
		Pirque	5.974	4.508	1.466	63	37	76
		Puente Alto	16	0	16	100	0	0
	Maipo	San José De Maipo	4.152	3.194	958	56	44	77
		Buín	5.431	4.250	1.181	57	43	78
		Calera De Tango	4.733	3.259	1.474	69	31	69
		Paine	17.961	12.817	5.144	67	33	71
	Melipilla	San Bernardo	2.691	1.569	1.122	72	28	58
		Alhué	2.420	2.041	379	30	70	84
		Curacaví	7.493	6.003	1.490	49	51	80
		María Pinto	3.541	2.493	1.048	69	31	70
		Melipilla	25.783	17.182	8.601	70	30	67
	Santiago	San Pedro	11.630	7.780	3.850	71	29	67
		Huechuraba	47	0	47	100	0	0
		La Florida	91	0	91	100	0	0
		La Reina	3	0	3	100	0	0
		Las Condes	4	0	4	100	0	0
		Lo Barnechea	1.574	1.280	294	44	57	81
		Maipú	1.368	926	442	71	29	68
		Peñalolén	7	0	7	100	0	0
		Pudahuel	2.870	1.860	1.010	74	26	65
	Talagante	Quilicura	279	193	86	48	52	69
		El Monte	2.664	2.006	658	61	39	75
		Isla De Maipo	4.058	3.331	727	42	58	82
		Padre Hurtado	3.429	2.982	447	24	76	87
		Peñaflor	3.406	2.705	701	51	49	79
Talagante		7.582	6.822	760	0	100	90	
O'Higgins	Cachapoal	Codegua	2.689	1.787	902	68	32	67
		Coinco	2.303	1.505	798	70	30	65
		Coltauco	5.312	3.924	1.388	67	33	74
		Doñihue	2.931	2.405	526	49	52	82
		Graneros	1.983	1.428	555	63	37	72
		Las Cabras	11.287	7.602	3.685	70	30	67
		Machalí	1.618	1.123	495	69	31	69
		Malloa	4.232	2.835	1.397	74	26	67
		Mostazal	3.272	2.365	907	70	30	72
		Olivar	1.788	1.392	396	50	50	78
		Peumo	1.143	837	306	54	46	73
		Pichidegua	4.873	3.600	1.273	64	36	74
		Quinta De Tilcoco	2.562	1.760	802	74	26	69
		Rancagua	5.545	3.992	1.553	63	38	72
		Rengo	6.027	4.620	1.407	62	38	77
		Requinoa	5.416	3.794	1.622	70	30	70
		San Vicente	9.973	6.719	3.254	71	30	67
	Cardenal Caro	La Estrella	3.054	2.155	899	64	36	71
		Litueche	4.436	3.296	1.140	63	37	74
		Marchihue	3.561	2.426	1.135	70	30	68
		Navidad	6.658	4.103	2.555	76	24	62
		Paredones	7.382	5.296	2.086	63	37	72
	Colchagua	Pichilemu	4.642	3.273	1.369	70	30	71
		Chimbarongo	6.007	4.306	1.701	65	35	72
		Chépica	2.867	2.031	836	69	31	71

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Total ruido	Viviendas Ruido permanente	Viviendas Ruido Satélites	Índice de clusterización
		Lolol	2.589	1.728	861	70	30	67
		Nancagua	3.149	2.217	932	66	34	70
		Palmilla	4.169	3.263	906	53	47	78
		Peralillo	2.808	2.038	770	66	35	73
		Placilla	2.927	2.128	799	67	33	73
		Pumanque	2.719	1.878	841	70	30	69
		San Fernando	5.442	3.705	1.737	69	31	68
		Santa Cruz	7.268	5.008	2.260	68	32	69
Maule	Cauquenes	Cauquenes	6.639	3.869	2.770	78	22	58
		Chanco	2.589	1.587	1.002	75	25	61
		Pelluhue	3.445	2.450	995	70	30	71
	Curicó	Curicó	8.498	5.420	3.078	74	26	64
		Hualañé	2.729	1.839	890	70	30	67
		Licantén	987	653	334	61	40	66
		Molina	4.749	3.610	1.139	60	40	76
		Rauco	2.639	1.733	906	71	29	66
		Romerol	5.388	3.289	2.099	75	25	61
		Sagrada Familia	5.024	3.594	1.430	62	38	72
		Teno	9.915	6.862	3.053	69	31	69
		Vichuquén	3.888	2.659	1.229	67	33	68
		Linares	Colbún	8.161	5.850	2.311	62	38
	Linares		11.741	7.113	4.628	78	22	61
	Longaví		10.915	6.684	4.231	76	24	61
	Parral		7.425	4.923	2.502	68	32	66
	Retiro		7.668	5.065	2.603	74	26	66
	San Javier		8.720	5.917	2.803	68	32	68
	Villa Alegre		3.934	2.522	1.412	72	28	64
	Yerbas Buenas		6.557	4.158	2.399	72	28	63
	Talca	Constitución	9.493	6.755	2.738	64	36	71
		Curepto	4.813	3.240	1.573	70	30	67
		Empedrado	1.067	625	442	77	23	59
		Maule	8.145	5.023	3.122	77	23	62
		Talca	6.115	4.159	1.956	70	30	68
		Pencahue	3.629	2.494	1.135	71	29	69
		Pelarco	3.590	2.176	1.414	75	25	61
		San Rafael	3.189	1.898	1.291	80	20	60
San Clemente		15.024	9.715	5.309	70	30	65	
		Río Claro	5.015	3.386	1.629	69	31	68

Bajo el objetivo de visualizar geoespacialmente las agrupaciones de viviendas rurales, se genera la Cartografía 4, que muestra cómo es la distribución de los clústeres según la proporción de viviendas en toda la macrozona centro y en función del distrito censal.

Cartografía 4: Agrupamiento de viviendas rurales Macrozona Centro



Por último, la Tabla 15 muestra el detalle de los índices de clasificación de clústeres según su tamaño de agrupamiento y nivel de consistencia, organizados a escala comunal, con el propósito de facilitar la comprensión de los resultados en las distintas escalas de análisis. En la tabla se evidencia que es la comuna de Melipilla la que presenta

la mayor cantidad de clusters (183), además, las comunas de Retiro y Yerbabuenas (provincia de Linares), son las que presentan un mayor porcentaje de consistencia (95,1% de consistencia) y en menor medida la comuna de Juan Fernández (50% de inconsistencia), aunque en este último caso se da por la baja presencia de clusters (únicamente 2). Respecto a las clasificaciones según el tamaño de los clústeres, las que presentan un mayor son: la comuna de Empedrado con “Muy bajo” (75%), Quilicura con “Bajo” (66,7%), Juan Fernández con “Medio” (50%), Calera con “Alto” (66,7%) y Talagante con “Muy alto” (33,3%), evidenciando que en general, en la macrozona centro los clusters poseen bajos tamaños (entre bajo a alto).

Tabla 15: Indicadores de clasificación a nivel de clusters, macrozona centro.

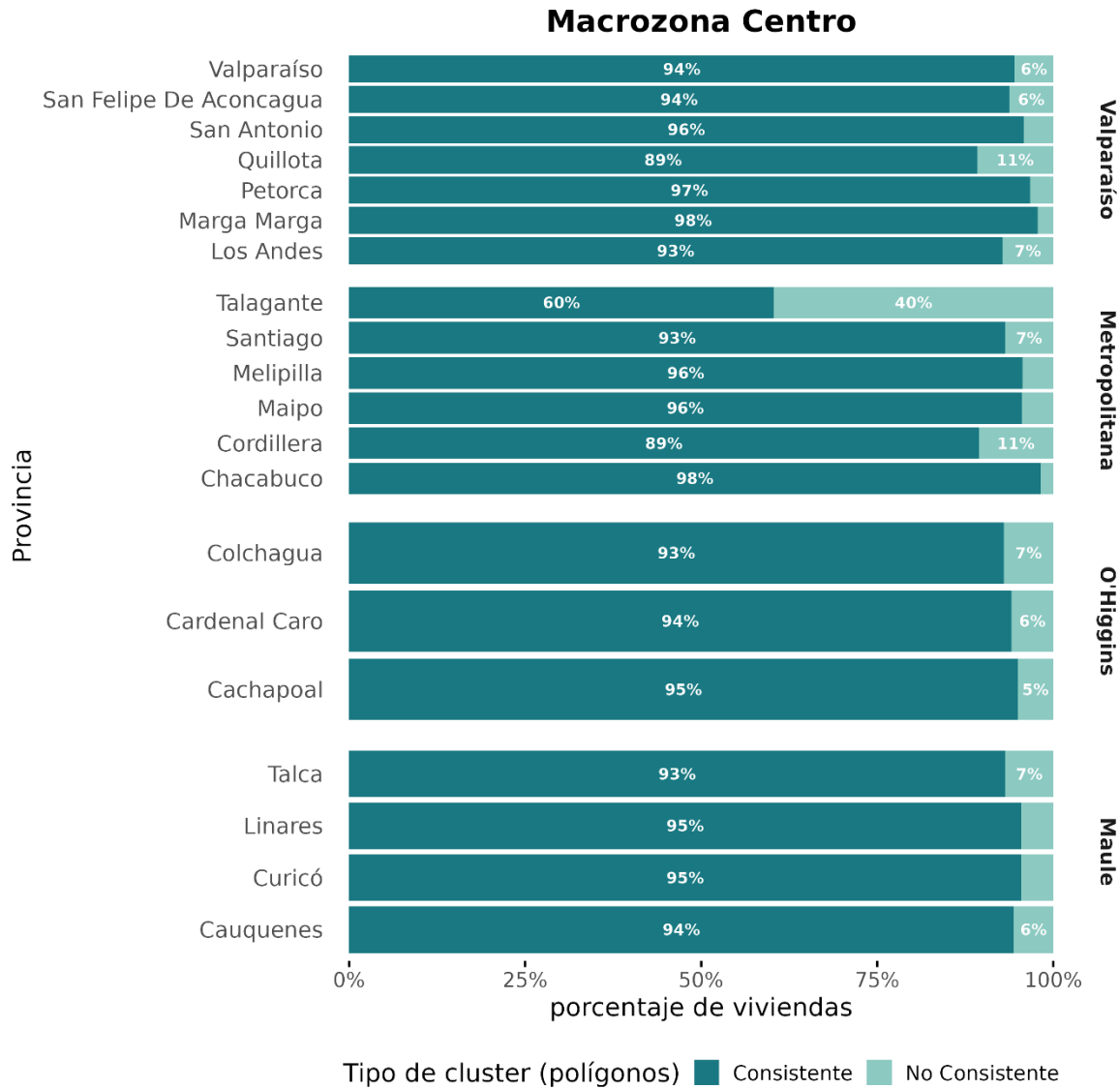
Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
Valparaíso	Los Andes	Calle Larga	18	44,4	27,8	22,2	5,6	0	94,4	5,6
		Los Andes	19	63,2	21,1	10,5	5,3	0	94,7	5,3
		Rinconada	9	44,4	22,2	33,3	0	0	88,9	11,1
		San Esteban	30	33,3	26,7	40	0	0	93,3	6,7
	Marga Marga	Limache	73	41,1	34,2	15,1	6,8	2,7	94,5	5,5
		Olmué	12	50	33,3	8,3	0	8,3	91,7	8,3
		Quilpué	15	40	26,7	20	6,7	6,7	93,3	6,7
		Villa Alemana	11	36,4	18,2	45,5	0	0	90,9	9,1
	Petorca	Cabildo	29	41,4	24,1	31	3,4	0	93,1	6,9
		La Ligua	98	36,7	25,5	22,4	10,2	5,1	94,9	5,1
		Papudo	9	22,2	55,6	22,2	0	0	88,9	11,1
		Petorca	24	37,5	33,3	25	4,2	0	91,7	8,3
		Zapallar	11	36,4	18,2	27,3	0	18,2	90,9	9,1
	Quillota	Calera	3	0	0	33,3	66,7	0	66,7	33,3
		Hijuelas	15	33,3	26,7	26,7	13,3	0	93,3	6,7
		La Cruz	7	57,1	28,6	0	0	14,3	85,7	14,3
		Nogales	24	58,3	29,2	12,5	0	0	91,7	8,3
		Quillota	36	36,1	30,6	33,3	0	0	94,4	5,6
	San Antonio	Algarrobo	17	29,4	41,2	11,8	11,8	5,9	94,1	5,9
		Cartagena	8	12,5	37,5	50	0	0	87,5	12,5
		El Quisco	13	46,2	30,8	23,1	0	0	92,3	7,7
		El Tabo	6	16,7	16,7	50	0	16,7	83,3	16,7
		San Antonio	24	33,3	37,5	20,8	4,2	4,2	91,7	8,3
		Santo Domingo	22	40,9	27,3	13,6	18,2	0	90,9	9,1
	San Felipe De Aconcagua	Catemu	22	40,9	27,3	27,3	4,5	0	90,9	9,1
		Llaillay	25	36	36	16	12	0	92	8
		Panquehue	9	44,4	11,1	44,4	0	0	88,9	11,1
		Putendo	28	35,7	35,7	21,4	7,1	0	92,9	7,1
		San Felipe	20	35	30	30	5	0	95	5
		Santa María	15	26,7	40	26,7	6,7	0	93,3	6,7
Valparaíso	Casablanca	45	33,3	26,7	31,1	8,9	0	93,3	6,7	
	Concón	9	11,1	22,2	44,4	22,2	0	88,9	11,1	
	Juan Fernández	2	0	50	50	0	0	50	50	
	Puchuncaví	28	53,6	21,4	17,9	3,6	3,6	92,9	7,1	
	Quintero	28	28,6	35,7	28,6	7,1	0	92,9	7,1	
	Valparaíso	15	53,3	26,7	6,7	0	13,3	93,3	6,7	
Metropolitana	Chacabuco	Colina	19	26,3	36,8	15,8	10,5	10,5	94,7	5,3
		Lampa	46	37	23,9	28,3	8,7	2,2	93,5	6,5
		Tiltil	33	42,4	33,3	21,2	3	0	93,9	6,1

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
	Cordillera	Pirque	31	32,3	32,3	22,6	6,5	6,5	93,5	6,5
		San José De Maipo	33	42,4	30,3	21,2	3	3	93,9	6,1
	Maipo	Buín	45	31,1	35,6	31,1	2,2	0	93,3	6,7
		Calera De Tango	30	46,7	20	30	0	3,3	93,3	6,7
		Paine	120	40,8	25,8	28,3	2,5	2,5	95	5
		San Bernardo	25	52	28	20	0	0	92	8
	Melipilla	Alhué	13	38,5	7,7	38,5	15,4	0	92,3	7,7
		Curacaví	30	30	23,3	20	13,3	13,3	93,3	6,7
		María Pinto	32	46,9	28,1	21,9	3,1	0	93,8	6,2
		Melipilla	183	42,6	30,1	22,4	3,8	1,1	94,5	5,5
		San Pedro	80	43,8	32,5	16,2	6,2	1,2	95	5
	Santiago	Lo Barnechea	14	35,7	28,6	35,7	0	0	92,9	7,1
		Maipú	9	44,4	22,2	22,2	11,1	0	88,9	11,1
		Pudahuel	28	57,1	25	17,9	0	0	92,9	7,1
		Quilicura	3	33,3	66,7	0	0	0	66,7	33,3
	Talagante	El Monte	23	34,8	34,8	30,4	0	0	91,3	8,7
		Isla De Maipo	18	22,2	27,8	27,8	16,7	5,6	94,4	5,6
		Padre Hurtado	15	13,3	20	46,7	20	0	93,3	6,7
		Peñaflor	15	6,7	53,3	26,7	6,7	6,7	93,3	6,7
		Talagante	3	33,3	0	33,3	0	33,3	66,7	33,3
O'Higgins	Cachapoal	Codegua	28	50	35,7	14,3	0	0	92,9	7,1
		Coinco	17	29,4	47,1	23,5	0	0	94,1	5,9
		Coltauco	37	45,9	16,2	35,1	2,7	0	94,6	5,4
		Doñihue	9	11,1	44,4	22,2	0	22,2	88,9	11,1
		Graneros	17	35,3	17,6	47,1	0	0	94,1	5,9
		Las Cabras	92	50	29,3	17,4	2,2	1,1	94,6	5,4
		Machalí	15	46,7	26,7	26,7	0	0	93,3	6,7
		Malloa	35	34,3	37,1	28,6	0	0	94,3	5,7
		Mostazal	23	34,8	26,1	39,1	0	0	91,3	8,7
		Olivar	12	25	58,3	8,3	0	8,3	91,7	8,3
		Peumo	12	41,7	41,7	16,7	0	0	91,7	8,3
		Pichidegua	40	47,5	25	27,5	0	0	95	5
		Quinta De Tilcoco	18	38,9	27,8	27,8	5,6	0	94,4	5,6
		Rancagua	42	50	23,8	21,4	2,4	2,4	92,9	7,1
		Rengo	43	44,2	30,2	18,6	4,7	2,3	93	7
		Requínoa	46	50	32,6	15,2	0	2,2	93,5	6,5
	San Vicente	80	42,5	32,5	22,5	2,5	0	95	5	
	Cardenal Caro	La Estrella	25	52	24	16	8	0	92	8
		Litueche	30	23,3	40	33,3	3,3	0	93,3	6,7
		Marchihue	27	40,7	37	18,5	3,7	0	92,6	7,4
		Navidad	42	59,5	19	14,3	2,4	4,8	92,9	7,1
		Paredones	46	47,8	19,6	26,1	2,2	4,3	93,5	6,5
		Pichilemu	36	44,4	25	25	5,6	0	94,4	5,6
	Colchagua	Chimbarongo	44	31,8	43,2	22,7	2,3	0	93,2	6,8
		Chépica	30	43,3	36,7	20	0	0	93,3	6,7
		Lolol	28	50	35,7	14,3	0	0	92,9	7,1
		Nancagua	25	36	40	16	8	0	92	8
		Palmilla	32	28,1	31,2	37,5	3,1	0	93,8	6,2
Peralillo		23	39,1	39,1	17,4	4,3	0	91,3	8,7	
Placilla		22	40,9	31,8	18,2	9,1	0	90,9	9,1	
Pumanque		23	39,1	43,5	13	4,3	0	91,3	8,7	
San Fernando	45	51,1	22,2	24,4	2,2	0	93,3	6,7		

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
Maule	Cauquenes	Santa Cruz	59	44,1	27,1	27,1	1,7	0	94,9	5,1
		Cauquenes	54	53,7	24,1	22,2	0	0	94,4	5,6
		Chanco	19	42,1	42,1	10,5	5,3	0	94,7	5,3
		Pelluhue	28	64,3	21,4	7,1	7,1	0	92,9	7,1
	Curicó	Curicó	71	54,9	26,8	15,5	1,4	1,4	94,4	5,6
		Hualañé	25	52	28	16	4	0	92	8
		Licantén	10	50	40	10	0	0	90	10
		Molina	33	54,5	18,2	21,2	0	6,1	93,9	6,1
		Rauco	24	50	29,2	20,8	0	0	91,7	8,3
		Romeral	49	57,1	30,6	6,1	6,1	0	93,9	6,1
		Sagrada Familia	40	32,5	40	25	2,5	0	95	5
		Teno	80	41,2	26,2	31,2	1,2	0	95	5
		Vichuquén	35	42,9	37,1	20	0	0	94,3	5,7
		Linares	Colbún	57	47,4	24,6	24,6	1,8	1,8	94,7
	Linares		116	55,2	32,8	10,3	1,7	0	94,8	5,2
	Longaví		89	53,9	21,3	23,6	1,1	0	94,4	5,6
	Parral		58	44,8	27,6	24,1	1,7	1,7	94,8	5,2
	Retiro		61	42,6	34,4	21,3	1,6	0	95,1	4,9
	San Javier		74	47,3	29,7	21,6	1,4	0	94,6	5,4
	Villa Alegre		36	58,3	16,7	25	0	0	94,4	5,6
	Yerbas Buenas		61	55,7	29,5	13,1	1,6	0	95,1	4,9
	Talca	Constitución	65	53,8	23,1	13,8	7,7	1,5	93,8	6,2
		Curepto	37	35,1	40,5	21,6	2,7	0	94,6	5,4
		Empedrado	12	75	8,3	16,7	0	0	91,7	8,3
		Maule	55	38,2	25,5	34,5	1,8	0	94,5	5,5
		Pelarco	25	36	40	20	4	0	92	8
		Pencahue	35	48,6	34,3	14,3	2,9	0	94,3	5,7
		Río Claro	43	48,8	23,3	27,9	0	0	93	7
San Clemente		119	46,2	32,8	19,3	0,8	0,8	95	5	
San Rafael		31	58,1	29	12,9	0	0	93,5	6,5	
Talca	51	39,2	37,3	21,6	2	0	94,1	5,9		

Al diferenciar los resultados de clusterización según el grado de consistencia (Figura 37) se obtienen proporciones similares entre las distintas regiones de la macrozona centro, obteniendo menos de 10% de inconsistencia en las distintas provincias, a excepción de la provincia de Talagante, que alcanza el mayor grado de inconsistencia (35%), Quillota y Cordillera (ambas alcanzan 11%). En estas tres provincias mencionadas, el alto número de viviendas, densidad y distancia en ellas son factores condicionantes en el resultado obtenido. En efecto, son polígonos que se distribuyen heterogéneamente respecto a su densidad local, siendo ambos casos interesantes de evaluar al ser agrupaciones potenciales de ser subdivididas posteriormente.

Figura 37: Proporción de tipos de clúster según consistencia en Macrozona Centro.

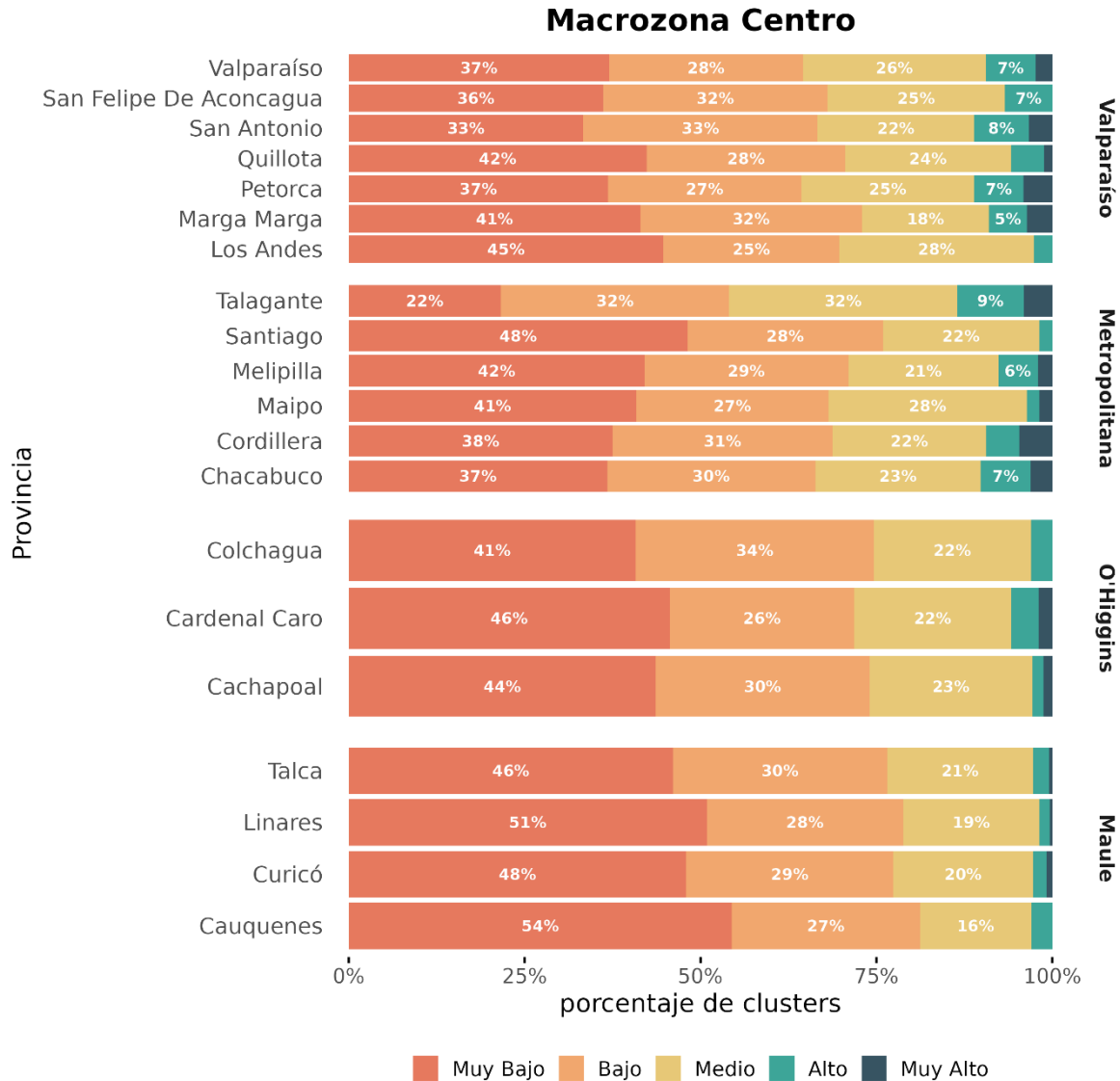


Las etiquetas de los segmentos menores al 5% fueron removidas para mejorar la visualización del gráfico

Finalmente, si se clasifican los clústeres según su tamaño de agrupamiento (Figura 38), se evidencia que la mayoría de las provincias presentan un tamaño “Muy bajo”, aunque en un rango variado, encontrando porcentajes desde 54% (Cauquenes) a 22% (Talagante). Esto puede deberse a la métrica utilizada para definir el modelo (*MinPts* de 20 viviendas), favoreciendo la búsqueda de agrupaciones con menor tamaño. En este sentido, se observan una alta proporción de clusters clasificados con tamaños “Bajo” y “Medio”, reflejando que en general, las viviendas al interior de los clusters abarcan las 50 a 300 viviendas. Asimismo, la clase de “Alto” tamaño son variadas, alcanzando los mayores valores en las provincias de Talagante (16%) y San Antonio (7,78%), indicando que todas estas agrupaciones donde se logran superar como mínimo 301 viviendas, reflejando su gran tamaño. De manera similar, en la clase “Muy alto” destaca el caso de

las provincias de Cordillera, San Antonio y Talagante (4,69% a 4,05%), reflejando que logran superar las 600 viviendas en un elevado porcentaje respecto al resto de provincias y regiones.

Figura 38: Clasificación de los clústeres según su tamaño de agrupamiento en la Macrozona Centro.



Las etiquetas de los segmentos menores al 5% fueron removidas para mejorar la visualización del gráfico

## 6.4. Macrozona Sur

En la Macrozona Sur (regiones Ñuble, Biobío, La Araucanía y Los Ríos) el poblamiento rural es disperso y fragmentado. La depresión intermedia facilita asentamientos en ciertos tramos, pero la densa red de ríos, lagos y humedales quiebra la continuidad territorial, propiciando núcleos pequeños y separados. En total se clusterizan 382.225 viviendas rurales, alcanzando un 62,55% respecto al total de viviendas de la macrozona.

Respecto a las viviendas clusterizadas en la macrozona sur:

- La Araucanía concentra el mayor número de viviendas clusterizadas (160.385); Los Ríos registra la menor cifra (49.292).
- La proporción de viviendas clusterizadas por región es similar: Biobío 63,95%, Ñuble 62,34%, Los Ríos 61,84% y Araucanía 62%.
- A nivel comunal, el mayor volumen de viviendas clusterizadas se observa en Los Ángeles (44.391) y Villarrica (32.980).

A escala provincial (Figura 39) aparecen diferencia intrarregional explicadas por orografía, accidentes geográficos y usos del suelo (agrícola/forestal) que alternan dispersión y concentración.

- Menor clusterización: Ranco (Los Ríos) con 60,3%, indicando mayor dispersión.
- Mayor clusterización: Arauco (Biobío) con 65,9%, asociada a “bolsas” sin vivienda y concentración en sectores viables (matriz agrícola-forestal y áreas montañosas/vegetadas que restringen el espacio habitable).

Destacan además las comunas: Padre Las Casas (89,5%), Curanilahue (83,4%) y Coronel (81,9%), las que poseen porcentajes mayores de índices de clusterización. Respecto a las viviendas ruido permanentes en la macrozona sur, estas presentan un comportamiento similar, contando con una proporción cercana a un 70% en todas las provincias (Figura 40).

En cambio, las viviendas ruido clasificadas como satélite presentan un comportamiento levemente diferente, desde la provincia Ranco en Los Ríos (23,3%), hasta la provincia de Itata de la región de Ñuble (29,7%). Estos resultados indican que en general la mayoría de las agrupaciones se encuentran rodeadas de viviendas ruido con una distancia y densidad superior, por lo que el algoritmo logra diferenciarlas de los clústeres obtenidos (70%).

Alrededor de un 30% sí se encuentran cercanas en distancia y densidad a las agrupaciones, y que podrían ser viviendas potenciales de ser incorporadas en las agrupaciones si la métrica *MinPts* cambiara en magnitud.

Figura 39: Proporción de viviendas clusterizadas por provincia en la Macrozona Sur

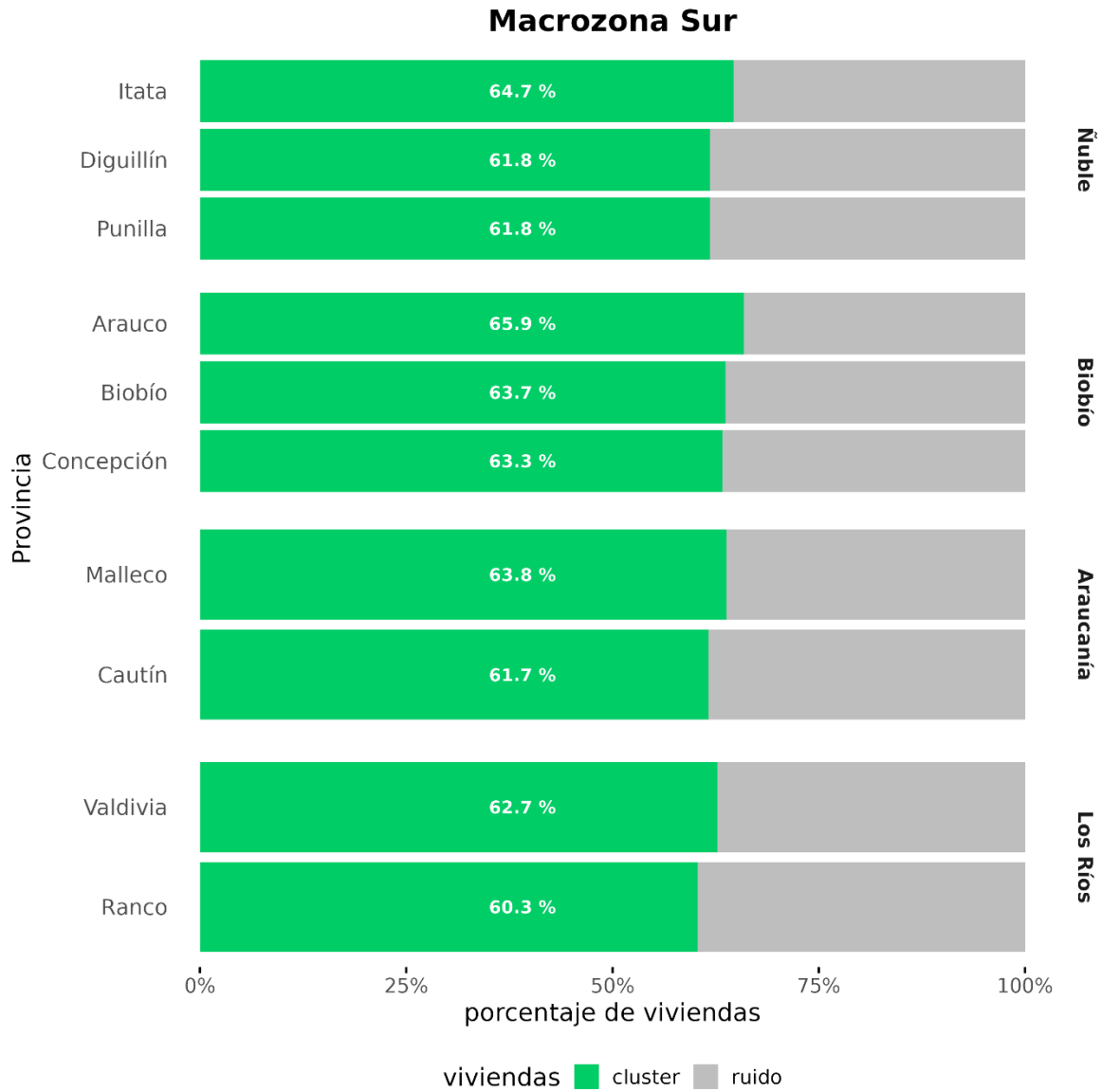
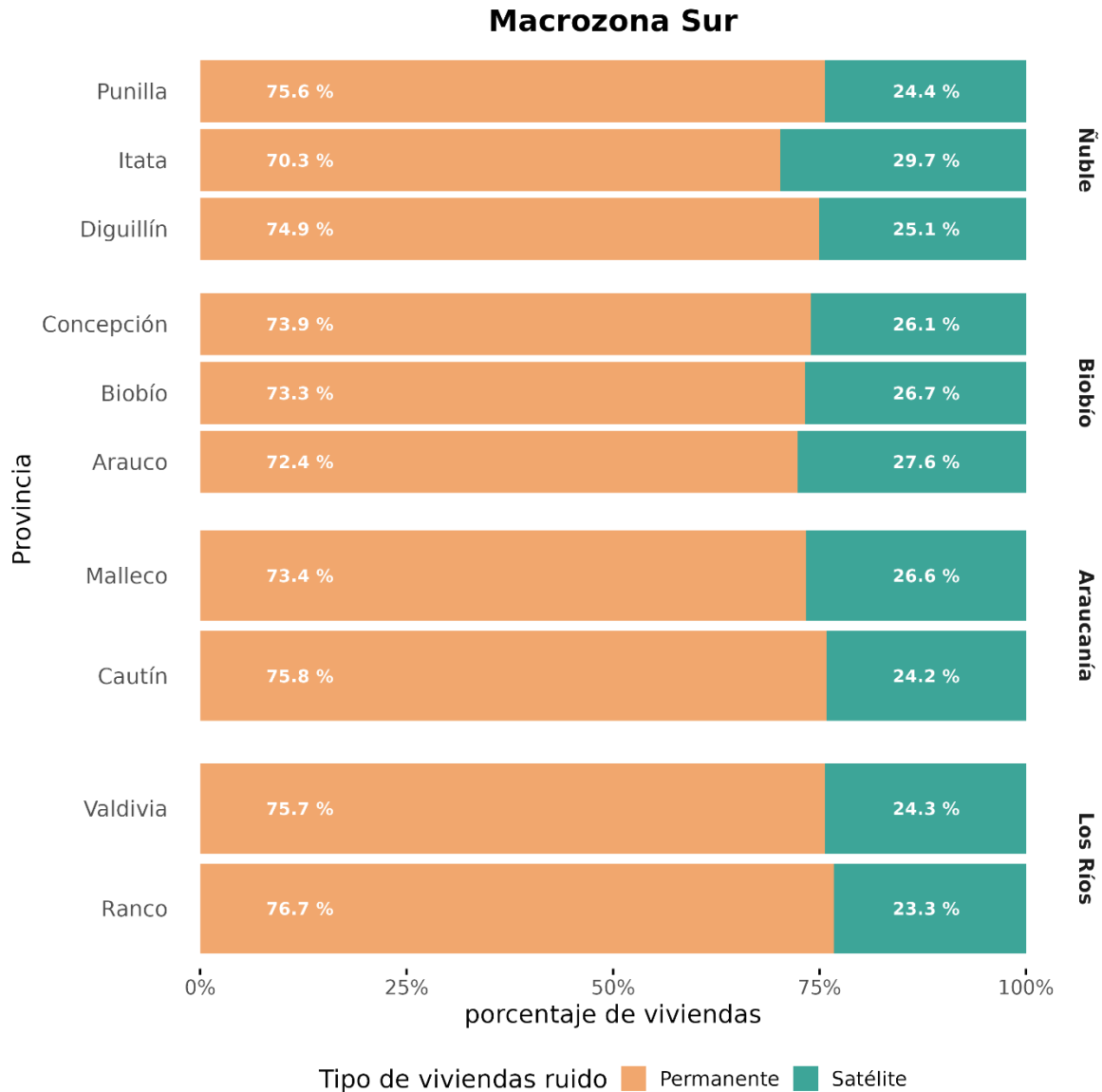


Figura 40: Proporción de viviendas ruido según su tipo en la Macrozona Sur.



De forma complementaria, la Figura 41 muestra a nivel provincial el rango de viviendas clusterizadas (mínimo, promedio y máximo):

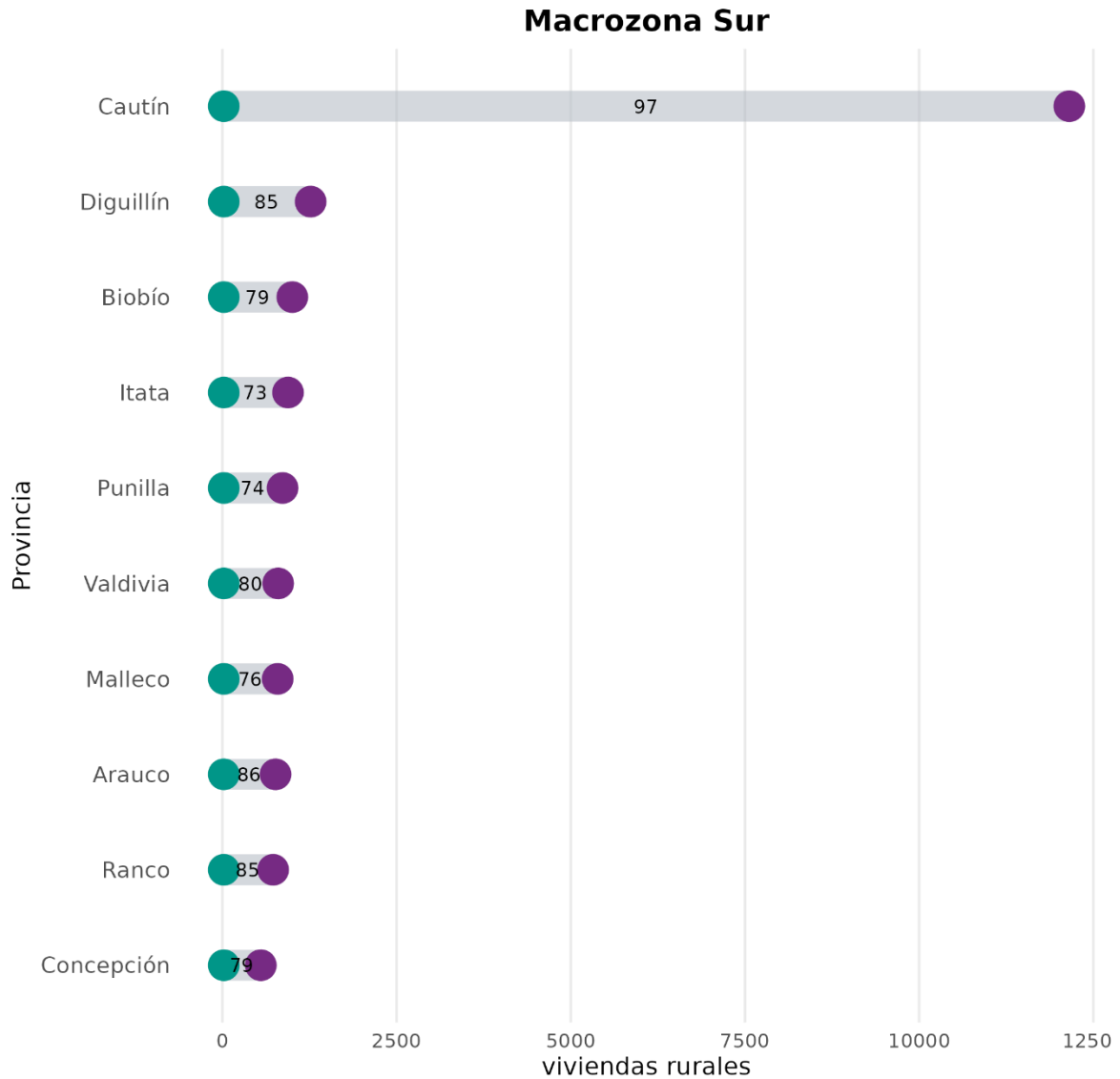
- El valor mínimo (representado en color verde) corresponde al umbral metodológico definido para *MinPts* de 20 viviendas, coincidiendo con este valor en todas las provincias.
- El valor máximo (representado en color morado) varía gradualmente entre provincias, desde alrededor de 552 viviendas clusterizadas en la provincia de Concepción hasta Diguillín con 1.267 viviendas, sugiriendo una configuración relativamente homogénea en toda la macrozona.

- Excepción de la provincia de Cautín: única que alcanza las 12.155 viviendas.

En general, destacan tres comunas de gran envergadura que sobrepasan las 10.000 viviendas rurales clusterizadas en la provincia de Cautín de la región de la Araucanía: la comuna de Villarrica (19.031), Padre Las Casas (16.552) y Pucón (20.438). En el caso de las comunas de Villarrica y Pucón, las viviendas clusterizadas se sitúan en cercanías del núcleo urbano, siendo sectores con una predominancia del sector turístico y, probablemente, los clusters puedan asociarse a viviendas vacacionales.

Un hallazgo interesante de destacar es que la macrozona Sur tiene una estructura de poblamiento rural más consolidada y dispersa que en el norte, por ejemplo, y eso da pie a detectar clústeres grandes (en contexto de macrozona sur), tanto en planicie como en valles cortos costeros, asociados a vialidad de jerarquía importante, y en torno a lagos relevantes donde la actividad turística es importante (lago Villarrica, lago Calafquen o lago Ranco). También en valles internos cordilleranos, en menor medida (valle río Allipen o valle del río Cautín).

Figura 41: Rango de viviendas (mínimo y máximo) de clusters por provincia en la Macrozona Sur



El valor dentro de cada barra corresponde al promedio de viviendas por cluster en cada provincia

Con el fin de comprender los resultados de la clusterización a distintas escalas, la Tabla 16 presenta un resumen de los índices previamente mencionados, incluyendo la subclasificación de las viviendas clasificadas como ruido y el índice de clusterización comunal. En la tabla se observa que es la comuna de Los Ángeles (Biobío) la que presenta un mayor número total de viviendas (44.391), clusterizadas (27.920) y de ruido (16.471). Por su parte, al analizar la subclasificación de las viviendas ruido, destaca la comuna de Chiguayante, que no presenta clusterización, por lo que las viviendas rurales siempre son ruido (100%). Destacando las comunas Panguipulli, Paillaco, Temuco, Río Bueno y Lanco, que son las que presentan mayores viviendas con ruido “permanente” (81,3-81,6%) y en Padre Las Casas (94,7%) clasificadas como “satélite”. Por último, esta última comuna también es la que presenta el mayor índice de clusterización (89,5%).

Tabla 16: Indicadores de clusterización a nivel de comuna, macrozona sur.

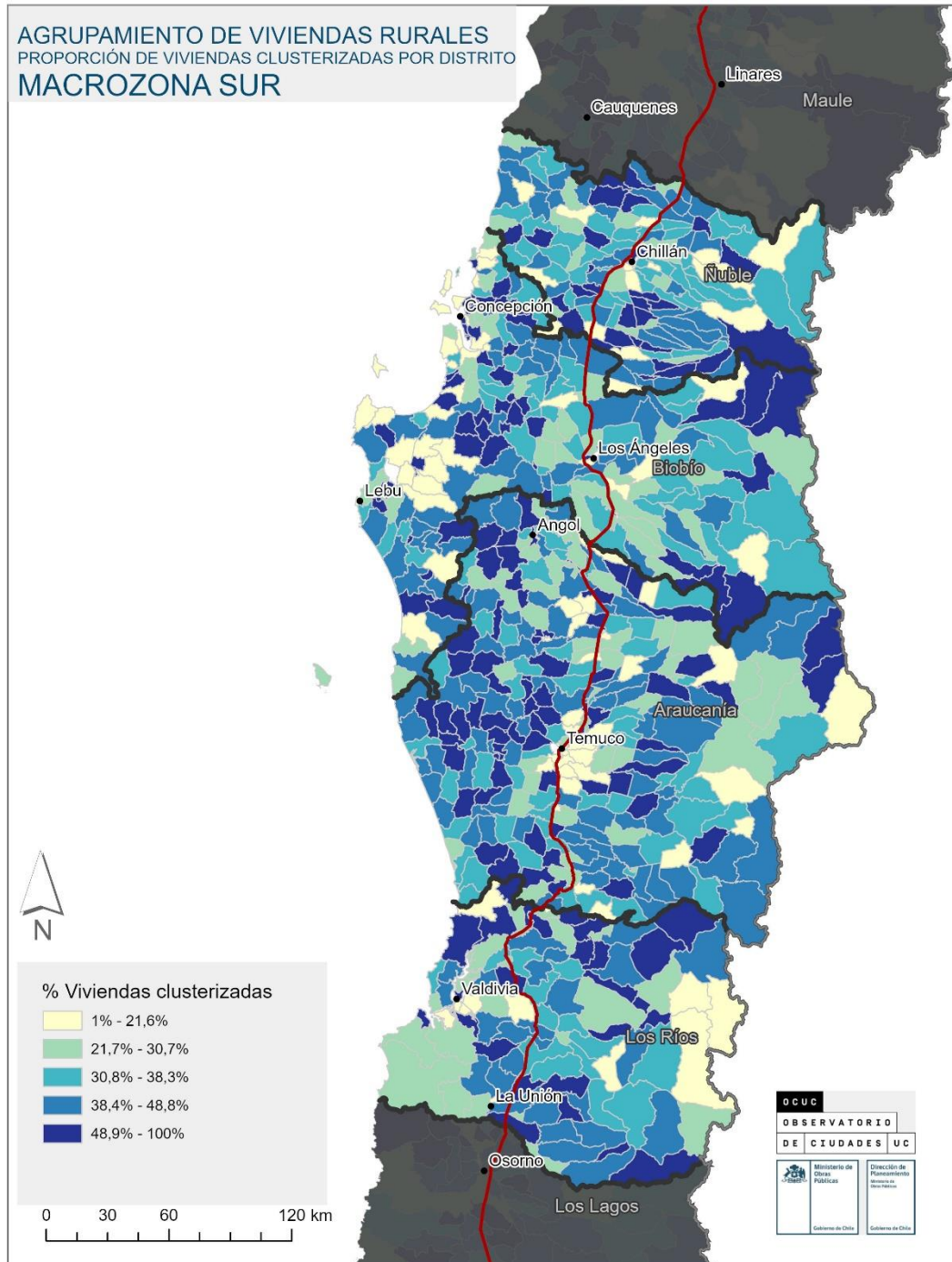
Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Total ruido	Viviendas Ruido permanente	Viviendas Ruido Satélites	Índice de clusterización
Ñuble	Diguillín	Bulnes	6.195	3.660	2.535	76	24	59
		Chillán	15.407	9.793	5.614	73	27	64
		Chillán Viejo	3.147	2.257	890	67	33	72
		El Carmen	5.489	3.224	2.265	80	20	59
		Pemuco	2.497	1.651	846	66	34	66
		Pinto	5.738	4.225	1.513	65	35	74
		Quillón	7.738	4.506	3.232	79	21	58
		San Ignacio	6.131	3.137	2.994	79	22	51
	Yungay	4.926	2.938	1.988	75	25	60	
	Itata	Cobquecura	3.516	2.255	1.261	68	32	64
		Coelemu	4.113	2.751	1.362	67	33	67
		Ninhue	2.642	1.928	714	67	34	73
		Portezuelo	2.555	1.741	814	72	28	68
		Quirihue	2.928	1.695	1.233	73	27	58
		Ránquil	3.184	2.061	1.123	76	24	65
		Treguaco	2.795	1.626	1.169	69	31	58
	Punilla	Coihueco	10.393	6.910	3.483	72	28	67
		San Carlos	10.650	6.286	4.364	76	24	59
		San Fabián	2.431	1.682	749	65	35	69
San Nicolás		5.288	3.239	2.049	77	23	61	
Ñiquén		8.061	4.638	3.423	80	20	58	
Biobío	Arauco	Arauco	5.244	3.653	1.591	65	35	70
		Cañete	8.291	4.797	3.494	80	20	58
		Contulmo	2.016	1.217	799	80	20	60
		Curanilahue	1.833	1.528	305	38	62	83
		Lebu	1.970	1.258	712	68	32	64
		Los Álamos	1.662	1.140	522	68	32	69
		Tirúa	3.306	2.432	874	68	32	74
	Biobío	Alto Biobío	7.209	4.795	2.414	72	29	67
		Antuco	1.706	1.081	625	70	30	63
		Cabrero	5.271	3.432	1.839	75	25	65
		Laja	4.956	3.112	1.844	74	26	63
		Los Ángeles	44.391	27.920	16.471	74	26	63
		Mulchén	5.125	3.709	1.416	59	41	72
		Nacimiento	5.316	3.109	2.207	74	27	59
		Negrete	2.531	1.457	1.074	75	25	58
		Quilaco	2.383	1.602	781	73	27	67
		Quilleco	3.300	2.348	952	69	31	71
		San Rosendo	959	564	395	75	25	59
		Santa Bárbara	6.494	4.380	2.114	71	29	67
	Tucapel	3.377	2.036	1.341	69	31	60	
Yumbel	14.505	8.913	5.592	76	24	61		
Concepción	Chiguayante	53	0	53	100	0	0	
	Concepción	2.925	2.044	881	71	29	70	
	Coronel	2.721	2.228	493	47	53	82	
	Florida	5.372	3.075	2.297	79	21	57	
	Hualpén	417	330	87	54	46	79	
	Hualqui	3.566	2.266	1.300	75	25	64	
	Lota	175	129	46	67	33	74	
	Penco	520	255	265	80	20	49	
San Pedro De La Paz	106	84	22	50	50	79		

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Total ruido	Viviendas Ruido permanente	Viviendas Ruido Satélites	Índice de clusterización
		Santa Juana	4.259	2.054	2.205	81	19	48
		Talcahuano	370	280	90	58	42	76
		Tomé	4.574	3.117	1.457	67	33	68
Araucanía	Cautín	Carahue	9.362	4.857	4.505	78	22	52
		Cholchol	4.399	2.313	2.086	80	20	53
		Cunco	8.107	4.689	3.418	79	21	58
		Curarrehue	5.820	3.418	2.402	78	22	59
		Freire	12.899	8.111	4.788	76	24	63
		Galvarino	6.179	3.531	2.648	77	23	57
		Gorbea	3.386	2.027	1.359	77	23	60
		Lautaro	10.310	5.998	4.312	75	25	58
		Loncoche	9.137	5.627	3.510	77	23	62
		Melipeuco	3.898	3.028	870	54	46	78
		Nueva Imperial	13.957	7.912	6.045	78	22	57
		Padre Las Casas	18.497	16.552	1.945	5	95	90
		Perquenco	3.032	1.808	1.224	77	23	60
		Pitrufquén	8.774	5.403	3.371	78	23	62
		Pucón	20.438	11.762	8.676	80	20	58
		Saavedra	9.654	5.675	3.979	77	23	59
		Temuco	12.407	6.973	5.434	81	19	56
		Teodoro Schmidt	9.929	5.813	4.116	79	21	59
		Toltén	4.309	2.617	1.692	77	24	61
		Vilcún	9.533	6.650	2.883	69	31	70
Villarrica	32.980	19.031	13.949	78	22	58		
Malleco	Angol	6.422	3.720	2.702	79	21	58	
	Collipulli	4.404	3.056	1.348	65	35	69	
	Curacautín	5.139	3.218	1.921	75	25	63	
	Ercilla	2.801	1.839	962	69	31	66	
	Lonquimay	5.206	3.606	1.600	71	29	69	
	Los Sauces	2.500	1.387	1.113	75	25	56	
	Lumaco	3.317	1.833	1.484	78	22	55	
	Purén	2.706	1.675	1.031	75	25	62	
	Renaico	1.430	953	477	72	28	67	
	Traiguén	2.425	1.435	990	75	25	59	
Victoria	5.308	3.868	1.440	67	33	73		
Los Ríos	Ranco	Futrono	5.055	3.698	1.357	66	34	73
		La Unión	6.984	3.994	2.990	77	23	57
		Lago Ranco	5.227	3.285	1.942	72	28	63
		Río Bueno	11.370	6.296	5.074	81	19	55
	Valdivia	Corral	2.000	1.507	493	60	40	75
		Lanco	3.650	2.028	1.622	81	19	56
		Los Lagos	7.160	4.710	2.450	72	29	66
		Mariquina	6.862	4.669	2.193	71	29	68
		Máfil	2.272	1.396	876	77	23	61
		Paillaco	4.379	2.563	1.816	81	19	59
Panguipulli	12.725	6.823	5.902	82	18	54		
Valdivia	12.024	8.323	3.701	68	32	69		

La información anteriormente mencionada, queda plasmada en la Cartografía 5, donde se ilustra, por distrito censal de la macrozona sur, el porcentaje de viviendas organizada

en clústeres. Cada distrito se codifica según su nivel de agrupamiento, y en gris se marcan aquellos sectores en los que el modelo no reconoció ningún clúster.

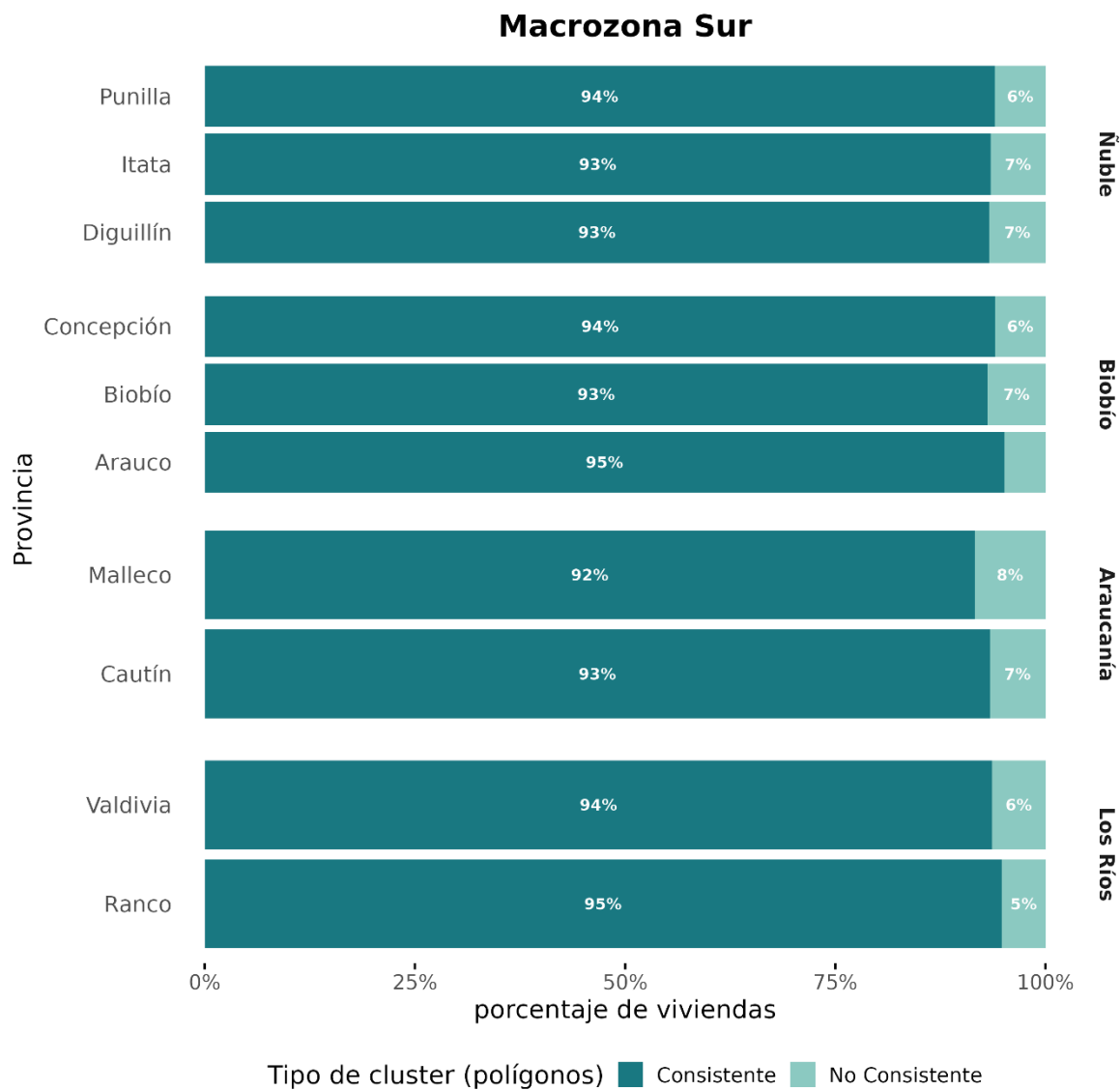
Cartografía 5. Agrupamiento de viviendas rurales por comuna en la Macrozona Sur



Al analizar los clusters según su grado de consistencia (Figura 42), se logra encontrar que la mayoría de las provincias presentan una alta consistencia, desde la provincia Malleco (92%), hasta Arauco y Ranco (95%). Esto indica que la provincia de Malleco es la que presenta el mayor grado de inconsistencia (8%), el cual sigue siendo menor en nivel, pero considerable respecto al resto de provincias de esta macrozona.

En general, los resultados indican que los polígonos obtenidos del modelo son en su mayoría agrupaciones donde las viviendas se distribuyen homogéneamente respecto a su densidad local, resultando en un buen indicador. En cambio, el caso de la provincia de Malleco refleja que la densidad se comporta de forma heterogénea en un 8%, siendo agrupaciones potenciales de ser subdivididas si los parámetros del modelo cambiaran.

Figura 42: Proporción de tipos de clúster según consistencia en Macrozona Sur.

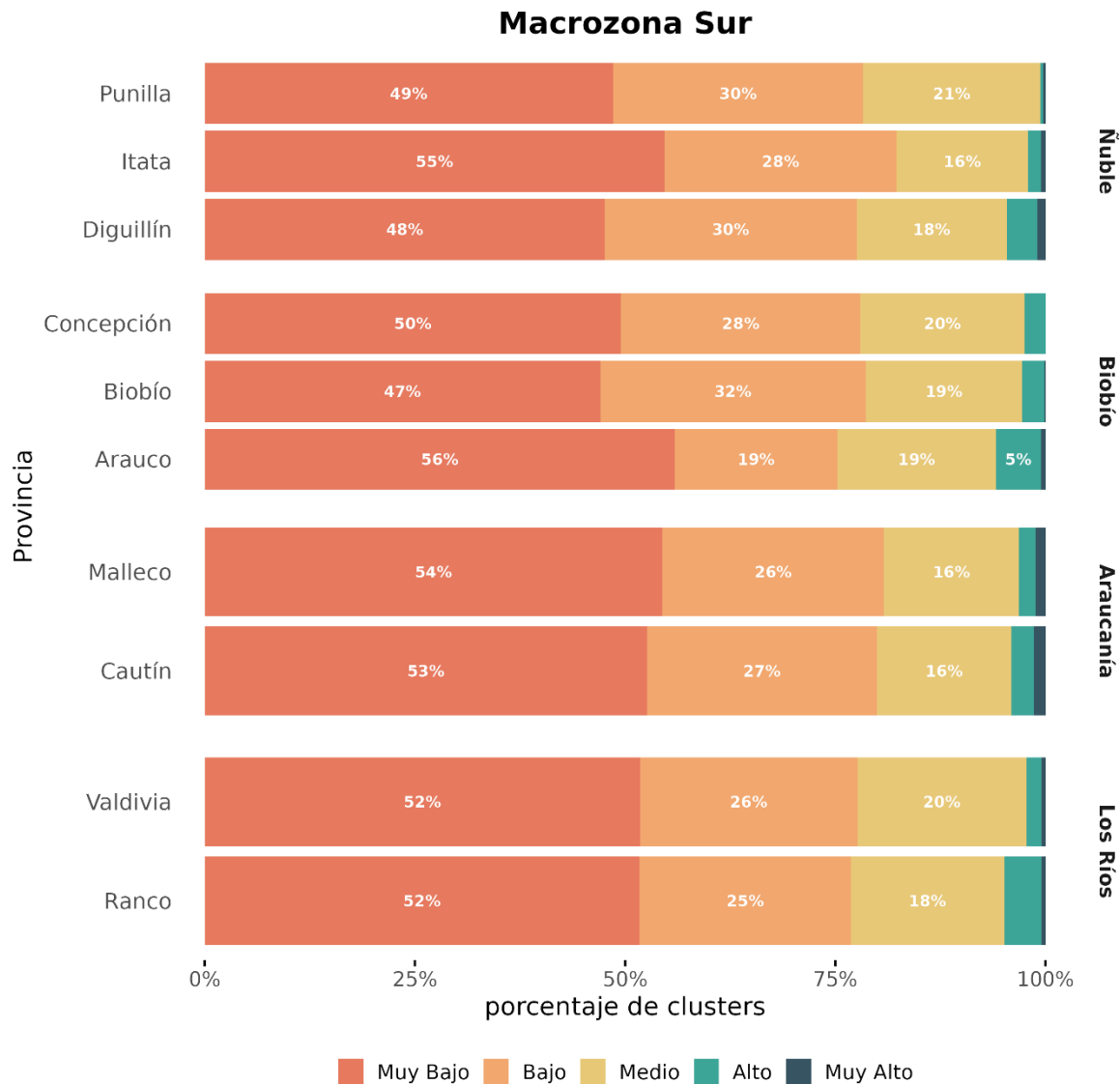


Las etiquetas de los segmentos menores al 5% fueron removidas para mejorar la visualización del gráfico

Por último, si se clasifican los clústeres según su tamaño de agrupamiento (Figura 43) se logra evidenciar que en la mayoría de las provincias presentan un tamaño “Muy bajo”, desde un 47% en Biobío hasta un 56% en Arauco. Esto puede deberse al parámetro *MinPts* dado en el modelo (20 viviendas), por lo que podría favorecer la búsqueda de agrupaciones con un menor tamaño.

En efecto, en general la macrozona sur presenta en su mayoría agrupaciones clasificadas con tamaños “Muy bajo”, “Bajo” y “Medio”. En cambio, las clases “Alto” y “Muy alto” solo se dan en la provincia de Arauco de la región del Biobío en un mayor porcentaje (5%) respecto al resto de las provincias, pero continuando bajo respecto al resto de tamaños. Esto puede deberse al ser la provincia con un mayor número de viviendas clusterizadas, 72,4%.

Figura 43: Clasificación de los clústeres según su tamaño de agrupamiento en la Macrozona Sur.



Las etiquetas de los segmentos menores al 5% fueron removidas para mejorar la visualización del gráfico

Por último, la Tabla 17 muestra el detalle de los índices de clasificación de clústeres según su tamaño de agrupamiento y nivel de consistencia. En este sentido, la comuna Los Ángeles es la que presenta la mayor cantidad de clusters (283). Además, las comunas que presentan mayor porcentaje en cada subclasificación por tamaño son: San Pedro de la Paz con “Muy bajo” (100%), Talcahuano con “Bajo” (60%), Concepción con “Medio” (38,9%), Melipeuco con “Alto” (18,8%) y Padre Las Casas con “Muy alto” (50%). Finalmente, Negrete y Galvarino son las comunas con los clusters con mayor consistencia (95,2 y 95,1% consistencia, respectivamente), en cambio Lota junto con San Pedro de La Paz son las que menos consistencia presentan (33,3% inconsistencia).

Tabla 17: Indicadores de clasificación a nivel de clusters, macrozona sur.

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
Ñuble	Diguillín	Bulnes	53	49,1	28,3	20,8	1,9	0	94,3	5,7
		Chillán	94	39,4	30,9	21,3	7,4	1,1	94,7	5,3
		Chillán Viejo	22	54,5	18,2	18,2	4,5	4,5	90,9	9,1
		El Carmen	53	56,6	32,1	9,4	1,9	0	94,3	5,7
		Pemuco	18	44,4	33,3	16,7	5,6	0	94,4	5,6
		Pinto	30	56,7	10	23,3	3,3	6,7	93,3	6,7
		Quillón	56	42,9	35,7	19,6	1,8	0	94,6	5,4
		San Ignacio	51	52,9	37,3	7,8	2	0	94,1	5,9
	Yungay	37	43,2	29,7	24,3	2,7	0	94,6	5,4	
	Itata	Cobquecura	32	68,8	9,4	18,8	3,1	0	93,8	6,2
		Coelemu	36	55,6	25	16,7	2,8	0	94,4	5,6
		Ninhue	15	53,3	26,7	13,3	0	6,7	93,3	6,7
		Portezuelo	24	33,3	45,8	20,8	0	0	91,7	8,3
		Quirihue	36	66,7	27,8	5,6	0	0	94,4	5,6
		Ránquil	29	51,7	24,1	24,1	0	0	93,1	6,9
	Punilla	Treguaco	20	40	45	10	5	0	95	5
		Coihueco	74	37,8	36,5	23	1,4	1,4	94,6	5,4
		San Carlos	88	50	25	25	0	0	94,3	5,7
		San Fabián	23	47,8	26,1	26,1	0	0	91,3	8,7
		San Nicolás	55	60	27,3	12,7	0	0	94,5	5,5
Ñiquén	69	49,3	31,9	18,8	0	0	94,2	5,8		
Biobío	Arauco	Arauco	40	52,5	22,5	15	10	0	95	5
		Cañete	60	56,7	16,7	23,3	3,3	0	95	5
		Contulmo	19	63,2	15,8	21,1	0	0	94,7	5,3
		Curanilahue	12	41,7	16,7	25	16,7	0	91,7	8,3
		Lebu	24	66,7	20,8	12,5	0	0	91,7	8,3
		Los Álamos	16	62,5	18,8	18,8	0	0	93,8	6,2
		Tirúa	15	40	26,7	13,3	13,3	6,7	93,3	6,7
	Biobío	Alto Biobío	72	48,6	34,7	15,3	1,4	0	94,4	5,6
		Antuco	16	43,8	37,5	18,8	0	0	93,8	6,2
		Cabrero	48	43,8	41,7	12,5	2,1	0	93,8	6,2
		Laja	45	53,3	26,7	20	0	0	93,3	6,7
		Los Ángeles	283	41,3	29	24,4	4,9	0,4	94,7	5,3
		Mulchén	50	50	30	16	4	0	94	6
		Nacimiento	45	51,1	26,7	22,2	0	0	93,3	6,7
		Negrete	21	42,9	28,6	28,6	0	0	95,2	4,8
		Quilaco	25	64	20	12	4	0	92	8
		Quilleco	26	38,5	34,6	26,9	0	0	92,3	7,7
		San Rosendo	10	50	40	10	0	0	90	10
		Santa Bárbara	63	50,8	30,2	19	0	0	93,7	6,3

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
Concepción		Tucapel	31	45,2	45,2	9,7	0	0	93,5	6,5
		Yumbel	130	53,1	33,8	10	3,1	0	94,6	5,4
	Concepción	Concepción	18	38,9	22,2	38,9	0	0	94,4	5,6
		Coronel	18	38,9	22,2	33,3	5,6	0	94,4	5,6
		Florida	43	48,8	30,2	20,9	0	0	93	7
		Hualpén	5	40	40	20	0	0	80	20
		Hualqui	26	53,8	19,2	19,2	7,7	0	92,3	7,7
		Lota	3	66,7	33,3	0	0	0	66,7	33,3
		Penco	6	66,7	33,3	0	0	0	83,3	16,7
		San Pedro De La Paz	3	100	0	0	0	0	66,7	33,3
		Santa Juana	33	60,6	24,2	15,2	0	0	93,9	6,1
		Talcahuano	5	40	60	0	0	0	80	20
		Tomé	40	42,5	37,5	15	5	0	95	5
Araucanía	Cautín	Carahue	71	59,2	19,7	19,7	1,4	0	94,4	5,6
		Cholchol	35	54,3	31,4	14,3	0	0	94,3	5,7
		Cunco	60	48,3	26,7	23,3	1,7	0	95	5
		Curarrehue	57	71,9	12,3	15,8	0	0	94,7	5,3
		Freire	62	41,9	29	21	3,2	4,8	93,5	6,5
		Galvarino	61	60,7	29,5	9,8	0	0	95,1	4,9
		Gorbea	26	57,7	26,9	11,5	3,8	0	92,3	7,7
		Lautaro	89	58,4	27	13,5	1,1	0	94,4	5,6
		Loncoche	57	56,1	26,3	10,5	5,3	1,8	94,7	5,3
		Melipeuco	16	37,5	25	12,5	18,8	6,2	93,8	6,2
		Nueva Imperial	101	49,5	35,6	11,9	1	2	95	5
		Padre Las Casas	4	25	25	0	0	50	75	25
		Perquenco	25	40	36	24	0	0	92	8
		Pitrufquén	59	39	37,3	18,6	3,4	1,7	94,9	5,1
		Pucón	144	51,4	28,5	16	3,5	0,7	94,4	5,6
		Saavedra	47	42,6	31,9	19,1	0	6,4	93,6	6,4
		Temuco	88	60,2	23,9	12,5	2,3	1,1	94,3	5,7
		Teodoro Schmidt	62	46,8	25,8	19,4	8,1	0	93,5	6,5
		Toltén	32	53,1	31,2	6,2	9,4	0	93,8	6,2
		Vilcún	43	46,5	20,9	23,3	2,3	7	93	7
	Villarrica	246	54,1	26	17,1	2,4	0,4	94,7	5,3	
	Malleco	Angol	64	62,5	21,9	14,1	1,6	0	93,8	6,2
		Collipulli	39	61,5	15,4	20,5	0	2,6	94,9	5,1
		Curacautín	46	56,5	28,3	13	2,2	0	93,5	6,5
		Ercilla	11	9,1	45,5	27,3	9,1	9,1	90,9	9,1
		Lonquimay	32	37,5	34,4	18,8	6,2	3,1	93,8	6,2
		Los Sauces	27	63	29,6	7,4	0	0	92,6	7,4
		Lumaco	27	55,6	22,2	22,2	0	0	92,6	7,4
		Purén	23	56,5	26,1	13	4,3	0	91,3	8,7
		Renaico	13	38,5	38,5	23,1	0	0	92,3	7,7
Traiguén	24	45,8	50	4,2	0	0	91,7	8,3		
Victoria	43	60,5	14	20,9	2,3	2,3	93	7		
Los Ríos	Ranco	Futrono	35	51,4	20	20	5,7	2,9	94,3	5,7
		La Unión	50	46	28	22	4	0	94	6
		Lago Ranco	34	44,1	32,4	14,7	8,8	0	94,1	5,9
		Río Bueno	84	58,3	22,6	16,7	2,4	0	94	6
	Valdivia	Corral	18	38,9	38,9	22,2	0	0	94,4	5,6
		Lanco	32	43,8	46,9	9,4	0	0	93,8	6,2
		Los Lagos	69	58	23,2	17,4	1,4	0	94,2	5,8
		Mariquina	48	52,1	18,8	27,1	0	2,1	93,8	6,2

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
		Máfil	23	56,5	26,1	17,4	0	0	91,3	8,7
		Paillaco	32	40,6	37,5	18,8	3,1	0	93,8	6,2
		Panguipulli	92	62	20,7	15,2	2,2	0	94,6	5,4
		Valdivia	84	44	22,6	28,6	3,6	1,2	94	6

## 6.5. Macrozona Austral

La macrozona austral abarca el extremo sur de Chile e incluye las regiones de Los Lagos, Aysén y Magallanes. Estas dos últimas destacan por ser de las más extensas y con menor densidad de población del país. A pesar de ello, en las últimas décadas se ha observado un proceso sostenido de urbanización, concentrando parte de la población en pequeños núcleos urbanos (Cancino et al., 2021). En conjunto, esta macrozona representa el 13,9% de las viviendas rurales a nivel nacional.

El patrón de poblamiento está fuertemente condicionado por las barreras geográficas y la orografía, lo que genera un territorio fragmentado. La excepción se encuentra en las provincias de Llanquihue y Osorno, donde existe una mayor continuidad. En general, las viviendas rurales se localizan en bordes de lago, valles o sectores planos más aptos para el asentamiento, como ocurre en el valle Simpson.

En total, en la macrozona austral se agruparon 148.456 viviendas rurales en 1.658 clusters de viviendas rurales. La Región de Los Lagos concentra la mayor parte de estas agrupaciones (1.320 clusters), mientras que Aysén (196) y Magallanes (142) presentan un número considerablemente menor. En términos relativos, el 64,42% de las viviendas rurales de la macrozona se encuentra agrupado en clusters:

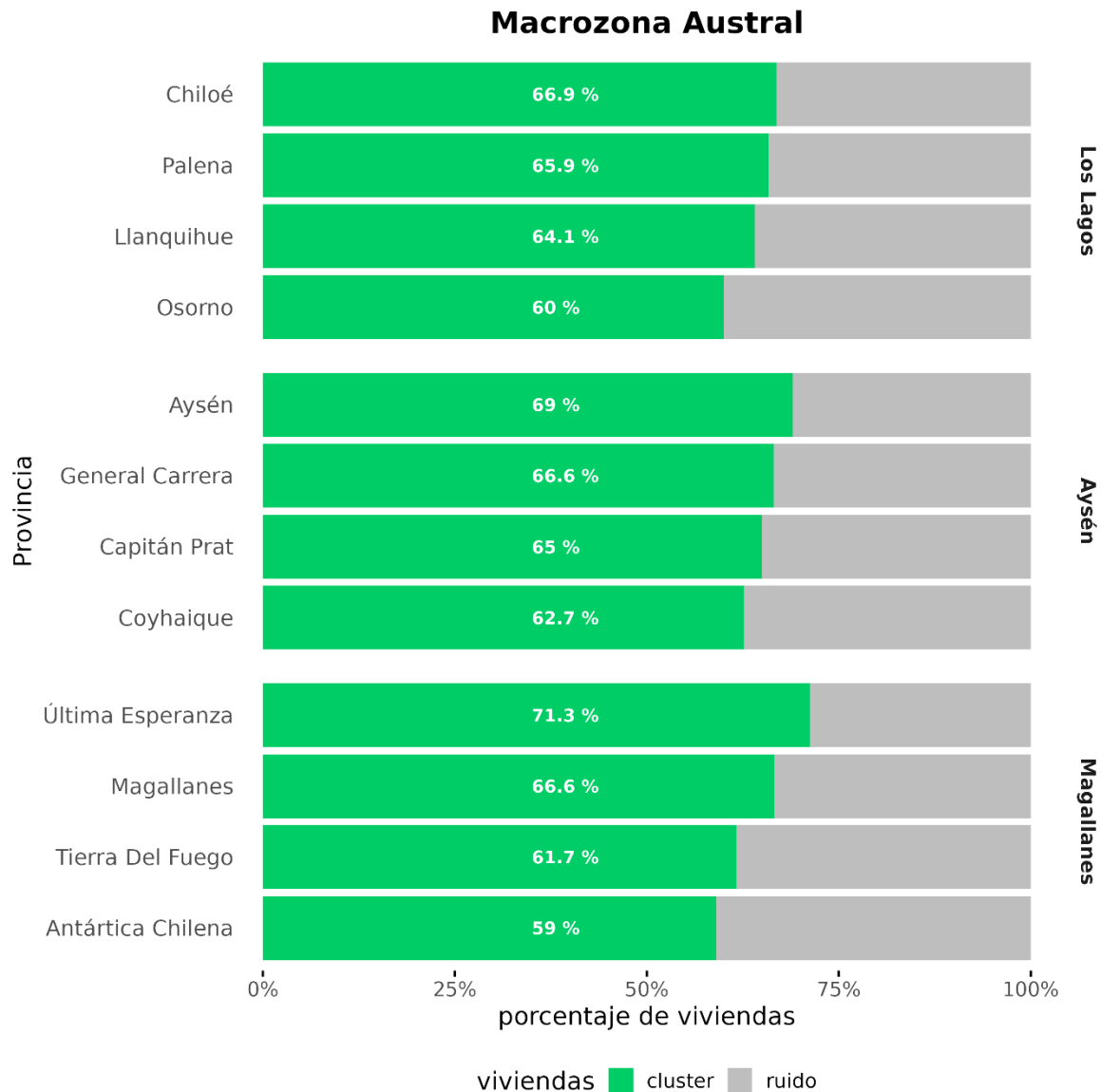
- Los Lagos: 64,14%
- Aysén: 65,09%
- Magallanes: 66,39%

A nivel nacional, la Región de Los Lagos, destaca por dos aspectos:

- Es una de las que presenta mayor proporción de viviendas clasificadas como ruido (67.098 viviendas ruido, después de la Araucanía con 98.280).
- Se ubica en el sexto lugar nacional en número de viviendas clusterizadas (120.023 viviendas clusterizadas, después de Araucanía con 160.385).

Esto refleja que, si bien existen numerosas viviendas agrupadas bajo los criterios del modelo, una parte significativa de las viviendas rurales de Los Lagos no cumplió los criterios mínimos para conformar un clúster, resultando en “ruido”, reflejando una alta heterogeneidad en su poblamiento.

Figura 44. Proporción de viviendas clusterizadas por provincia en la Macrozona Austral



A escala provincial (Figura 44), las mayores diferencias se observan entre territorios. Las provincias con menor proporción de viviendas clusterizadas son Antártica Chilena (59%), Osorno (60%) y Tierra Del Fuego (61,7%). En contraste, las proporciones más altas corresponden a Última Esperanza (71,3%), Aysén (69%) y Chiloé (66,9%).

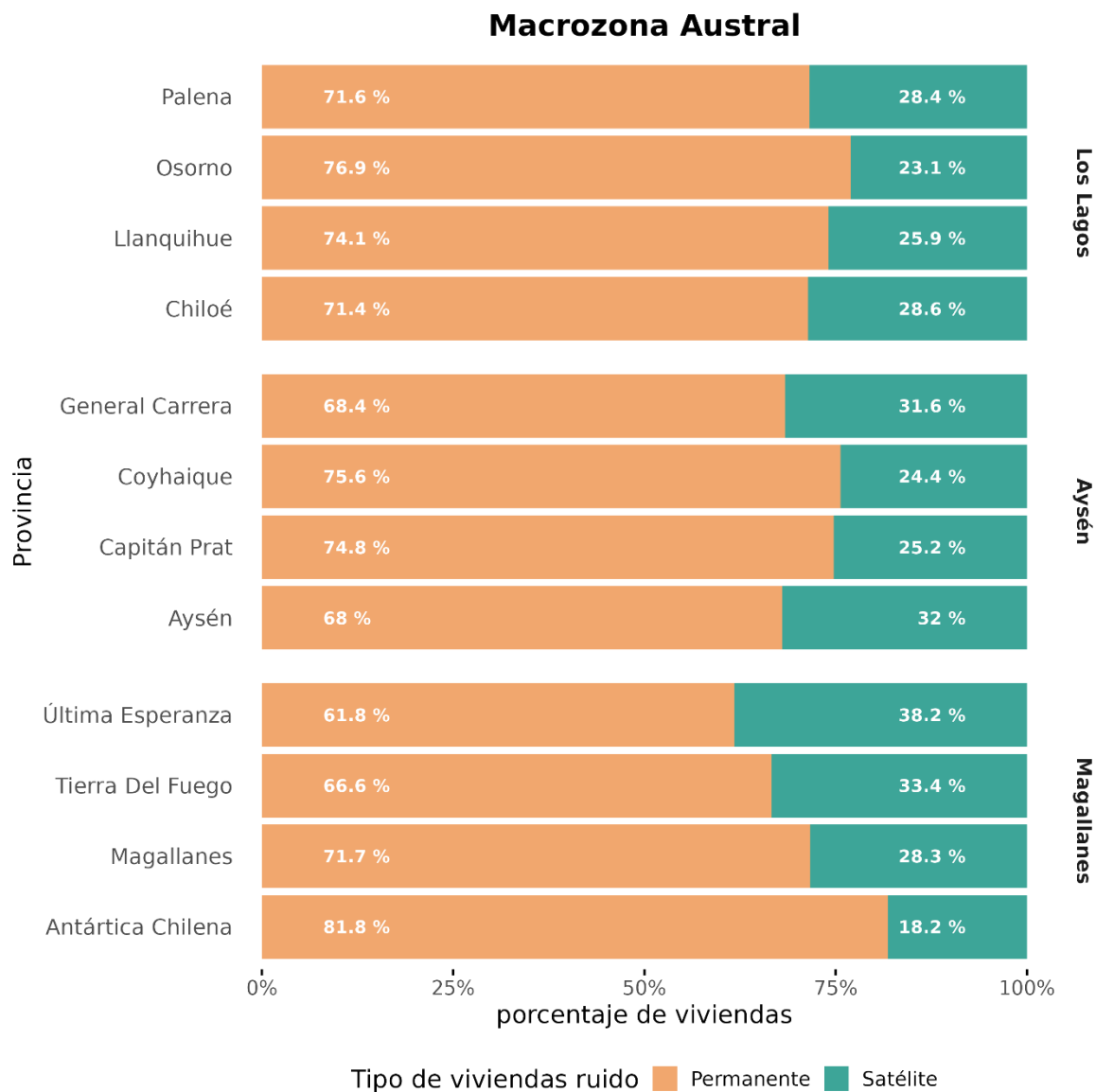
Estos resultados reflejan la diversidad geográfica y territorial del sur austral: mientras algunas zonas logran concentrar viviendas en núcleos relativamente compactos, otras se ven limitadas por las grandes distancias y condiciones orográficas, lo que reduce la posibilidad de formar clusters estables.

En la macrozona austral, las viviendas clasificadas como ruido permanente presentan proporciones variables entre 61,8% y 81,8% en todas las provincias (Figura 45). Estas

diferencias son más marcadas en la Región de Magallanes, donde se observan contrastes importantes entre provincias, a diferencia de Los Lagos y Aysén, que muestran un comportamiento más homogéneo.

Por su parte, las viviendas ruido tipo satélite aparecen en menor proporción, aunque siguen la misma tendencia de variabilidad territorial. Nuevamente, la región de Magallanes es la que concentra los mayores contrastes, con rangos que van desde 18,2% (Antártica Chilena) hasta 38,2% (Última Esperanza). Demostrando que gran parte de las agrupaciones se encuentran rodeadas de viviendas ruido que, por su distancia o baja densidad, son claramente diferenciadas por el algoritmo.

Figura 45: Proporción de viviendas ruido según su tipo en la Macrozona Austral.

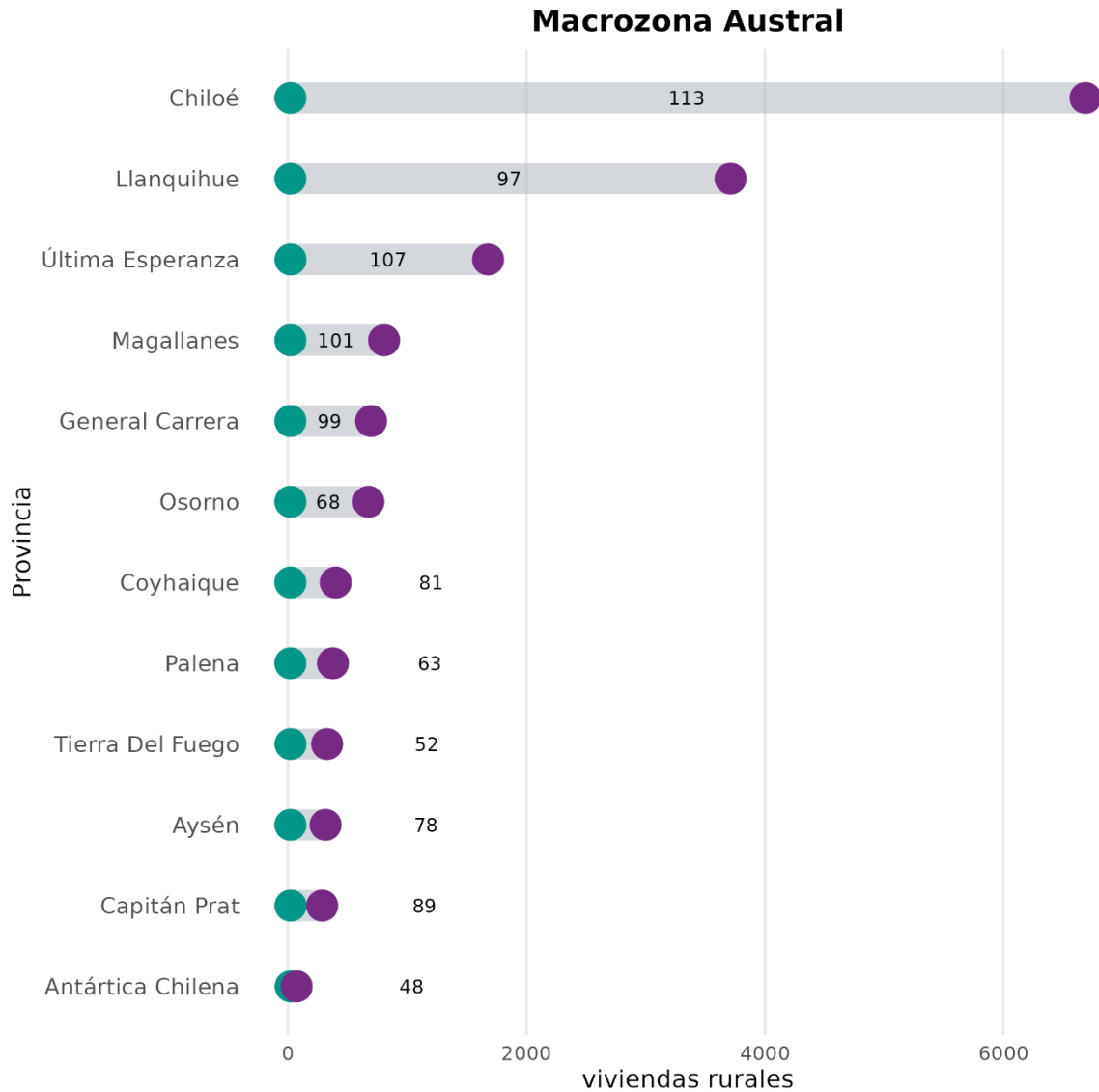


De forma complementaria, la Figura 46 presenta, a escala provincial, el rango de viviendas clusterizadas (mínimo, promedio y máximo). El valor mínimo (en verde) coincide en todas las provincias con el umbral metodológico definido por MinPts = 20 viviendas. En contraste, el máximo (en morado) muestra una variación gradual entre provincias: desde 22 viviendas en Antártica Chilena (mínimo) hasta alrededor de 6.687 viviendas en Chiloé (máximo).

A nivel comunal, destacan tres casos que superan las 1.500 viviendas clusterizadas:

- Chiloé (Los Lagos): Clústeres dispersos en el territorio apto, usualmente en pequeños núcleos rurales asociados a usos agropecuarios.
- Última Esperanza (Magallanes): Clústeres concentrados casi exclusivamente en el interior sur de la provincia (sectores de Puerto Natales, lago Balmaceda y lago Sarmiento), con presencia en áreas próximas al Parque Nacional Torres del Paine. En este caso, la actividad turística (deportes de montaña y expediciones) parece contribuir a una mayor densidad de viviendas en torno a estos polos.
- Llanquihue (Los Lagos): Presenta un máximo de viviendas clusterizadas de 3.711.

Figura 46: Rango de viviendas (mínimo y máximo) de clusters por provincia en la Macrozona Austral



El valor dentro de cada barra corresponde al promedio de viviendas por cluster en cada provincia

En términos regionales, las provincias vinculadas a núcleos urbanos principales tienen a exhibir mayor volumen de vivienda rural y, por lo tanto, más clústeres: particularmente Llanquihue, Osorno y Chiloé. Este patrón es consistente con lo descrito por Cancino et al. (2021): una mejor conectividad e infraestructura favorece la consolidación de núcleos rurales de mayor densidad, fenómeno también observado en la macrozona norte.

Con el propósito de integrar los resultados a distintas escalas, la Tabla 18 resume los principales índices de clusterización a nivel comunal, incorporando la subclasificación de viviendas ruidas (permanente y satélite) y el índice de clusterización. En la tabla se observa que es la comuna de Puerto Montt la que presenta un mayor número total de viviendas (29.632), clusterizadas (19.964) y de ruido (9.668). Por su parte, al analizar la subclasificación de las viviendas ruidas, destaca la comuna de Curaco de Vélez con

mayor porcentaje de viviendas ruido “permanente” (86,3%), y Castro (100%) con mayor porcentaje de viviendas clasificadas como “satélite”, dado que no se encontraron viviendas ruido permanente, siendo en todos los casos satélite. Por último, de igual forma las comunas de Castro y Puqueldón son las que presentan el mayor índice de clusterización (90 y 89,9%, respectivamente).

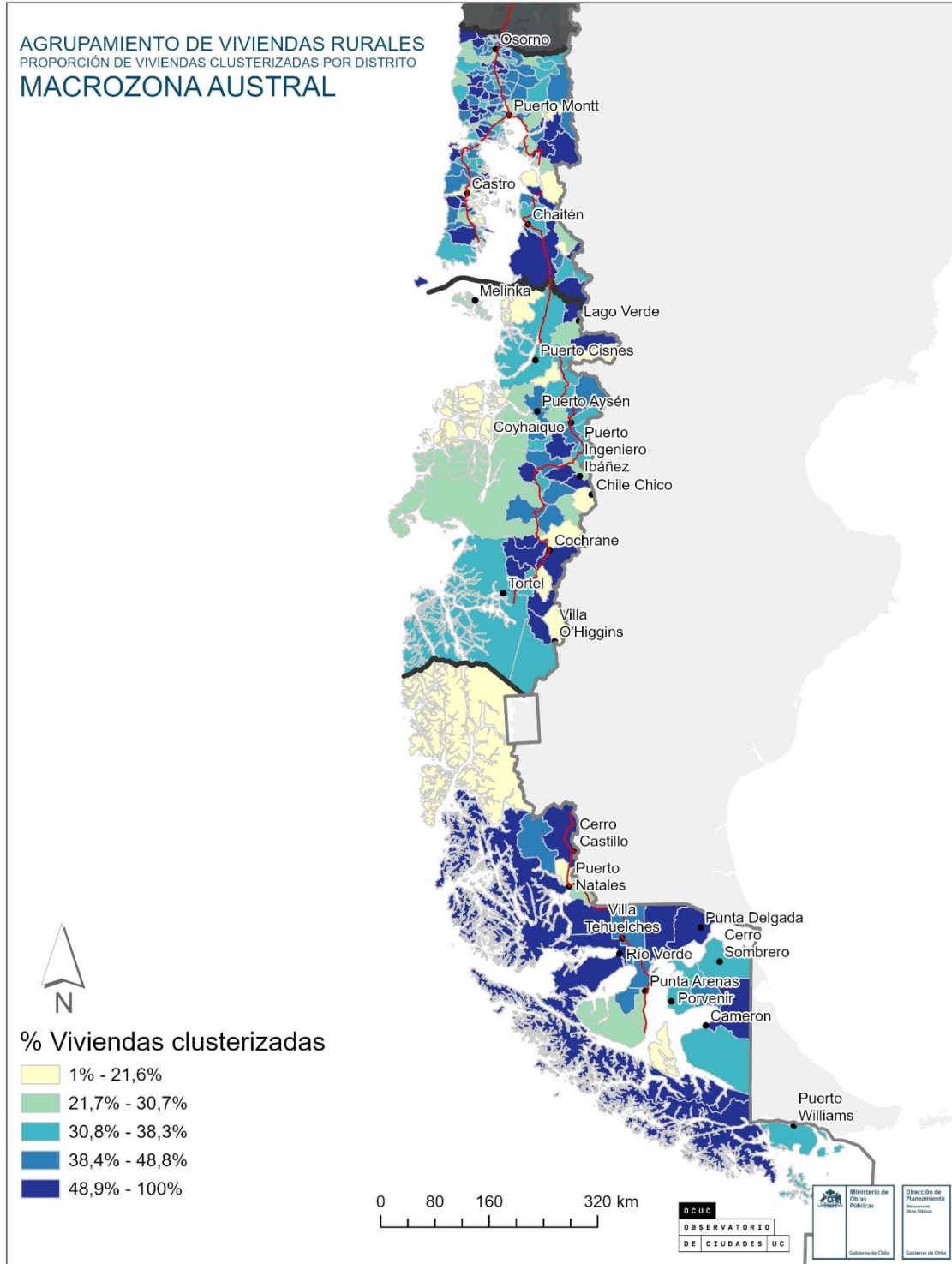
Tabla 18: Indicadores de clusterización a nivel comuna, macrozona austral.

Región	Provincia	Comuna	Total De Viviendas	Viviendas Clusterizadas	Total Ruido	Viviendas Ruido Permanente	Viviendas Ruido Satélites	Índice De Clusterización
Los Lagos	Chiloé	Ancud	9.953	5.605	4.348	79	21	56
		Castro	8.161	7.344	817	0	100	90
		Chonchi	7.953	4.243	3.710	81	19	53
		Curaco De Vélez	1.896	934	962	86	14	49
		Dalcahue	5.569	2.847	2.722	81	19	51
		Puqueldón	3.134	2.817	317	1	99	90
		Queilén	3.437	2.964	473	29	71	86
		Quellón	7.825	4.885	2.940	75	25	62
		Quemchi	5.127	3.129	1.998	74	26	61
	Quinchao	3.863	3.306	557	26	74	86	
	Llanquihue	Calbuco	10.649	7.047	3.602	73	27	66
		Cochamó	3.346	2.037	1.309	73	27	61
		Fresia	3.938	2.005	1.933	78	22	51
		Frutillar	4.716	2.840	1.876	76	24	60
		Llanquihue	2.401	1.287	1.114	83	17	54
		Los Muermos	6.627	3.219	3.408	84	16	49
		Mauilín	6.024	3.628	2.396	77	23	60
		Puerto Montt	29.632	19.964	9.668	72	28	67
	Osorno	Puerto Varas	14.413	10.343	4.070	65	35	72
		Osorno	8.931	5.379	3.552	77	23	60
		Puerto Octay	5.213	3.084	2.129	75	25	59
		Purranque	4.837	2.742	2.095	77	23	57
		Puyehue	5.821	3.721	2.100	73	27	64
		Río Negro	4.876	3.003	1.873	77	23	62
		San Juan De La Costa	5.579	3.318	2.261	78	22	60
	Palena	San Pablo	4.797	2.798	1.999	81	20	58
		Chaitén	2.490	1.675	815	67	33	67
		Futaleufú	769	516	253	76	24	67
Hualaihué		4.288	2.834	1.454	72	28	66	
Aysén	Aysén	Palena	856	509	347	77	23	60
		Aysén	2.777	1.872	905	72	28	67
		Cisnes	2.372	1.677	695	63	37	71
	Capitán Prat	Guaitecas	134	97	37	68	32	72
		Cochrane	752	424	328	84	16	56
		O'higgins	551	408	143	67	33	74
	Coyhaique	Tortel	474	323	151	62	38	68
		Coyhaique	11.328	7.024	4.304	76	24	62
	General Carrera	Lago Verde	947	670	277	72	29	71
Chile Chico		3.058	2.088	970	68	32	68	
Magallanes	Antártica Chilena	Río Ibáñez	3.036	1.968	1.068	68	32	65
		Cabo De Hornos	161	95	66	82	18	59
		Laguna Blanca	457	276	181	82	18	60
		Punta Arenas	8.161	5.720	2.441	69	31	70
		Río Verde	311	92	219	76	24	30
San Gregorio	446	154	292	86	14	35		

Región	Provincia	Comuna	Total De Viviendas	Viviendas Clusterizadas	Total Ruido	Viviendas Ruido Permanente	Viviendas Ruido Satélites	Índice De Clusterización
	Tierra Del Fuego	Porvenir	2.022	1.185	837	68	32	59
		Primavera	1.696	1.106	590	64	36	65
		Timaukel	602	374	228	70	30	62
	Última Esperanza	Natales	3.408	2.561	847	58	43	75
		Torres Del Paine	634	319	315	73	27	50

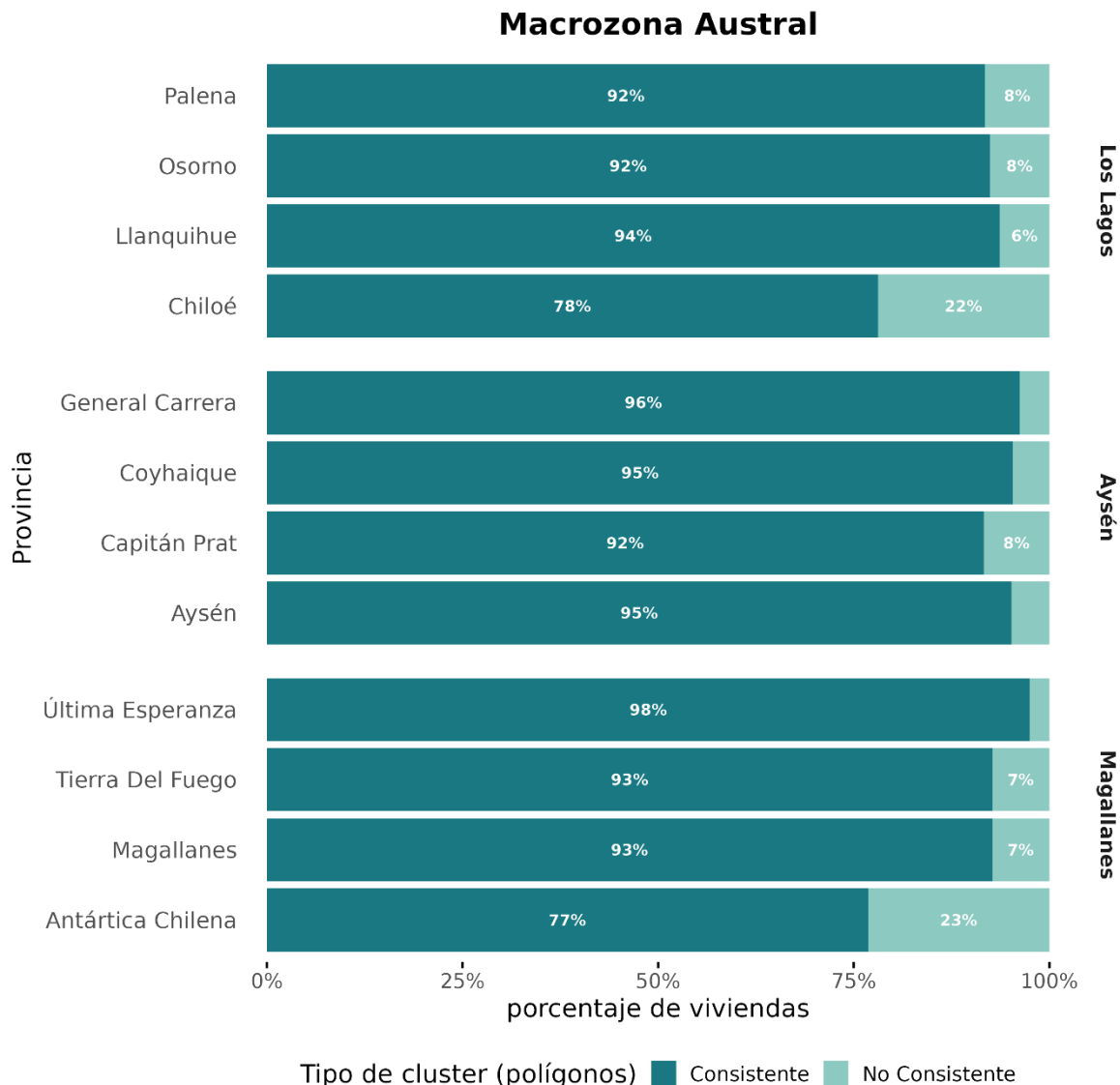
Los resultados mencionados anteriormente se georreferencian a continuación, en la Cartografía 6 donde se evidencia la proporción de viviendas clusterizadas según el distrito censal en la macrozona austral. En la cartografía es posible observar que las regiones de Aysén y Magallanes presentan algún área sin clasificar, las cuales se encuentran en sectores costeros. En cambio, los mayores porcentajes de clusterización se observan en los núcleos urbanos, donde las condiciones geográficas, de infraestructura y climáticas son mayormente favorables.

Cartografía 6: Agrupamiento de viviendas en la Macrozona Austral



Al analizar los clusters según su grado de consistencia (Figura 47), la mayoría de las provincias presenta alta consistencia interna de los clusters. Los valores van desde Antártica Chilena (77%) hasta Última Esperanza (98% cada una). Los territorios con mayor inconsistencia relativa son Antártica Chilena (23%), Chiloé (17%) y Tierra del Fuego (14%). En términos prácticos, esto indica que, salvo esas excepciones, los clústeres agrupan viviendas con densidades locales similares, con menos de 10% de inconsistencia en la mayoría de los casos; es decir, el patrón interno de cada agrupación es estable y homogéneo (a excepción de Antártica Chilena y Chiloé).

Figura 47: Proporción de tipos de clúster según consistencia en Macrozona Austral.

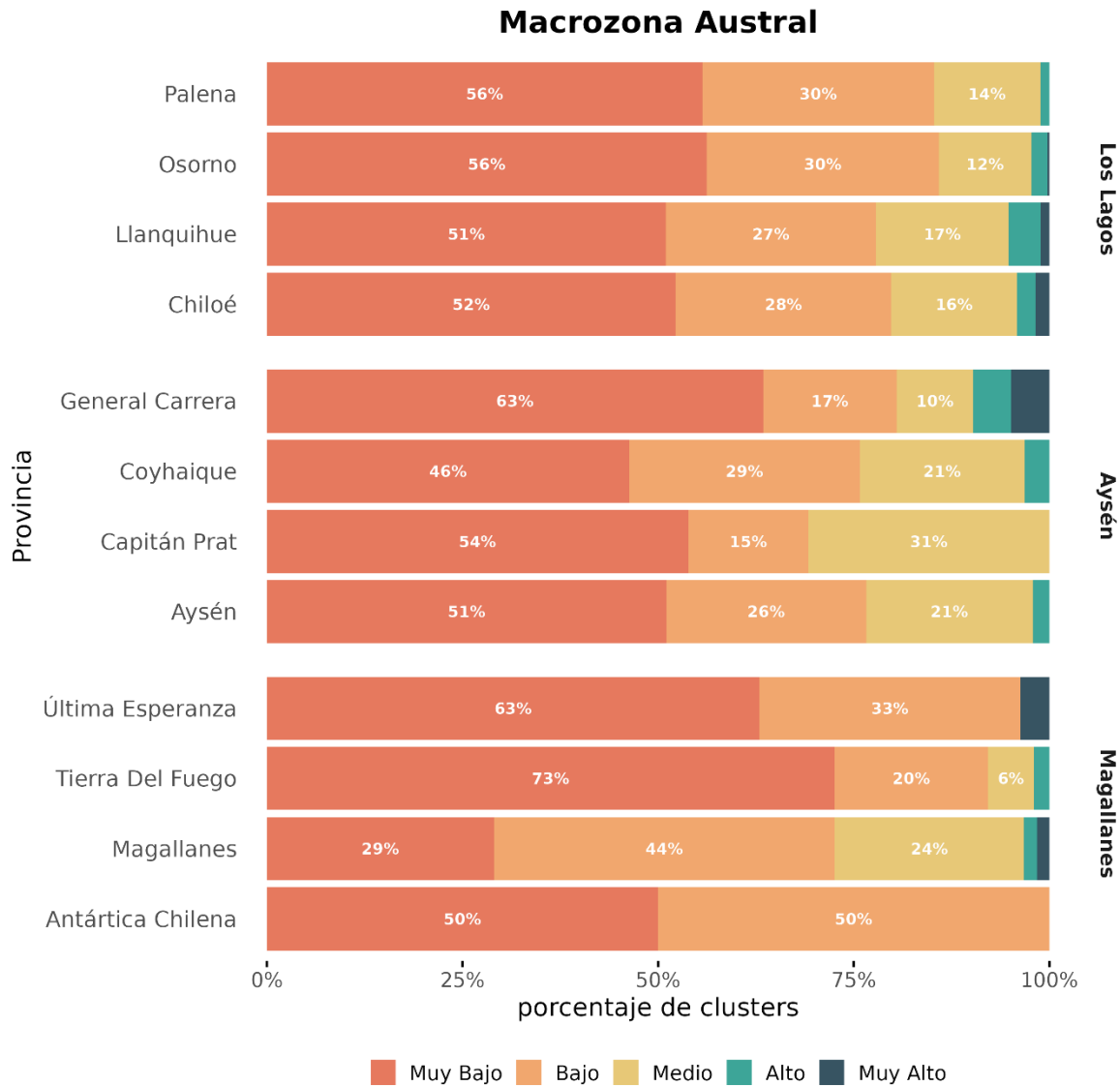


Las etiquetas de los segmentos menores al 5% fueron removidas para mejorar la visualización del gráfico

Finalmente, si se clasifican los clústeres según su tamaño de agrupamiento (Figura 48), se evidencia predominancia de la clase de tamaño “Muy bajo”, aunque con amplitud territorial: dese 73% de los clústeres en Tierra del Fuego hasta 29% en Magallanes

(provincia). Le siguen con frecuencia las clases “Bajo” y “Medio”, lo que sugiere que el grueso de las agrupaciones concentra entre 50 y 300 viviendas. Las clases “Alto” ( $\geq 300$  viviendas) y “Muy Alto” ( $\geq 600$  viviendas) son minoritarias, aunque es posible encontrar clústeres con estas categorías pese al contexto geográfico restrictivo.

Figura 48: Clasificación de los clústeres según su tamaño de agrupamiento en la Macrozona Austral.



Las etiquetas de los segmentos menores al 5% fueron removidas para mejorar la visualización del gráfico

Por último, la Tabla 19 muestra el detalle de los índices de clasificación de clústeres según su tamaño de agrupamiento y nivel de consistencia, organizados a escala comunal, con el propósito de facilitar la comprensión de los resultados en las distintas escalas de análisis. En este sentido, la comuna de Los Ángeles es la que presenta la mayor cantidad de clusters (283). Además, las comunas que presentan mayor porcentaje en cada subclasificación por tamaño son: San Pedro De la Paz con “Muy bajo” (100%), Talcahuano con “Bajo” (60%), Concepción con “Medio” (38,9%), Melipeuco con

“Alto” (18,8%) y Padre Las Casas con “Muy alto” (50%). Finalmente, Negrete es la comuna con los clusters con mayor consistencia (95,2% consistencia), en cambio Lota y San Pedro De La Paz son las que menos consistencia presentan (33,3% inconsistencia).

Tabla 19: Indicadores de clasificación a nivel de clusters, macrozona austral.

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
Los Lagos	Chiloé	Ancud	86	47,7	36	16,3	0	0	94,2	5,8
		Castro	2	0	0	0	0	100	50	50
		Chonchi	62	50	35,5	12,9	1,6	0	93,5	6,5
		Curaco De Vélez	11	54,5	27,3	9,1	9,1	0	90,9	9,1
		Dalcahue	47	55,3	27,7	17	0	0	93,6	6,4
		Puqueldón	2	0	0	50	0	50	50	50
		Queilén	9	66,7	22,2	0	0	11,1	88,9	11,1
		Quellón	76	63,2	15,8	21,1	0	0	94,7	5,3
		Quemchi	32	53,1	28,1	6,2	9,4	3,1	93,8	6,2
	Quinchao	10	10	10	40	30	10	90	10	
	Llanquihue	Calbuco	62	46,8	25,8	21	3,2	3,2	93,5	6,5
		Cochamó	26	57,7	23,1	15,4	3,8	0	92,3	7,7
		Fresia	33	63,6	24,2	12,1	0	0	93,9	6,1
		Frutillar	36	52,8	25	19,4	2,8	0	94,4	5,6
		Llanquihue	23	56,5	30,4	13	0	0	91,3	8,7
		Los Muermos	51	56,9	29,4	11,8	2	0	94,1	5,9
		Mauullín	52	55,8	28,8	13,5	1,9	0	94,2	5,8
		Puerto Montt	184	49,5	25	16,8	7,1	1,6	94,6	5,4
	Puerto Varas	74	40,5	31,1	23	4,1	1,4	94,6	5,4	
	Osorno	Osorno	75	49,3	30,7	17,3	2,7	0	94,7	5,3
		Puerto Octay	49	69,4	20,4	6,1	4,1	0	93,9	6,1
		Purranque	43	46,5	46,5	7	0	0	93	7
		Puyehue	49	65,3	22,4	8,2	2	2	93,9	6,1
		Río Negro	45	48,9	40	8,9	2,2	0	93,3	6,7
		San Juan De La Costa	43	51,2	25,6	20,9	2,3	0	93	7
	Palena	San Pablo	50	64	24	12	0	0	94	6
		Chaitén	28	53,6	32,1	14,3	0	0	92,9	7,1
Futaleufú		9	55,6	33,3	11,1	0	0	88,9	11,1	
Hualaihué		45	60	26,7	11,1	2,2	0	93,3	6,7	
Aysén	Aysén	Palena	6	33,3	33,3	33,3	0	0	83,3	16,7
		Aysén	25	52	28	16	4	0	92	8
		Cisnes	19	42,1	26,3	31,6	0	0	94,7	5,3
	Capitán Prat	Guaitecas	3	100	0	0	0	0	66,7	33,3
		Cochrane	6	66,7	16,7	16,7	0	0	83,3	16,7
		O'Higgins	4	50	25	25	0	0	75	25
	Coyhaique	Tortel	3	33,3	0	66,7	0	0	66,7	33,3
		Coyhaique	88	46,6	30,7	19,3	3,4	0	94,3	5,7
	General Carrera	Lago Verde	7	42,9	14,3	42,9	0	0	85,7	14,3
		Chile Chico	22	63,6	18,2	9,1	4,5	4,5	90,9	9,1
Río Ibáñez		19	63,2	15,8	10,5	5,3	5,3	94,7	5,3	
Magallanes	Antártica Chilena	Cabo De Hornos	2	50	50	0	0	0	50	50
	Magallanes	Laguna Blanca	4	25	50	25	0	0	75	25
		Punta Arenas	54	29,6	40,7	25,9	1,9	1,9	94,4	5,6
		Río Verde	2	50	50	0	0	0	50	50
		San Gregorio	2	0	100	0	0	0	50	50

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters (%)					Nivel de consistencia (%)	
				Muy bajo	Bajo	Medio	Alto	Muy alto	Cons.	Incons.
	Tierra Del Fuego	Porvenir	20	70	15	10	5	0	95	5
		Primavera	23	78,3	17,4	4,3	0	0	91,3	8,7
		Timaukel	8	62,5	37,5	0	0	0	87,5	12,5
	Última Esperanza	Natales	19	57,9	36,8	0	0	5,3	94,7	5,3
		Torres Del Paine	8	75	25	0	0	0	87,5	12,5

## 7. FOCALIZACIÓN DE LA INVERSIÓN

En este capítulo se presenta un modelo territorial, implementado en una Toolbox, para caracterizar los territorios rurales del país y focalizar la inversión en infraestructura pública. El enfoque integra variables ambientales, socioeconómicas, demográficas y de accesibilidad. La focalización se entiende como el diseño, evaluación e inversión en proyectos de infraestructura basados en evidencia cuantitativa y cualitativa, con el fin de maximizar el impacto social y económico, reducir ineficiencias y mejorar la equidad territorial.

Durante el estudio se evaluaron diversas alternativas y enfoques metodológicos para focalizar, priorizar y caracterizar la inversión pública, realizada por el MOP en territorios rurales. Estas aproximaciones se alinearon con diseños estándar (principalmente el análisis multicriterio) y la caracterización mediante técnicas de cluster, cada uno con ventajas y limitaciones.

El análisis multicriterio (Gomez et al., 2005; Voogd, 1982) permite priorizar de forma directa a través de puntajes ponderados por dimensiones y/o variables, identificando qué territorios requieren mayor atención. Sin embargo, al momento de explicar la brecha específica o la razón de dicha prioridad, la información suele quedar diluida en ponderaciones y promedios, lo que dificulta la correcta visualización de los factores que afectan el puntaje y, en consecuencia, pueden comprometer el objetivo inicial del modelo.

Ante ello, se propone una aproximación mixta modular, evaluando distintos elementos en cada sección, graficada en la Figura 49. Siendo las siguientes fases:

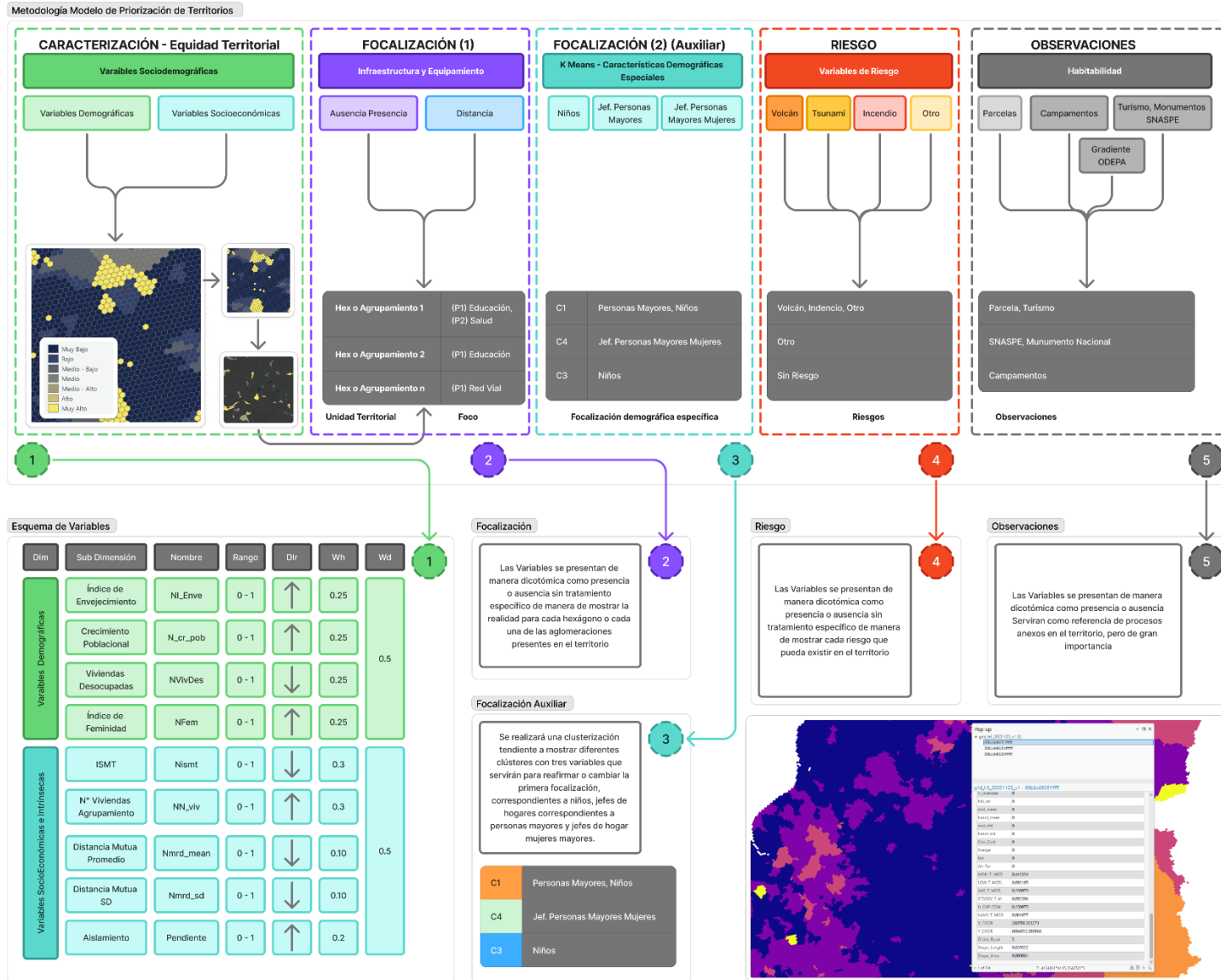
1. Caracterización de la equidad territorial, basada en dos dimensiones – demográfica y socioeconómica– mediante análisis multicriterio.
2. Focalización de infraestructura y equipamiento, con variables dicotómicas (presencia/ausencia), para identificar tipologías inexistentes o a gran distancia y vincular esas carencias con la equidad territorial desde la perspectiva de la infraestructura.
3. Focalización auxiliar mediante *k-means*, orientada a caracterizar grupos prioritarios (adultos mayores jefes/jefas de hogar y menores de 14 años).
4. Evaluación de riesgos y atributos rurales relevantes, considerando la presencia/ausencia de riesgos específicos y características como campamentos, parcelas de agrado y la categoría urbano-rural (ODEPA, 2025), entre otros.

El modelo considera las variables presentadas en la Tabla 20, organizadas en dimensiones. Como se señaló, estas dimensiones no solo cumplen un rol de priorización; según su tipología, determinan la incorporación al modelo mediante análisis multicriterio o caracterización discreta. Además, aportan observaciones

complementarias para la focalización, enriqueciendo la interpretación de brechas y necesidades territoriales.

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, etc.). No se limita al ingreso de los hogares o pobreza, sino de un paquete de atributos demográficos y socioeconómicos intrínsecos que posicionan a cada territorio en mejores o peores condiciones de partida para acceder a servicios, políticas e infraestructura. El indicador por sí sólo no evalúa la priorización, pero se realiza un esquema metodológico que encadena todos los factores de un esquema multicriterio.

Figura 49: Esquema general del modelo de Priorización de Territorios



*Tabla 20: Variables modelo de focalización*

<b>Categoría</b>	<b>Nombre variable</b>	<b>Unidad</b>	<b>Fuente</b>	<b>Escala</b>
Infraestructura y equipamiento	Establecimientos de salud	Distancia (m)	MINEDUC (2021), MINSAL (2024)	Subcomunal
Infraestructura y equipamiento	Establecimientos educacionales	Distancia (m)	MINEDUC (2021)	Subcomunal
Infraestructura y equipamiento	Planta de tratamiento de aguas servidas (AS)	Presencia/ausencia	SISS 2022	Subcomunal
Infraestructura y equipamiento	Territorios operacionales	Presencia/ausencia	SISS 2025	Subcomunal
Infraestructura y equipamiento	Red vial	Presencia/ausencia	MOP 2025	Subcomunal
Infraestructura y equipamiento	Aeródromos	Distancia (m)	MOP 2014	Subcomunal
Infraestructura y equipamiento	Embalses	Presencia/ausencia	MOP 2024	Subcomunal
Infraestructura y equipamiento	Infraestructura portuaria y costera	Distancia (m)	MOP 2023	Subcomunal
Infraestructura y equipamiento	Puentes	Presencia/ausencia	MOP 2025	Subcomunal
Infraestructura y equipamiento	Red Aeroportuaria	Distancia (m)	MOP 2025	Subcomunal
Infraestructura y equipamiento	Agua Potable (SSR)	Presencia/ausencia	MOP (2024)	Subcomunal
Demografía	Variación cantidad de habitantes	Porcentaje	INE (2017-2024)	Comunal
Demografía	Índice de envejecimiento	Porcentaje	INE (2017-2024)	Comunal
Demografía	Viviendas desocupadas	Cantidad	INE (2017-2024)	Comunal
Demografía	Variación cantidad de viviendas	Porcentaje	INE (2017-2024)	Comunal
Socioeconómico	Indicador Sociomaterial Territorial	Índice normalizado	OCUC (2017)	Entidad rural
Demografía	Índice de Feminidad	Tasa (ratio)	INE (2024)	Subcomunal
Demografía	Jefatura de Hogar - Personas Mayores	Cantidad	INE (2024)	Subcomunal
Habitabilidad	Campamentos	Presencia/ausencia	MINVU (2024)	Subcomunal
Habitabilidad	Parcelas de agrado	Presencia/ausencia	MINVU (2024)	Subcomunal
Amenazas naturales	Amenaza volcánica	Presencia/ausencia	SERNAGEOMIN (2024)	Intercomunal
Amenazas naturales	Amenaza de tsunami e inundaciones (cota 30)	Presencia/ausencia	IGM (2023)	Intercomunal
Amenazas naturales	Recurrencia de incendios forestales	Cantidad	CONAF (2024)	Intercomunal
Amenazas naturales	Afectaciones a la infraestructura (SIEMOP)	Presencia/ausencia	MOP	Subcomunal
Amenazas naturales	Catastro de remociones en masa	Presencia/ausencia	SENAPRED (2024)	Subcomunal

## 7.1. Estandarización de las variables

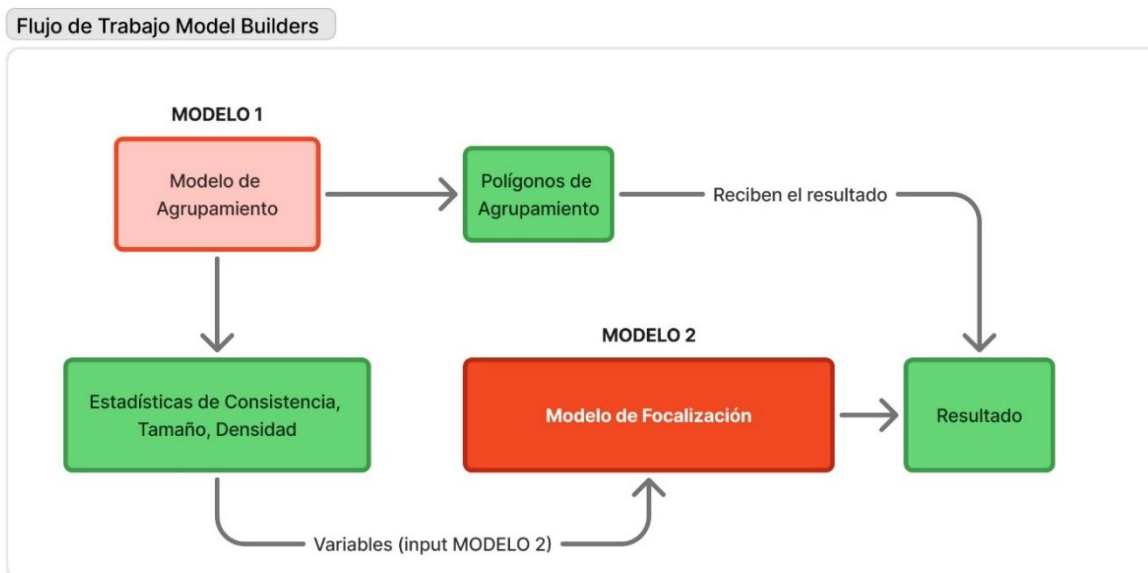
Las variables señaladas en la Tabla 20, fueron procesadas siguiendo los siguientes criterios:

- a) Para el caso de las variables de presencia / ausencia los valores siempre fueron tradicionalmente etiquetados bajo la regla dicotómica 0 y 1. Esto dio pie para estandarizar el resto de las variables con el criterio mínimo máximo (0-1)
- b) Las variables de distancia no cubren todo el territorio nacional (600.000 del 1.000.000 de orígenes) para este caso se dejó todas las variables sin conectividad con un valor sobre el máximo. Esta medida se está analizando desde el punto de vista de la distribución ya que produce algunas inconsistencias derivadas de sesgo y curtosis por la cantidad de valores máximos.
- c) Para las variables con distribuciones muy alejadas de la normal se realizó una transformación Box-Cox.
- d) Todas las variables, sin excepción fueron normalizadas entre 0 y 1.
- e) Las variables con tendencia contrarias o negativas se les aplicó su inverso, es decir  $1-x$ , con lo que su distribución queda entre 0-1 de acuerdo con su concepto.
- f) Para efectos de entendimiento las variables originales fueron dejadas en las bases e inmediatamente al lado, tienen su versión normalizada con el sufijo N. Este punto no aplica para las variables de presencia y ausencia.

## 7.2. Método de cálculo

La Figura 50 presenta el esquema metodológico que permite ver el flujo de trabajo con el que el Model Builder opera. Este modelo será empaquetado para su ejecución automática considerando eventuales cambios en el modelo de agrupamiento. Esto significa que para todos los efectos el modelo de agrupamiento será considerado como un input en el modelo de focalización y priorización (Figura 50). Asimismo, considerarán instrucciones para reemplazar variables en el caso de la información secundaria, así como un instructivo, para la normalización, direccionamiento y eventual transformación de distribuciones de estas.

Figura 50: Flujo de trabajo Model Builders.



En el esquema (Figura 50) se evidencian en primer lugar la bajada de la priorización, que está constituida por las dimensiones demográficas y socioeconómica. Estas variables median la vulnerabilidad como es el caso del ISMT, el índice de envejecimiento, el índice de feminidad y los hogares monoparentales con jefaturas mujeres mayores a 65 años. Además, se acompaña de variables intrínsecas de tamaño derivadas del modelo de agrupamiento. Esta dimensión opera a través de análisis multicriterio y permite obtener un índice de priorización que es alojado a escala de los hexágonos y posteriormente se lleva a escala de los polígonos de agrupamiento.

De esta manera, este primer resultado nos entrega la prioridad del modelo, es decir lugares que, por su vulnerabilidad y tamaño, deben ser motivo de prioridad.

El segundo apartado evalúa la focalización y para esto analiza de forma discreta las brechas de los territorios, principalmente relevando la situación de presencia y ausencia y/o distancia a equipamientos e infraestructura crítica. Con esto podemos poner un apellido a la vulnerabilidad, respondiendo la pregunta de que es necesario focalizar. Esto se entregará a partir de una caracterización discreta que además ordenará las primera tres prioridades de los territorios en cuestión, por ejemplo, expresando como primera prioridad la necesidad de colegios (p1) y embalses (p2).

### 7.3. Escalas de análisis

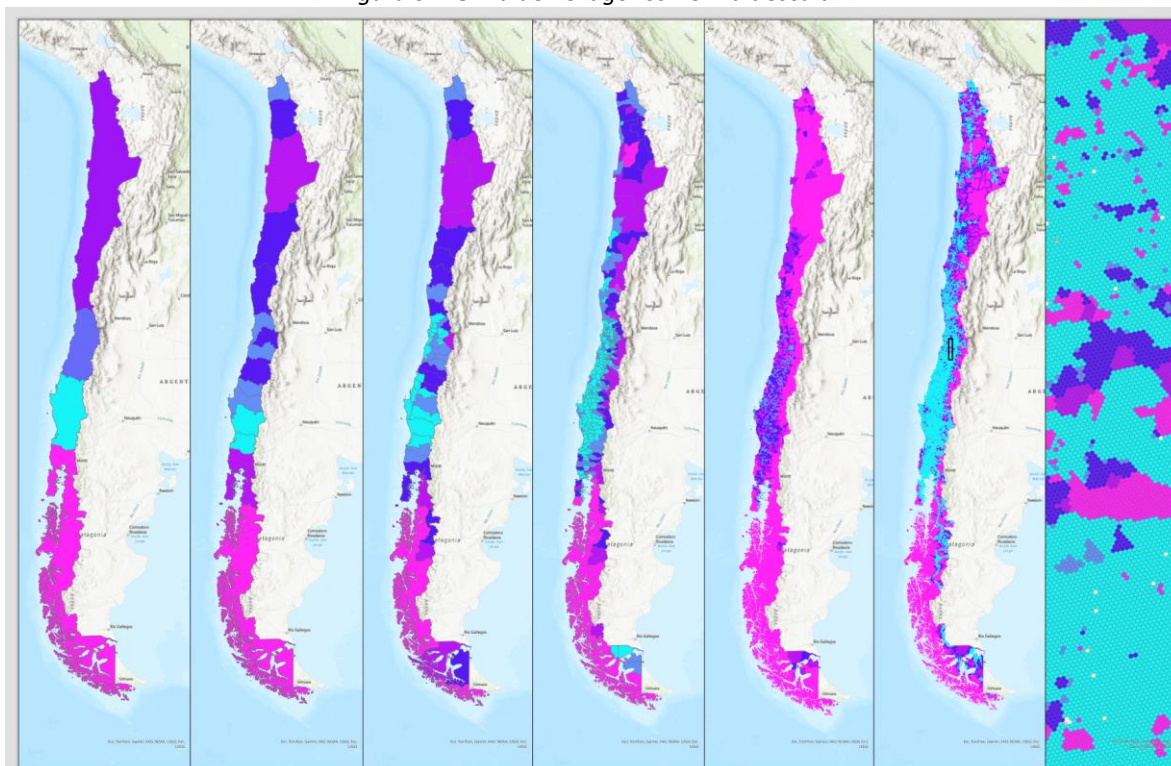
Bajo el objetivo de obtener representatividad de la información secundaria a priorizar y dada la gran variedad de datos, se escogió una unidad especial que fuese regular, de escala acotada y no dependiente de las divisiones administrativas. Para ello se utilizó el sistema H3 —específicamente en resolución L8—, por su carácter exhaustivo a nivel nacional, su identificador único e inmutable y su facilidad para anexar otras unidades de referencia. La resolución L8 entrega una malla de hexágonos regulares que cubre todo

el territorio nacional; cada celda tiene un área aproximada de 0,74 km<sup>2</sup> (≈74 ha), lo que se traduce en poco más de 1,1 millones de hexágonos para Chile.

Sobre cada hexágono se agregaron atributos de escalas administrativas y operativas (distrito censal, comuna, provincia, región y macrozonas ministeriales), manteniendo la trazabilidad entre niveles (Figura 51). Esta trazabilidad permite que cada unidad espacial (hexágono) pueda vincularse de manera estable y reproducible con las distintas escalas administrativas, al estar asociadas como atributos. Esto permite interactuar con las escalas según sea el interés del análisis. Los resultados de priorización estarán disponibles para todo el territorio.

En el flujo del modelo, la malla H3 – L8 sirve como soporte común para integrar los insumos de los modelos de agrupamiento y las salidas en la escala final definida. Sin perjuicio de lo anterior, el modelo generará también salidas complementarias (no finales) en todas las unidades espaciales mencionadas, junto con una capa general a nivel L8 que conserva el máximo detalle.

Figura 51: Grilla de hexágonos H3 multiescalar.

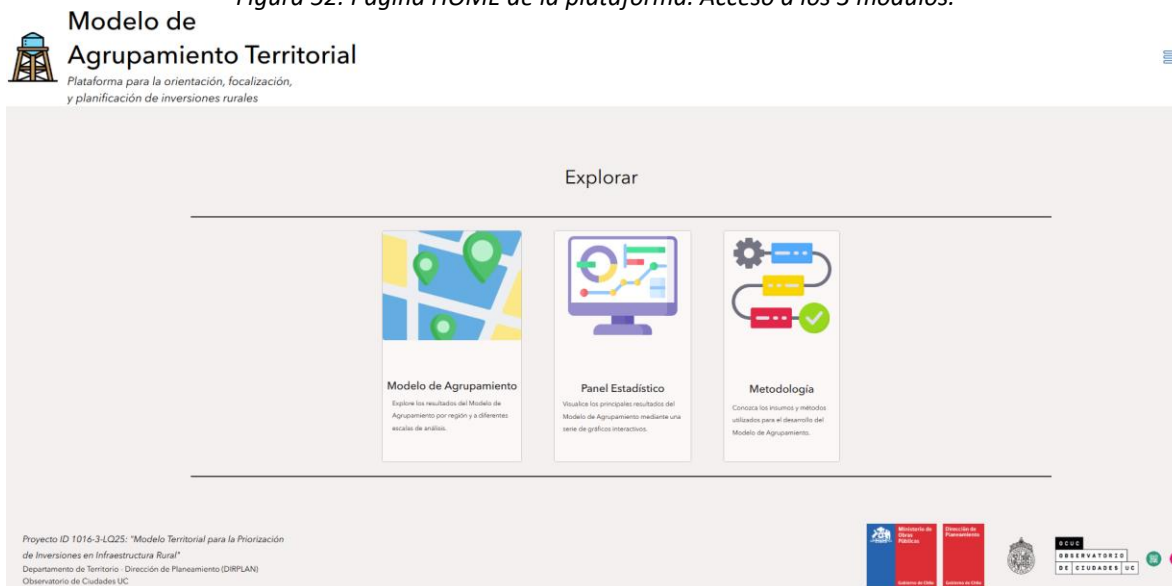


#### 7.4. Visualización de resultados

Con el objetivo de hacer operativos los resultados del ejercicio de focalización de la inversión, se implementó una plataforma WebGIS pública en ArcGIS Online (ESRI). La visualización fue diseñada como una experiencia guiada que combina navegación tipo portal (herramienta Experience Builder) con visores de exploración y síntesis,

permitiendo pasar desde una lectura general de resultados a una interpretación de detalle por territorio y clúster.

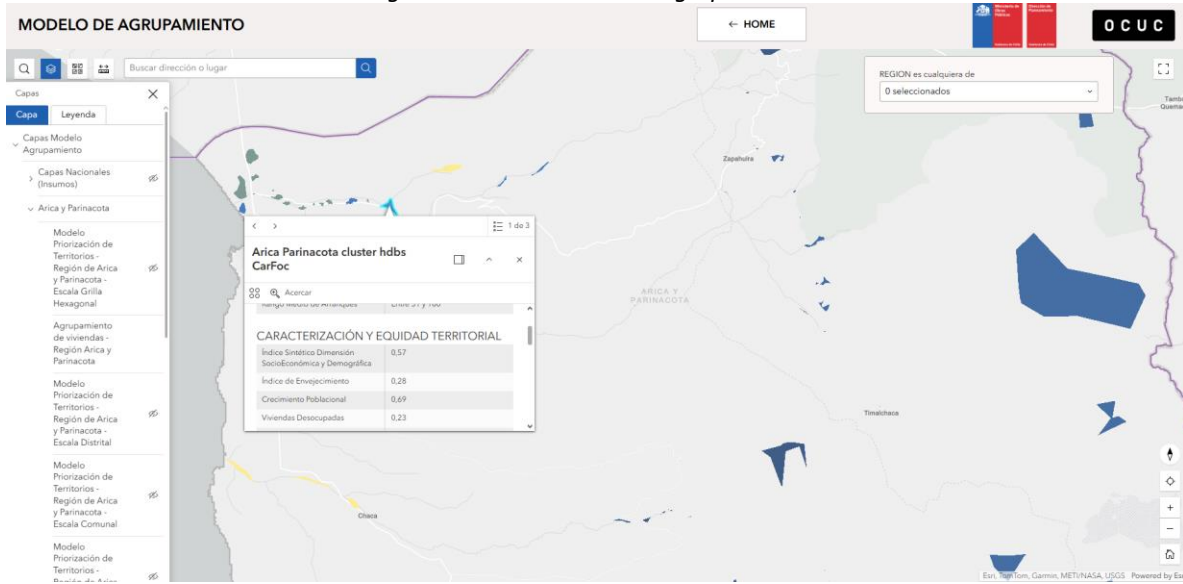
Figura 52: Página HOME de la plataforma. Acceso a los 3 módulos.



La plataforma se organiza desde una página de inicio titulada 'Modelo de Agrupamiento Territorial - Plataforma para la orientación, focalización, y planificación de inversiones rurales', que funciona como menú principal. Desde ahí se accede a tres módulos, cada uno con un rol distinto (exploración espacial / síntesis estadística / respaldo metodológico):

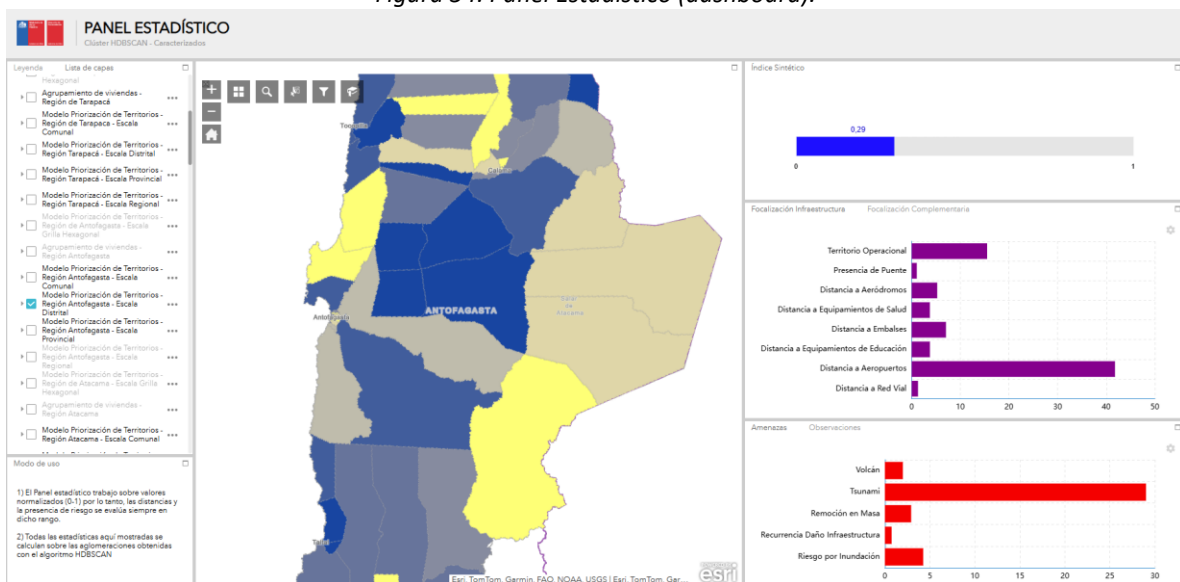
1. Modelo de Agrupamiento: visor cartográfico orientado a la exploración espacial de resultados por región y a distintas escalas de análisis (grilla hexagonal H3/regional/provincial/comunal/distrital/clústeres). Permite visualizar la clasificación del Índice Sintético (dimensión socioeconómica y demográfica) en categorías (de muy bajo a muy alto con herramientas básicas de navegación y de manera interactiva. Incluye filtro regional, facilitando comparaciones y lectura rápida por territorios. Además, se puede indagar en la información contenida de cada entidad espacial agrupada en seis bloques: Identidad de Agrupamiento / Caracterización y Equidad Territorial / Focalización Infraestructura y Equipamiento / Focalización Complementaria / Amenazas / Observaciones.

Figura 53: Visor Modelo de Agrupamiento.



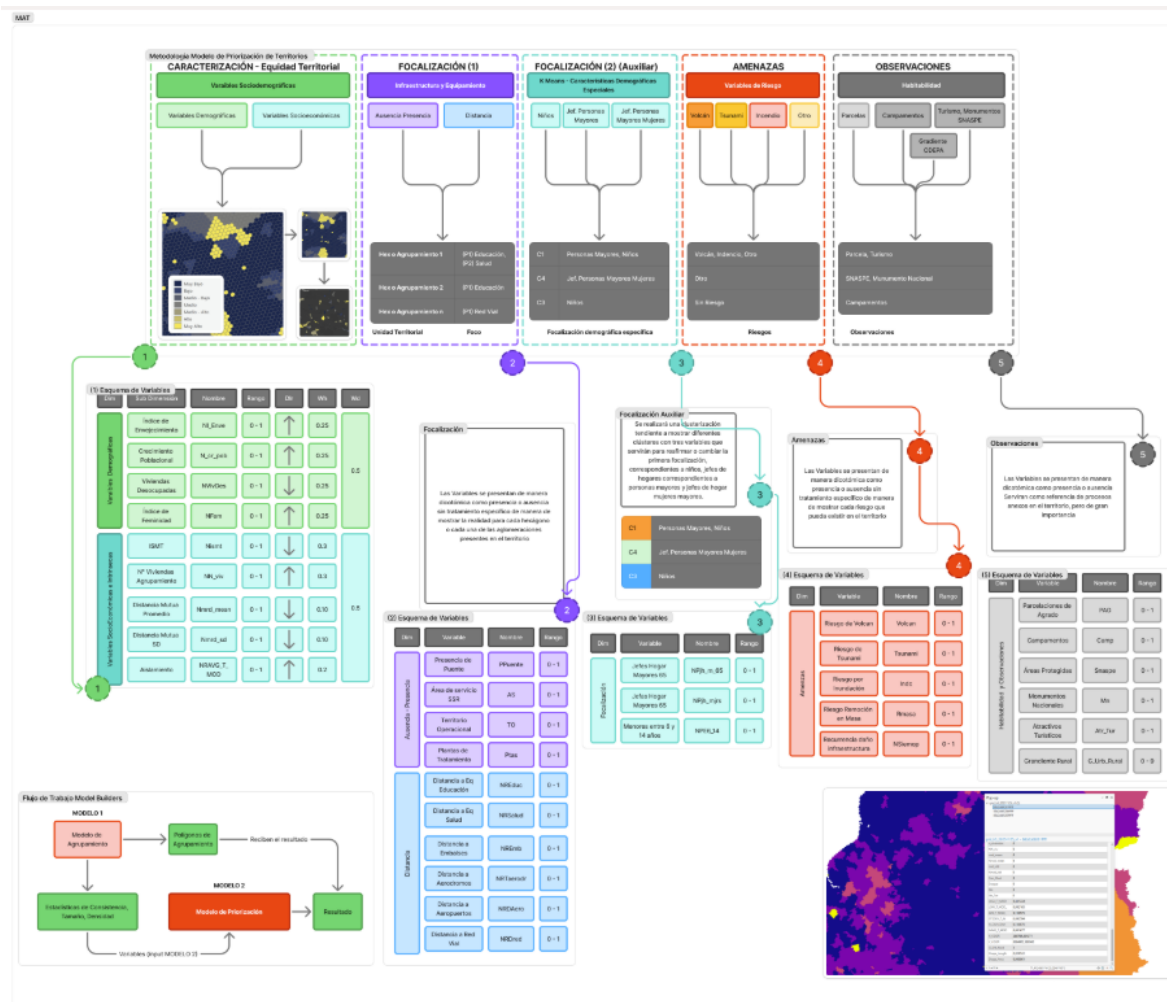
2. Panel Estadístico: visor analítico que integra mapa y gráficos interactivos para interpretación, según el espacio geográfico seleccionado. Componiéndose, además de las herramientas típicas de un visor, de un panel de síntesis con Índice Sintético (barra 0–1), gráficos de Focalización de Infraestructura y Focalización Complementaria, y un bloque adicional de Amenazas/Observaciones, que contextualiza la priorización.
3. Metodología: sección de soporte que explica insumos, variables y lógica del modelo desarrollado, reforzando la transparencia del resultado y su adecuado entendimiento para los usuarios de perfil técnico.

Figura 54: Panel Estadístico (dashboard).



La información se organiza para una lectura tipo “ver–filtrar–explicar”, donde el usuario parte desde el HOME y selecciona el módulo según su necesidad (explorar resultados en mapa, revisar síntesis estadística o consultar la metodología), luego acota el análisis filtrando por región y activando capas según escala en el visor de Modelo de Agrupamiento o controlando capas y escala en el Panel Estadístico, y finalmente, al seleccionar un territorio o clúster, la plataforma presenta evidencia comparable en un orden lógico: primero el Índice Sintético (0–1), después las brechas relevantes de focalización (infraestructura y complementaria) y, como contexto, amenazas y observaciones, mientras los pop-ups aportan atributos del territorio (comuna, consistencia del clúster y métricas asociadas) para asegurar trazabilidad y comprensión del resultado.

Figura 55: Flujo metodológico integrado en plataforma.



Esta visualización apoya la focalización de inversión del MOP porque permite identificar territorios con mayor prioridad relativa, entender qué brechas específicas explican esa condición, facilitando comparar resultados entre territorios.

## 8. CREACIÓN DE UN TOOLBOX PARA ARCGIS PRO

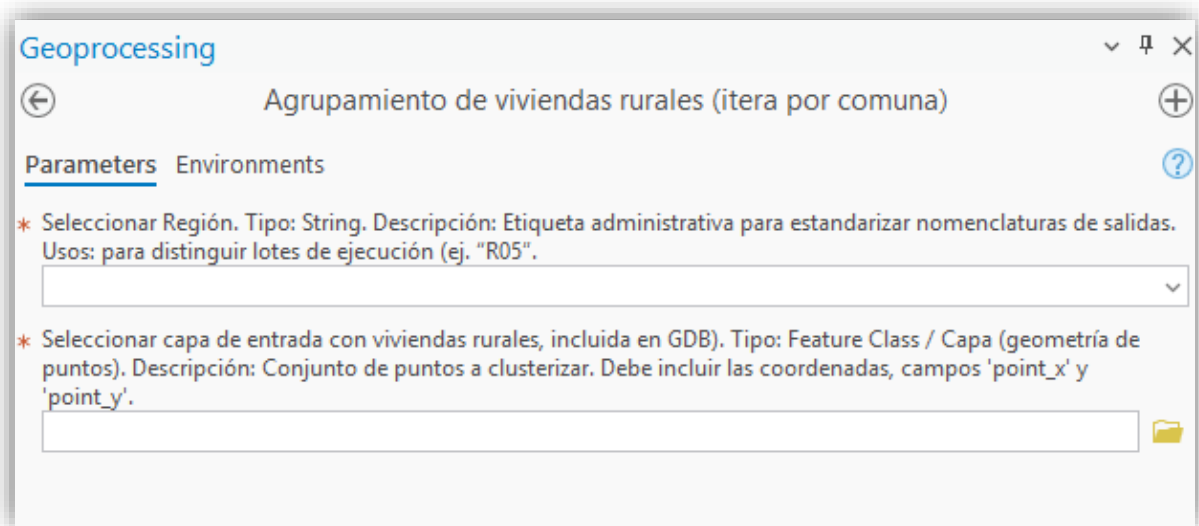
Considerando que el modelo de agrupamiento de viviendas basado en HDBSCAN se desarrolló nativamente en Python, se creó una herramienta tipo Toolbox que permite implementar este algoritmo de forma fácil e interactiva utilizando directamente la aplicación ArcGIS Pro.

Esta herramienta facilita que cualquier usuario con acceso a ArcGIS Pro pueda recalcular de manera automática todo el flujo metodológico definido en la Figura 11, identificando agrupaciones de viviendas mediante HDBSCAN.

Esto se realiza mediante la detección estandarizada de clústeres de vivienda, la estimación de métricas de consistencia y la construcción de polígonos envolventes (concave hull) que representen el tejido territorial de cada agrupación. El ámbito de trabajo es regional, y el procesamiento se realiza comuna a comuna dentro de la región seleccionada. La metodología implementada ha sido detallada en la sección homóloga.

El flujo general de la herramienta comienza con la selección de la región a procesar, y de la capa de viviendas rurales de esa región. La herramienta verifica que el sistema de referencia cartográfico entrante sea SIRGAS-Chile 2016 (EPSG: 9153), en caso contrario y de manera automática corrige el sistema de referencia. De modo, que tanto el análisis como las salidas mantengan la coherencia espacial (Figura 56).

Figura 56. Interfaz de la herramienta contenida en Toolbox para la agrupación de viviendas



El núcleo del procesamiento utiliza el algoritmo HDBSCAN. Este método se ejecuta comuna a comuna, evitando sesgos cuando existen grandes diferencias de densidad entre territorios. Sobre las agrupaciones resultantes, la herramienta, calcula estadísticas descriptivas por clúster. Finalmente, con los puntos válidos de cada clúster

se genera un polígono concave hull, delimitando contornos de robusta fidelidad la forma del asentamiento definido.

Las salidas de la herramienta de la toolbox se escriben en la misma geodatabase que contiene el insumo espacial de las viviendas rurales de una región, dentro de un Feature Dataset creado automáticamente con el nombre de `Resultados_<REGIÓN>`. Se generan dos capas de resultados a nivel regional:

- Puntos de viviendas con atributos analíticos (identificador de clúster inicial y final, probabilidad, outlier, MRD, tipo de ruido).
- Polígonos (concave hull) por clúster, enriquecidos con las estadísticas calculadas.

En términos de interoperabilidad y mantenimiento, la toolbox se integra de forma nativa con ArcGIS Pro (formato .pyt) y está pensada para trabajar en un entorno clonado del Python del mismo ArcGIS Pro. Las dependencias analíticas principales son Geopandas, Hdbscan, Shapely y Fiona (además de Numpy/Pandas), todas instalables en el clon. Este enfoque permite replicar el entorno en otras estaciones de trabajo con pasos guiados (clonado del entorno base de Pro y posterior instalación de librerías), minimizando fricciones y garantizando reproducibilidad. Por otro lado, el diseño modular del pipeline –lectura, reproyección, agrupamiento por comuna, calibración, generación de hulls y exporte final- facilita ajustes selectivos cuando se requieran cambios de parámetros o nuevas métricas.

La herramienta ha sido concebida para uso operativo por equipos técnicos y tomadores de decisión. Su interfaz expone solo los parámetros necesarios (elección de región y capa de insumo), automatizando el resto (CRS, creación de Feature Dataset de resultados, nomenclatura de capas resultado, estadísticas, exportación final). De este modo se reduce la variabilidad entre usuarios y estandariza la producción de clústeres para su consumo directo en flujos SIG, dashboards o reportes temáticos. Finalmente, la toolbox está documentada con mensajes claros en la ventana de geoprocésamiento y convenciones de salida consistentes.

En documento *MANUAL USUARIO TOOLBOX MOP\_AGRUPACIÓN VIVIENDAS RURALES.pdf* se ha desarrollado un manual con el objetivo de que los usuarios técnicos de la herramienta tengan disponible toda la información necesaria para hacer uso de esta.

Del mismo modo, y, por último, se hace entrega de dos archivos Toolbox con dos objetivos diferenciados:

- **MODELO\_FOCALIZACIÓN\_INVERSIONES\_MOP\_versión\_SedesRegionales.pyt**: versión destinada a los funcionarios técnicos de sedes regionales DIRPLAN con una herramienta con menos opciones y más cerrada para obtener resultados más estandarizados al diseño del proceso general metodológico.

- **MODELO\_FOCALIZACIÓN\_INVERSIONES\_MOP\_versión\_SedeCentral.pyt:** se trata de una versión con algunos parámetros opcionales abiertos que da opción a desarrollar mayor capacidad de testeo de resultados y experimentación. Herramienta destinada a DIRPLAN central que actúa como contraparte técnica del proyecto, y por tanto con conocimiento sobre la metodología empleada y como se aplica.

## 9. CONSIDERACIONES

El modelo de agrupamiento implementado mediante el algoritmo HDBSCAN constituye una herramienta clave dentro del objetivo general del estudio: caracterizar territorialmente el grado de agrupamiento de las viviendas rurales en Chile, como base para orientar la planificación estratégica de infraestructura.

A diferencia de los métodos tradicionales este enfoque permite detectar agrupamientos sin imponer supuestos sobre la forma o tamaño de los clusters, lo que favorece una clasificación más ajustada a la diversidad de patrones de asentamiento rural.

Debido a que HDBSCAN se basa en una densidad jerárquica de puntos y no requerir un número predeterminado de clusters, los resultados obtenidos por este algoritmo son más robustos ya que permiten dar cuenta de las diferentes realidades que existen en cuanto a la distribución de viviendas rurales a lo largo del territorio.

A partir del trabajo realizado en terreno, la elaboración de un Modelo de Agrupamiento clusterizado y el Modelo de Focalización Territorial, se desprenden las siguientes conclusiones, hallazgos y consideraciones.

### a) Cobertura geoespacial de viviendas rurales como insumo estructural del modelo

Los resultados del Modelo de Agrupamiento dependen directamente de la calidad, consistencia y vigencia de la capa de viviendas rurales. La base utilizada en este estudio –derivada de la APC 2023 e integrada con las viviendas de Aldeas INE– representa el estado más completo disponible al momento del análisis, por lo que es indispensable mantener actualizada la capa geoespacial, más aún con la incertidumbre de que este producto siga siendo actualizado por el INE. Esto refuerza la necesidad de que DIRPLAN administre una capa maestra de viviendas rurales, actualizada periódicamente con nuevas fuentes (como Overture Maps u otros catastros sectoriales), de modo de asegurar una versión oficial y consistente. Sin esta capa estable, cualquier futura ejecución del modelo podría generar resultados no comparables con los de este estudio.

### b) Consideraciones metodológicas para el uso de HDBSCAN

El modelo confirma que HDBSCAN es un algoritmo robusto para capturar patrones de poblamiento rural en territorios altamente heterogéneos gracias a su capacidad de ajustar automáticamente la distancia de búsqueda. Sin embargo, su aplicación correcta exige mantener criterios metodológicos estables y documentados. Entre ellos, destaca el uso de la escala comunal, la aplicación de filtros de densidad local mediante la distancia mutua de alcance (mrd), la clasificación de consistencia, y la distinción entre ruido permanente y ruido satélite. Asimismo, el parámetro MinPts=20, consensuado con DIRPLAN y los equipos regionales, define la unidad mínima de aglomeración para la

inversión en infraestructura pública; no obstante, el modelo generará resultados distintos si este valor se modifica. Aunque HDBSCAN tiene resultados óptimos a nivel nacional, otros algoritmos basados en DBSCAN podrían complementar análisis locales específicos cuando las condiciones territoriales lo requieran.

### **c) Necesidad de fortalecer y ampliar el Modelo de Focalización**

El Modelo de Focalización se beneficia de la estructura espacial generada por el agrupamiento, pero su capacidad explicativa depende de incorporar información actualizada y más diversa. Existe un amplio margen para enriquecer sus variables con nuevos datos de equipamiento, accesibilidad, amenazas naturales, condiciones socioeconómicas y habitabilidad. Una focalización más sensible a la dinámica rural —y no solo a sus atributos estructurales— permitirá que el índice de equidad territorial evolucione y se alinee mejor con cambios demográficos, transformaciones del uso de suelo y variaciones en la provisión de servicios.

### **d) Integración del modelo hacia la cartera de proyectos y el ciclo de inversión**

Los resultados del estudio permiten avanzar más allá del diagnóstico territorial, sentando las bases para una cartera preliminar de proyectos. Los clústeres y su caracterización asociada pueden funcionar como unidades operativas que apoyen la programación plurianual y la prefactibilidad, especialmente cuando se integran con sectores como salud, educación, vialidad, saneamiento y agua potable rural. La integración del modelo al Plan Nacional de Infraestructura Pública permitiría conectar la priorización territorial con decisiones efectivas de inversión pública.

De este modo, el Modelo Territorial es una herramienta útil para focalizar sectores rurales que requieren mayor atención en la disponibilidad de infraestructura y servicios públicos.

### **e) Gobernanza técnica para asegurar la sostenibilidad del modelo**

La vigencia y continuidad del Modelo Territorial Rural dependen de establecer una gobernanza técnica clara al interior de DIRPLAN, responsable de resguardar la integridad metodológica y operativa de las herramientas desarrolladas. Para ello, resulta indispensable contar con un profesional o equipo encargado de mantener la toolbox, administrar y actualizar las capas maestras —especialmente la capa de viviendas rurales—, ajustar parámetros cuando corresponda, sistematizar y responder a la retroalimentación regional, y coordinar validaciones periódicas en terreno que permitan verificar la coherencia de los resultados frente a cambios reales en el poblamiento rural.

Asimismo, se recomienda ejecutar una actualización anual del modelo, utilizando la toolbox y los insumos más recientes, con el fin de monitorear cambios en la distribución

territorial de viviendas, mantener la trazabilidad entre versiones y asegurar consistencia en los análisis comparativos.

Esta gobernanza debe también garantizar la disponibilidad y usabilidad de la herramienta para los equipos regionales, asegurando que los analistas puedan ajustar parámetros según las particularidades de sus territorios. Esto incluye la posibilidad de modificar el valor de MinPts conforme a requerimientos regionales, adaptar la escala de análisis (comunal, provincial o subcomunal) y aplicar el modelo a contextos específicos donde se requiera mayor detalle. De este modo, el modelo conserva una estructura metodológica común a nivel nacional, pero permite la flexibilidad necesaria para responder a la diversidad territorial del país

#### **f) Valor metodológico y operacional del estudio**

El estudio no solo entrega dos modelos funcionales, sino también una arquitectura metodológica escalable, una toolbox replicable y un marco de equidad territorial que permite orientar la inversión pública con mayor precisión, transparencia y fundamento técnico. Su continuidad dependerá de mantener actualizados los datos, conservar coherencia metodológica, fortalecer la gobernanza institucional y avanzar hacia su integración en la cartera de proyectos del MOP. Bajo estas condiciones, el modelo tiene el potencial de consolidarse como una herramienta estratégica para transformar las condiciones de vida en los territorios rurales del país.

## 10. BIBLIOGRAFÍA

Anselin, L. (2020). Métodos de agrupamiento basados en la densidad. [https://geodacenter.github.io/workbook/99\\_density/lab9b.html](https://geodacenter.github.io/workbook/99_density/lab9b.html)

Banco Central Chile. (2022). Índice de Ventas Diarias del Comercio Minorista (IVDCM) Fuentes, métodos y resultados. División Estadísticas. [https://www.bcentral.cl/documents/33528/3662096/Nota\\_Tecnica\\_IVDCM.pdf](https://www.bcentral.cl/documents/33528/3662096/Nota_Tecnica_IVDCM.pdf)

Cancino, R., Bustos, F., Mora, M., Mendoza, Y., Herrera, M. & Acevedo, C. (2021). Diagnóstico transversal y estudio de línea de base Macro Zona Sur Austral. <https://ciep.maps.arcgis.com/sharing/rest/content/items/ae786d9b19d945eb80159f472d6b449a/data>

Campello, R.J.G.B., Moulavi, D., Sander, J. (2013). Density-Based Clustering Based on Hierarchical Density Estimates. In: Pei, J., Tseng, V.S., Cao, L., Motoda, H., Xu, G. (eds) Advances in Knowledge Discovery and Data Mining. PAKDD 2013. Lecture Notes in Computer Science, vol 7819. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-37456-2\\_14](https://doi.org/10.1007/978-3-642-37456-2_14)

Campello, R.J.G.B., Moulavi, D., Zimek, A. y Sandler, J. (2015). Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection. *ACM Transactions on Knowledge Discovery from Data*, 10(1): 1-51. <https://doi.org/10.1145/2733381>

Castillo, M., Núñez, A. & Galarza, S. (2024). Simulación de la operación de un bus eléctrico de 18 metros en rutas de Santiago de Chile. [https://theicct.org/wp-content/uploads/2024/08/ID-59-%E2%80%933-Simulacion-Santiago\\_final.pdf](https://theicct.org/wp-content/uploads/2024/08/ID-59-%E2%80%933-Simulacion-Santiago_final.pdf)

Ester, M., Kriegel, H.P., Sander, J. y Xu, X. (1996). Un algoritmo basado en densidad para descubrir clusters en grandes bases de datos espaciales con ruido. En *Actas de KDD-96*: 226–31.

ESRI. (2025a). Density-based Clustering (Spatial Statistics). <https://pro.arcgis.com/en/pro-app/3.4/tool-reference/spatial-statistics/densitybasedclustering.htm>

ESRI. (2025b). How Density-based Clustering works. <https://pro.arcgis.com/en/pro-app/3.4/tool-reference/spatial-statistics/how-density-based-clustering-works.htm>

García, F., Contreras, D., Vergara, F. & Andaur, X. (2024). Caracterización de sismos en el norte de Chile utilizando técnicas de clustering para la identificación de posibles puntos de primera respuesta. *Ingeniare. Revista chilena de ingeniería*, 32(8). <https://www.scielo.cl/pdf/ingeniare/v32/0718-3305-ingeniare-32-8.pdf>

Gan, J. & Tao, Y. (2017). On the Hardness and Approximation of Euclidean DBSCAN. *ACM Transactions on Database Systems (TODS)*, 42 (3). <http://dx.doi.org/10.1145/3083897>

Gómez, M., Barredo, J. I., & García, A. (2005). *Sistemas de Información Geográfica y evaluación multicriterio en la ordenación del territorio*, 2ª edición. RA-MA S.A. Editorial y Publicaciones. [http://www.amazon.es/Información-Geográfica-evaluación-multicriterio-ordenación/dp/8478976736/ref=sr\\_1\\_1/278-7474965-0339149?s=books&ie=UTF8&qid=1444701094&sr=1-1&keywords=Barredo+Cano%2C+José+Ignacio](http://www.amazon.es/Información-Geográfica-evaluación-multicriterio-ordenación/dp/8478976736/ref=sr_1_1/278-7474965-0339149?s=books&ie=UTF8&qid=1444701094&sr=1-1&keywords=Barredo+Cano%2C+José+Ignacio)

Illacanchi, P. (2021). DBSCAN: Explorando la densidad espacial de siniestros fatales para la identificación y priorización de intervenciones en la Red Vial Nacional. Ministerio de Transportes y Comunicaciones. <https://www.gob.pe/institucion/mtc/informes-publicaciones/4185441-documento-de-trabajo-001-dbscan-explorando-la-densidad-espacial-de-siniestros-fatales-para-la-identificacion-y-priorizacion-de-intervenciones-en-la-red-vial-nacional>

Instituto Nacional de Estadísticas (INE) (2025). Resultados Nacionales Censo 2024. <https://censo2024.ine.gob.cl/resultados/>

Instituto Nacional de Estadísticas (INE), Ministerio de Vivienda y Urbanismo (MINVU), Subsecretaría de Desarrollo Regional y Administrativo (SUBDERE) & Secretaría de Planificación de Transporte (SECTRA). (2020). Metodología para determinar las áreas funcionales de Chile. Gobierno de Chile. [https://www.ine.gob.cl/docs/default-source/geodatos-abiertos/publicaciones/indicadores-urbanos/metodolog%C3%ADa-para-determinar-las-%C3%A1reas-funcionales-de-chile.pdf?sfvrs-n=4bc6da9e\\_4](https://www.ine.gob.cl/docs/default-source/geodatos-abiertos/publicaciones/indicadores-urbanos/metodolog%C3%ADa-para-determinar-las-%C3%A1reas-funcionales-de-chile.pdf?sfvrs-n=4bc6da9e_4)

Jeon, Y. Yoo, J., Lee, J. & Yoon, S. (2017). NC-Link: A New Linkage Method for Efficient Hierarchical Clustering of Large-Scale Data. En *IEEE Access*, (5), pp. 5594-5608. <https://doi.org/10.1109/ACCESS.2017.2690987>

Li, T., Johansen, K. & McCabe, M.F. (2022). A machine learning approach for identifying and delineating agricultural fields and their multi-temporal dynamics using three decades of Landsat data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 186: 83-101. <https://doi.org/10.1016/j.isprsjprs.2022.02.002>

Liu, P., Zhou, D., & Wu, N. (2007). VDBSCAN: Varied Density Based Spatial Clustering of Applications with Noise. 2007 International Conference on Service Systems and Service Management, 1-4. <https://doi.org/10.1109/ICSSSM.2007.4280175>

McInnes, L. & Healy, J. (2017). Accelerated Hierarchical Density Based Clustering. *IEEE International Conference on Data Mining Workshops (ICDMW)*, New Orleans, LA, USA, 2017, pp. 33-42. <https://doi.org/10.1109/ICDMW.2017.12>

McInnes, L., Healy, J. y Astels, S. (2017). hdbscan: Hierarchical density based clustering, *Journal of Open Source Software*, 2(11), 205. <https://doi.org/10.21105/joss.00205>

Namaganda, J. & Mutale, J. (2020). An Optimal Rural Community PV Microgrid Design Using Mixed Integer Linear Programming and DBSCAN Approach. *SAIEE Africa Research Journal*, 111(39), 111-119. <http://dx.doi.org/10.23919/SAIEE.2020.9142604>

Nicolis, O., Delgado, L., Peralta, B. et al. (2024). Space-time clustering of seismic events in Chile using ST-DBSCAN-EV algorithm. *Environ Ecol Stat* 31, 509–536. <https://doi.org/10.1007/s10651-023-00594-3>

Ng, K., Chen, C. & Jenelius, E. (2025). The 15-minute city around one's trajectory: Evaluating food accessibility for transit users in Stockholm, Sweden. *Journal of Transport Geography*, 127. <https://doi.org/10.1016/j.jtrangeo.2025.104283>

Oficina de Estudios y Políticas Agrarias (ODEPA) (2022). Lineamientos para el uso de las definiciones de ruralidad en Chile – Informe Final, Ministerio de Agricultura. <https://bibliotecadigital.ciren.cl/items/c7b7ab39-3e98-4aac-90e2-796e8cc5892e>

Oficina de Estudios y Políticas Agrarias (ODEPA) (2025). La gradiente urbano-rural en Chile: Una propuesta de clasificación del territorio a escala sub-comunal con relevancia para el ciclo de políticas públicas en Chile. Ministerio de Agricultura de Chile.

Orellana, A., Díaz, D., Irrarázaval, F., Moreno, D., 2021. Descentralización e inversión pública en infraestructura: propuestas para priorizar brechas de desarrollo territorial a nivel subnacional. *Temas de la Agenda Pública*, 16(140), 1-10. Centro de Políticas Públicas UC. [https://politicaspUBLICAS.uc.cl/web/content/uploads/2021/05/Paper-No-140\\_web-3.pdf](https://politicaspUBLICAS.uc.cl/web/content/uploads/2021/05/Paper-No-140_web-3.pdf)

Rahimi, I., Mifei, L., Choon, J., Pamuspusan, D., Huang, Y., He, B., Cai, A., Reza, M. & Gandomi, A. (2025). Optimizing renewable energy site selection in rural Australia: Clustering algorithms and energy potential analysis. *Energy Conversion and Management: X*, 25. <https://doi.org/10.1016/j.ecmx.2024.100855>

Ramírez, M. (2022). VISTRAFFIC: Herramienta para análisis del tráfico vehicular en Santiago. Tesis de grado para optar al título de Ingeniero Civil en Computación. Universidad de Chile. <https://repositorio.uchile.cl/handle/2250/189200>

Reda, H., Mohapatra, S. K., Das, T. K., & Dash, S. K. (2024). Electric bus arrival and charging station placement assessment using machine learning techniques. *International Journal of Sustainable Engineering*, 17(1), 278–294. <https://doi.org/10.1080/19397038.2024.2333563>

Said, S., Aminuddin, R., Zainal-Abidin, N., Nasir, S. & Ibrahim, A. (2022). Visualizing COVID-19 Vaccination Rate and Vaccination Centre in Malaysia using DBSCAN Clustering model. 2022 IEEE International Power and Renewable Energy Conference (IPRECON), Kollam, India. <https://doi.org/10.1109/IPRECON55716.2022.10059495>

Sander, J., Ester, M., Kriegel, H.P. y Xu, X. (1998). Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications. *Minería de datos y descubrimiento de conocimiento*, 2: 169–94. <https://doi.org/10.1023/A:1009745219419>

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Trans. Database Syst.*, 42(3), 19:1-19:21. <https://doi.org/10.1145/3068335>

Subsecretaría de Desarrollo Regional y Administrativo (SUBDERE). (2020). Asentamientos humanos rurales en Chile, una clasificación comunal: Una aproximación desde el análisis espacial de la concentración y dispersión de la población. Gobierno de Chile. <https://proactiva.subdere.gov.cl/handle/123456789/514>

Voogd, H. (1982). Multicriteria evaluation with mixed qualitative and quantitative data. *Environment and Planning B: Planning and Design*, 9(2), 221-236. <https://doi.org/10.1068/b090221>

Wang, X. & Ma, Q. (2024). Spatial Distribution Pattern and Influencing Factors of Bed-and-breakfasts (B&Bs) from the Perspective of Urban-rural Differences: A Case Study of Jiadong Peninsula, China. *Chin. Geogr. Sci.* 34, 752–763. <https://doi.org/10.1007/s11769-024-1448-3>

Zhou S., Deng Y., Wang R., Li, N. & Si, Q. (2017). "Effective Mapping of Urban Areas Using ENVISAT ASAR, Sentinel-1A, and HJ-1-C Data. *IEEE Geoscience and Remote Sensing Letters*, 14(6): 891-895. <https://doi.org/10.1109/LGRS.2017.2686419>

## 11. ANEXOS

### 11.1. Anexo 1. Script de código R desarrollado para la homologación y estandarización de datos geospaciales.

```
# Meta -----
# Title: Estandarización de datos base a normas MOP - APC 2023
# Project: Modelo territorial de priorización MOP
# Author: OCUC
# Creation date: 16-06-2025
# Notes -----
# Este script tiene como objetivo ilustrar el flujo de trabajo de homologación de datos usando la
# actualización pre-censal del 2023. La ejecución resulta en una serie de shapefiles adaptados a la
# normativa de la IDE MOP.
# Este script no representa una versión final del trabajo completo, sino que un ejemplo
# ilustrativo con un dataset específico; cada dataset contemplado en la Geodatabase final cuenta o
# contará con una lógica propia dentro de este script.
# Libs -----
# Lectura de paquetes utilizados en el script. El uso de xfun::pkg_attach permite instalar los
# paquetes que no se encuentren instalados en el sistema, y cargar los que ya estén instalados.
install.packages('xfun')
suppressMessages({
xfun::pkg_attach( c('roperators', 'magrittr', 'arrow', 'cli', 'dplyr', 'sf', 'stringr',
'janitor'), install = T, message = F )
})
# Workspace -----
# Set working directory
setwd('/nexus/work/ocuc25/mop')
# Remove scientific notation
options(scipen = 999)
```

```
# Functions -----
# SECCIÓN A
# Se definen funciones auxiliares para facilitar la inferencia de un CUT a partir de otro.
region_desde_comuna <- function(x) str_sub(str_pad(x, 5, 'left', '0'), 1, 2)
provincia_desde_comuna <- function(x) str_sub(str_pad(x, 5, 'left', '0'), 1, 3)
region_desde_provincia <- function(x) str_sub(str_pad(x, 3, 'left', '0'), 1, 2)
comuna_desde_zona <- function(x) str_pad(x, 11, 'left', '0')
zona_desde_manzent <- function(x) str_pad(x, 14, 'left', '0')
# Se define una función para facilitar la homologación de un string de texto al estándar
# planteado por IDE MOP.
idemop_str <- function(x) {
```

```
x |> str_replace_all(' ', '_') |> str_to_upper() |> iconv(from = 'UTF-8', to = 'ASCII//TRANSLIT')
}

# Se definen funciones para homologar los datos y los nombres de variables, así como el sistema
de coordenadas. Una para cada nivel administrativo de la APC 2023.

apc23_aldea <- function(sf_aldea) {

std_apc23_aldea <- sf_aldea |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO = provincia_desde_comuna(CUT),
NOM_PRO = idemop_str(N_PROVINCI), CUT_COM = str_pad(CUT, 5, 'left', '0'), NOM_COM =
idemop_str(N_COMUNA), NOMBRE = idemop_str(N_ALDEA), # NOM_ASCII = iconv(NOMBRE_UTF8, from = 'UTF-
8', to = 'ASCII//TRANSLIT'), # NOM_ORI = N_ALDEA, FECHA_ACT = NA_character_, LEVANT =
NA_character_ ) |> st_transform(crs_epsg)

std_apc23_aldea$ID_IDE <- sapply(1:nrow(std_apc23_aldea), function(x) gibberish(15, sym = F))

return(std_apc23_aldea)

}

apc23_comunal <- function(sf_comunal) {

std_apc23_comunal <- sf_comunal |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO = provincia_desde_comuna(CUT),
NOM_PRO = idemop_str(N_PROVINCI), CUT_COM = str_pad(CUT, 5, 'left', '0'), NOM_COM =
idemop_str(N_COMUNA), # NOM_ASCII = iconv(NOM_UTF8, from = 'UTF-8', to = 'ASCII//TRANSLIT'), #
NOM_ORI = N_COMUNA, FECHA_ACT = NA_character_, LEVANT = NA_character_ ) |> st_transform(crs_epsg)

std_apc23_comunal$ID_IDE <- sapply(1:nrow(std_apc23_comunal), function(x) gibberish(15, sym = F))

return(std_apc23_comunal)

}

apc23_distrital <- function(sf_distrital) {

std_apc23_distrital <- sf_distrital |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO = provincia_desde_comuna(CUT),
NOM_PRO = idemop_str(N_PROVINCI), CUT_COM = str_pad(CUT, 5, 'left', '0'), NOM_COM =
idemop_str(N_COMUNA), COD_DIS = COD_DISTRI, CUT_DIS = paste0(CUT_COM, str_pad(COD_DISTRI, 2,
'left', '0')), NOMBRE = idemop_str(N_DISTRITO), # NOMBRE_ASCII = iconv(NOMBRE_UTF8, from = 'UTF-
8', to = 'ASCII//TRANSLIT'), # NOMBRE_ORIGINAL = N_DISTRITO, TIPO = idemop_str(TIPO_DISTR),
FECHA_ACT = NA_character_, LEVANT = NA_character_ )

std_apc23_distrital$ID_IDE <- sapply(1:nrow(std_apc23_distrital), function(x) gibberish(15, sym =
F))

return(std_apc23_distrital)

}

apc23_eje_vial <- function(sf_eje_vial) {

std_apc23_eje_vial <- sf_eje_vial |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO = provincia_desde_comuna(CUT),
NOM_PRO = idemop_str(N_PROVINCI), CUT_COM = str_pad(CUT, 5, 'left', '0'), NOM_COM =
idemop_str(N_COMUNA), CLASE_COM = CLASE_COMU, CLASE_URB = CLASE_URBA, NOMBRE =
idemop_str(NOMBRE), # NOMBRE_ASCII = iconv(NOMBRE_UTF8, from = 'UTF-8', to = 'ASCII//TRANSLIT'),
# NOMBRE_ORIGINAL = N_EJE_VIAL ALIAS = idemop_str(ALIAS), ROLE = idemop_str(ROL), FECHA_ACT =
NA_character_, LEVANT = NA_character_ ) |> st_transform(crs_epsg)

std_apc23_eje_vial$ID_IDE <- sapply(1:nrow(std_apc23_eje_vial), function(x) gibberish(15, sym =
F))

return(std_apc23_eje_vial)

}

}
```

```

apc23_limite_urbano <- function(sf_limite_urbano) {

std_apc23_limite_urbano <- sf_limite_urbano |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO = provincia_desde_comuna(CUT),
NOM_PRO = idemop_str(N_PROVINCI), CUT_COM = str_pad(CUT, 5, 'left', '0'), NOM_COM =
idemop_str(N_COMUNA), NOMBRE = idemop_str(N_URBANO), TIPO = idemop_str(TIPO), CATEGORIA =
idemop_str(CATEGORIA), FECHA_ACT = NA_character_, LEVANT = NA_character_ ) |>
st_transform(crs_epsg)

std_apc23_limite_urbano$ID_IDE <- sapply(1:nrow(std_apc23_limite_urbano), function(x)
gibberish(15, sym = F))

return(std_apc23_limite_urbano)

}

apc23_manzanas <- function(sf_manzana_aldea, sf_manzana_urbana) {

std_apc23_manzana_aldea <- sf_manzana_aldea |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO = provincia_desde_comuna(CUT),
NOM_PRO = idemop_str(N_PROVINCI), CUT_COM = str_pad(CUT, 5, 'left', '0'), NOM_COM =
idemop_str(N_COMUNA), COD_DIS = COD_DISTRI, COD_ZONLOC = COD_LOCALI, COD_ENT = COD_ENTIDA,
COD_MANZ = COD_MANZAN, NOM_POB = idemop_str(N_ALDEA), # NOM_ASCII = iconv(NOMBRE_UTF8, from =
'UTF-8', to = 'ASCII//TRANSLIT'), # NOM_ORI = N_ALDEA CUT_MANZ = str_pad(MANZENT_TX, 11, 'left',
'0'), TIPO = 'RURAL', FECHA_ACT = NA_character_, LEVANT = NA_character_ ) |>
st_transform(crs_epsg)

std_apc23_manzana_urbana <- sf_manzana_urbana |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO = provincia_desde_comuna(CUT),
NOM_PRO = idemop_str(N_PROVINCI), CUT_COM = str_pad(CUT, 5, 'left', '0'), NOM_COM =
idemop_str(N_COMUNA), COD_DIS = COD_DISTRI, COD_ZONLOC = COD_ZONA, COD_ENT = NA_integer_,
COD_MANZ = COD_MANZAN, NOM_POB = idemop_str(N_URBANO), CUT_MANZ = str_pad(Mzent_TX, 11, 'left',
'0'), TIPO = 'URBANA', FECHA_ACT = NA_character_, LEVANT = NA_character_ ) |>
st_transform(crs_epsg)

std_apc23_manzana <- rbind(std_apc23_manzana_aldea, std_apc23_manzana_urbana) |>
arrange(CUT_MANZ)

std_apc23_manzana$ID_IDE <- sapply(1:nrow(std_apc23_manzana), function(x) gibberish(15, sym = F))

return(std_apc23_manzana_aldea)

}

apc23_edificaciones_rurales <- function(sf_edificaciones_rurales, sf_distritos) {

std_apc23_edificaciones_rurales <- sf_edificaciones_rurales |> transmute( ID_IDE = NA_character_,
CUT_REG = region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO =
provincia_desde_comuna(CUT), NOM_PRO = idemop_str(N_PROVINCI), CUT_COM = str_pad(CUT, 5, 'left',
'0'), NOM_COM = idemop_str(N_COMUNA), FUENTE = if_else( FUENTE == 'APC 2023', 'APC_2023',
idemop_str(FUENTE) ), USO = idemop_str(USO_EDIFIC), FECHA_ACT = NA_character_, LEVANT =
NA_character_ )

std_apc23_edificaciones_rurales$ID_IDE <- sapply(1:nrow(std_apc23_edificaciones_rurales),
function(x) gibberish(15, sym = F))

return(std_apc23_edificaciones_rurales)

}

idemop_export <- function(x) {

x_df <- get(x, envir = .GlobalEnv) x_stringy <- str_sub(x, 5)

x_TG <- x_df |> select(ID_IDE, geometry)

if (st_crs(x_TG)$epsg != crs_epsg) x_TG <- st_transform(x_TG, crs_epsg)

x_TL <- x_df |> st_drop_geometry()

```

```
x_name <- file_names$file_name[file_names$df_name == x_stringy]

x_TG_path <- str_glue('estandarizacion/export/TG/{x_name}') x_TL_path <-
str_glue('estandarizacion/export/TL') x_zip_path <-
str_glue('estandarizacion/export/ZIP/{x_name}.zip')

if (!dir.exists(x_TG_path)) dir.create(x_TG_path)

if (!dir.exists('estandarizacion/export/TL'))
dir.create(str_glue('estandarizacion/export/TL/{x_name}'), recursive = T)

x_TG |> st_write(str_glue('{x_TG_path}/{x_name}.shp'), append = F) x_TL |>
write.csv(str_glue('{x_TL_path}/{x_name}.csv'), row.names = F, append = F)

zip0( zipfile = x_zip_path, files = c( list.files( path =
str_glue('estandarizacion/export/TG/{x_name}'), full.names = T ),
str_glue('estandarizacion/export/TL/{x_name}.csv') ) )
}

# Global vars -----
# Constantes de uso regular dentro del script.
# ----- Data paths

path_limite_comunal <- '/nexus/work/ocuc25/mop/estandarizacion/data/LÍMITE COMUNAL' path_cuencas
<- '/nexus/work/ocuc25/mop/estandarizacion/data/Cuencas_BNA' path_red_vial <-
'/nexus/work/ocuc25/mop/estandarizacion/data/Red_Vial_2021_MOP' path_apc23 <-
'/nexus/work/ocuc25/mop/data_avances_23_06'

# ----- Parameters

crs_epsg <- 9153
```

```
# Data -----
# Sección B

# ----- Reference

# Lectura de datos de referencia.

comunas <- read_parquet('/nexus/work/db_temp/datos_ine17/comunas.parquet') provincias <-
read_parquet('/nexus/work/db_temp/datos_ine17/provincias.parquet') regiones <-
read_parquet('/nexus/work/db_temp/datos_ine17/regiones.parquet')

file_names <- read.csv('/nexus/work/ocuc25/mop/estandarizacion/file_names.csv')

# ----- Cuencas

shp_cuencas <- st_read(str_glue('{path_cuencas}/{findShp(path_cuencas)}'))

# ----- Red vial

shp_red_vial <- st_read(str_glue('{path_red_vial}/{findShp(path_red_vial)}'))

# ----- ISMT 2022

shp_isit22 <- st_read('/nexus/work/ocuc25/mop/sharepoint/4. DATOS
GEOESPACIALES/INSUMOS_BASE/SOCIOECONOMICO/%C3%8Dndice_Socio_Material_y_Territorial_(ACTUALIZACI%C
3%93N_METODOL%C3%93GICA_2022_-_DATOS_CENSO_2017).shp')

# ----- Acuíferos SHAC

shp_acuíferos_shac <-
st_read('estandarizacion/data/Acuíferos_SHAC_2024_DGA/Acuíferos_SHAC/INV_Acuíferos_SHAC.shp')

# ----- Acuíferos y vegas protegidas
```

```
shp_acuiferos_vegas <-
st_read('estandarizacion/data/Acuiferos_y_Vegas_Protegidas_2024_DGA/Acuiferos_y_Vegas_Protegidas/
Acuiferos_Protegidos_Vegas_y_Bofedales.shp')

shp_vegas_protegidas <-
st_read('estandarizacion/data/Acuiferos_y_Vegas_Protegidas_2024_DGA/Acuiferos_y_Vegas_Protegidas/
Vegas_Protegidas.shp')

# ----- Aeródromos
shp_aerodromos <- st_read('estandarizacion/data/AERODROMOS/aerodromos_chile_2014.shp')

# ----- Bomberos
shp_bomberos <-
st_read('estandarizacion/data/CIAS_BOMBEROS/layer_companias_de_bomberos_20231110080349.shp')

# ----- Carabineros
shp_carabineros <-
st_read('estandarizacion/data/COMISARIAS_CARABINEROS/layer_cuarteles_20220309024623.shp')

# ----- Establecimientos Educación Escolar
shp_establecimientos_educ <-
st_read('estandarizacion/data/ESTABLECIMIENTOS_EDUCACION_ESCOLAR_2021/layer_establecimientos_educ
acion_escolar_20220309024120.shp')

# ----- Establecimientos salud
shp_establecimientos_salud <-
st_read('estandarizacion/data/ESTABLECIMIENTOS_SALUD_2024/layer_establecimientos_de_salud_2024071
2091725.shp')

# ----- Red hidrográfica
shp_red_hidro <-
st_read('estandarizacion/data/Red_Hidrografica_Nacional_2017_BCN/Red_Hidrografica.shp')

# ----- Red vial INE
shp_red_vial_ine <- st_read('estandarizacion/data/RED_VIAL_INE_2024/RED_VIAL_INE_2024.shp')

# ----- Servicios sanitarios rurales
shp_servicios_sanitarios_rurales <-
st_read('estandarizacion/data/Servicios_Sanitarios_Rurales_2025/Servicios_Sanitarios_Rurales_2025
_04_29.shp')

# ----- Territorio operacional
shp_territorio_operacional <-
st_read('estandarizacion/data/TERRITORIO_OPERACIONAL2024/TO_2024_12.shp')
```

```
# Main -----
# SECCIÓN C
# ----- Estandarización de cuencas
shp_cuencas$REGION_AD <- NA_character_
for (i in 1:nrow(shp_cuencas)) {
cuenca <- shp_cuencas[i, ]
region_primaria <- cuenca$COD_REG region_secundaria <- cuenca$COD_REGION
if (region_primaria == region_secundaria) {
```

next

```

} else if (length(str_split(region_secundaria, ',')[[1]]) > 1) {
regiones_adicionales <- str_split(region_secundaria, ',')[[1]] |>
  str_subset(paste0('^', region_primaria, '$'), negate = T)

if (length(regiones_adicionales) > 1) stringy <- str_c(regiones_adicionales, collapse = ',')
else stringy <- regiones_adicionales

}

shp_cuencas$CUT_REGION_AD[i] <- stringy
}

std_cuencas <- shp_cuencas |> select( ID_IDE, COD_CUENCA = COD_CUEN, NOMBRE = NOM_CUEN, REGION =
COD_REG, REGION_AD, FUENTE, FECHA = FECHA_ACTU, LEVANTAMIE, geometry ) |> st_transform(crs_epsg)

st_write(std_cuencas, '/nexus/work/ocuc25/mop/estandarizacion/new_data/CUENCAS.shp')

----- Estandarización de red vial

std_red_vial <- shp_red_vial |> transmute( ID_IDE = NA_character_, CUT_REGION = case_when( REGION
== 'Región de Tarapacá' ~ '01', REGION == 'Región de Antofagasta' ~ '02', REGION == 'Región de
Atacama' ~ '03', REGION == 'Región de Coquimbo' ~ '04', REGION == 'Región de Valparaíso' ~ '05',
REGION == "Región del Libertador General Bernardo O'Higgins" ~ '06', REGION == 'Región del Maule'
~ '07', REGION == 'Región del Biobío' ~ '08', REGION == 'Región de La Araucanía' ~ '09', REGION
== 'Región de Los Lagos' ~ '10', REGION == 'Región Aysén del General Carlos Ibáñez del Cam' ~
'11', REGION == 'Región de Magallanes y La Antártica Chilena' ~ '12', REGION == 'Región
Metropolitana de Santiago' ~ '13', REGION == 'Región de Los Ríos' ~ '14', REGION == 'Región de
Arica y Parinacota' ~ '15', REGION == 'Región de Ñuble' ~ '16', T ~ NA_character_ ), NOM_REG =
NA_character_, ID_CAMINO = CODIGO_CAM, ROL = ROL, ROL_ID = ROL_ID, ROL_LABEL = ROL_LABEL, NOMBRE
= NOMBRE_CAM, CARPETA = CARPETA, CLASIFICAC = CLASIFICAC, CALZADA = CALZADA, ORIENTACION =
ORIENTACIO, KM_INICIO = KM_I, KM_FIN = KM_F, KM_TRAMO = KM_TRAMO, ENROLADO = if_else(ENROLADO ==
'Sí', T, F), CONCESIONA = if_else(CONCESIONA == 'Sí', T, F), geometry = geometry ) |>
st_transform(crs_epsg) # |>

st_make_valid()

std_red_vial$ID_IDE <- sapply(1:nrow(std_red_vial), function(x) gibberish(15, sym = F))
std_red_vial$NOM_REG <- regiones$nombre[match(std_red_vial$CUT_REGION, str_pad(regiones$ine17, 2,
'left', '0'))]

st_write(std_red_vial, '/nexus/work/ocuc25/mop/estandarizacion/new_data/RED_VIAL.shp')

# ----- Estandarización de actualización pre-censal 2023

for (i in list.files('data_avances_23_06', recursive = F, full.names = T)) {

region <- str_extract(i, 'R\d{2}')

# Lectura de datos e inmediata homologación mediante la aplicación de las funciones previamente
deifinidas.

region_std_aldea <- apc23_aldea(st_read(str_glue('{i}/Aldea.shp'))) region_std_comunal <-
apc23_comunal(st_read(str_glue('{i}/Comunal.shp'))) region_std_distrital <-
apc23_distrital(st_read(str_glue('{i}/Distrital.shp'))) region_std_eje_vial <-
apc23_eje_vial(st_read(str_glue('{i}/Eje_Vial.shp'))) region_std_limite_urbano <-
apc23_limite_urbano(st_read(str_glue('{i}/Limite_Urbano_Censal.shp'))) region_std_manzanas <-
apc23_manzanas(st_read(str_glue('{i}/Manzana_Aldea.shp')),
st_read(str_glue('{i}/Manzana_Urbana.shp'))) region_std_edificaciones_rurales <-
apc23_edificaciones_rurales(st_read(str_glue('{i}/Puntos_Edificacion_Rural.shp')))

# Si la región es R01, se inicializan las variables de almacenamiento. Si no, se concatenan los
datos a las variables ya existentes.

```

```

if (region == 'R01') {

std_aldea <- region_std_aldea
std_comunal <- region_std_comunal
std_distrital <- region_std_distrital
std_eje_vial <- region_std_eje_vial
std_limite_urbano <- region_std_limite_urbano
std_manzanas <- region_std_manzanas
std_edificaciones_rurales <- region_std_edificaciones_rurales

} else {

std_aldea <- rbind(std_aldea, region_std_aldea)
std_comunal <- rbind(std_comunal, region_std_comunal)
std_distrital <- rbind(std_distrital, region_std_distrital)
std_eje_vial <- rbind(std_eje_vial, region_std_eje_vial)
std_limite_urbano <- rbind(std_limite_urbano, region_std_limite_urbano)
std_manzanas <- rbind(std_manzanas, region_std_manzanas)
std_edificaciones_rurales <- rbind(std_edificaciones_rurales, region_std_edificaciones_rurales)

}

cli_alert(rmsg(str_glue('Processed region {region}')))

}

# ----- Estandarización de ISMT 2022

std_ismt22 <- shp_isit22 |> transmute( ID_IDE = NA_character_, CUT_REG = str_pad(cod_region, 2,
'left', '0'), NOM_REG = idemop_str(str_sub(region, 4, -1)), CUT_PRO = str_pad(cod_provin, 3,
'left', '0'), NOM_PRO = idemop_str(provincia), CUT_COM = str_pad(cod_comuna, 5, 'left', '0'),
NOM_COM = idemop_str(comuna), CUT_ZONA = str_pad(zona, 11, 'left', '0'), NOM_ZONA =
idemop_str(nombre), ISMTPN = ismtpn, Q1 = Q1, Q2 = Q2, Q3 = Q3, Q4 = Q4, Q5 = Q5, ALTO = Alto,
MEDIO = Medio, BAJO = Bajo, AB = AB, C1 = C1, C2 = C2, C3 = C3, D = D, E = E, IND_HACIN =
ind_hac, HACIN_CRI = hacin_cri, HACIN_MED = hacin_med, HACIN_NO = hacin_no, ALLEG = alleg,
ESCOLAR = escolar, MAT_ACCEPT = mat_acept, MAT_RECUP = mat_recup, MAT_IRREC = mat_irrec, IND_MATER
= ind_mat, FECHA_ACT = NA_character_, LEVANT = NA_character_, geometry = geometry ) |>
st_transform(crs_epsg)

std_ismt22$ID_IDE <- sapply(1:nrow(std_ismt22), function(x) gibberish(15, sym = F))

# ----- Estandarización de Límites Regionales

for (i in list.files('estandarizacion/data/LIMITE_REGIONAL', full.names = T)) {

region <- str_sub(i, -3, -1)

region_limite <- st_read(str_glue('{i}/REGION_C17.shp')) |> transmute( ID_IDE = NA_character_,
CUT_REG = str_pad(REGION, 2, 'left', '0'), NOM_REG = idemop_str(NOM_REGION), geometry = geometry
) |> st_transform(crs_epsg)

region_limite$ID_IDE <- sapply(1:nrow(region_limite), function(x) gibberish(15, sym = F))

if (region == 'R01') std_limite_regional <- region_limite if (region != 'R01')
std_limite_regional <- rbind(std_limite_regional, region_limite)

}

# ----- Estandarización de Acuíferos SHAC

std_acuíferos_shac <- shp_acuíferos_shac |> unselect('OBJECTID') |> transmute( ID_IDE = ID_IDE,
ID_UNICO = ID_UNICO, COD_SHAC = COD_SHAC, CUT_REG = str_pad(COD_REG, 2, 'left', '0'), NOM_REG =
idemop_str(REGION), COD_BNA = COD_BNA, NOM_BNA = idemop_str(NOM_BNA), NOM_ACUIF =
idemop_str(NOM_ACUIF), COD_BNA_SH = COD_BNA_SH, SHAC = idemop_str(SHAC), TIPO_LIMIT =
idemop_str(TIPO_LIMIT), RES_DGA = RES_DGA, F_RES_DGA = F_RES_DGA, F_TRAZON = F_TRAZON, F_PUB_DO =
F_PUB_DO, RESMODIF_1 = RESMODIF_1, PUBMODIF_1 = PUBMODIF_1, MOTIV_MOD1 = MOTIV_MOD1, RESMODIF_2 =
RESMODIF_2, PUBMODIF_2 = PUBMODIF_2, MOTIV_MOD2 = MOTIV_MOD2, RESMODIF_3 = RESMODIF_3, PUBMODIF_3
= PUBMODIF_3, MOTIV_MOD3 = MOTIV_MOD3, TIPOESTUDI = TIPOESTUDI, REFERENCIA = REFERENCIA,

```

```
F_UESTUDIO = F_UESTUDIO, FUENTE = FUENTE, FECHA_ACTU = FECHA_ACTU, LEVANTAMIE = LEVANTAMIE,
geometry = geometry ) |> st_transform(crs_epsg)

# ----- Estandarización de Acuíferos y Vegas Protegidas

std_vegas_protegidas <- shp_vegas_protegidas |> transmute( ID_IDE = ID_IDE, CUT_REG =
str_pad(COD_REG, 2, 'left', '0'), NOM_REG = idemop_str(REGION), CUT_PRO = str_pad(COD_PROV, 3,
'left', '0'), NOM_PRO = idemop_str(PROVINCIA), CUT_COM = str_pad(COD_COM, 5, 'left', '0'),
NOM_COM = idemop_str(COMUNA), COD_CUEN = COD_CUEN, COD_SUBC = COD_SUBC, COD_SSUBC = COD_SSUBC,
COD_VEG = COD_VEG, NUM_VEG = NUM_VEGA, NOM_VEG = idemop_str(NOM_VEG), EXISTENCIA =
idemop_str(EXISTENCIA), ORIGEN = idemop_str(ORIGEN), POSICION = idemop_str(POSICION), MAPA =
idemop_str(MAPA), UTM_ESTE = UTM_ESTE, UTM_NORTE = UTM_NORTE, FUENTE = FUENTE, FECHA_ACTU =
FECHA_ACTU, LEVANTAMIE = LEVANTAMIE, geometry = geometry ) |> st_transform(crs_epsg)

# ----- Estandarización de Aeródromos

std_aerodromos <- shp_aerodromos |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(cod_comuna), NOM_REG = idemop_str(Region), CUT_PRO =
provincia_desde_comuna(cod_comuna), NOM_PRO = idemop_str(Provincia), CUT_COM =
str_pad(cod_comuna, 5, 'left', '0'), NOM_COM = idemop_str(Comuna), NOMBRE =
idemop_str(Aerodromo), CATEGORIA = idemop_str(categoría), COD_OACI = cod_oaci, COD_IATA =
cod_iata, FUENTE = NA_character_, FECHA_ACT = NA_character_, LEVANTAMIE = NA_character_, geometry
= geometry ) |> st_transform(crs_epsg)

std_aerodromos$ID_IDE <- sapply(1:nrow(std_aerodromos), function(x) gibberish(15, sym = F))

# ----- Estandarización de Bomberos

std_bomberos <- shp_bomberos |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(ID_2002), NOM_REG = NA_character_, CUT_PRO = provincia_desde_comuna(ID_2002),
NOM_PRO = NA_character_, CUT_COM = str_pad(ID_2002, 5, 'left', '0'), NOM_COM = NA_character_,
CUT_CUERPO = str_pad(CUT_CUERPO, 5, 'left', '0'), CUT_CIA = CUT_CIA, COMPANIA =
idemop_str(COMPANIA), DIRECCION = idemop_str(DIRECCION), TELEFONO = TELEFONO, FUENTE =
NA_character_, FECHA_ACT = NA_character_, LEVANTAMIE = NA_character_, geometry = geometry ) |>
st_transform(crs_epsg)

std_bomberos$NOM_REG <- idemop_str(regiones$nombre_completo[match(std_bomberos$CUT_REG,
str_pad(regiones$region_16r, 2, 'left', '0'))]) std_bomberos$NOM_PRO <-
idemop_str(provincias$nombre[match(std_bomberos$CUT_PRO, str_pad(provincias$ine17, 3, 'left',
'0'))]) std_bomberos$NOM_COM <- idemop_str(comunas$nombre[match(std_bomberos$CUT_COM,
str_pad(comunas$ine17, 5, 'left', '0'))]) std_bomberos$ID_IDE <- sapply(1:nrow(std_bomberos),
function(x) gibberish(15, sym = F))

# ----- Estandarización de Carabineros

std_carabineros <- shp_carabineros |> transmute( ID_IDE = NA_character_, CUT_REG = str_pad(RG, 2,
'left', '0'), NOM_REG = idemop_str(NOMBRE_REG), NOM_PRO = idemop_str(NOMBRE_PRO), NOM_COM =
idemop_str(NOMBRE_COM), PREFECTURA = idemop_str(PREFECTURA), ZONA = idemop_str(ZONA), NOM_UNIDAD
= idemop_str(NOMBRE_UNI), TIPO_UNIDAD = idemop_str(TIPO_DE_UN), TIPO_VIA =
idemop_str(TIPO_DE_VI), NOM_VIA = idemop_str(NOMBRE_DE), NUMERO = NUMERO, HECHO = HECHO,
UNID_ENCRI = UNID_ENCRI, POINT_X = POINT_X, POINT_Y = POINT_Y, geometry = geometry ) |>
st_transform(crs_epsg) |> st_zm(drop = T, what = 'ZM')

std_carabineros$ID_IDE <- sapply(1:nrow(std_carabineros), function(x) gibberish(15, sym = F))

# ----- Estandarización de Establecimientos Educación Escolar

std_establecimientos_educ <- shp_establecimientos_educ |> transmute( ID_IDE = NA_character_,
CUT_REG = str_pad(COD_REG_RB, 2, 'left', '0'), NOM_REG = NA_character_, CUT_PRO =
str_pad(COD_PRO_RB, 3, 'left', '0'), NOM_PRO = NA_character_, CUT_COM = str_pad(COD_COM_RB, 5,
'left', '0'), NOM_COM = NA_character_, NOMBRE = idemop_str(NOM_RBD), DIRECCION =
idemop_str(DIRECCION), NUMERO = NUMERO, REFERENCIA = idemop_str(REFERENCIA), RBD = RBD, DGV_RBD =
DGV_RBD, ANNO = AGNO, TIPO_SOST = TIPO_SOST, MAT_PARV = MAT_PARV, MAT_BAS_RE = MAT_BAS_RE,
MAT_BAS_AD = MAT_BAS_AD, MAT_ESP = MAT_ESP, MAT_MHC_RE = MAT_MHC_RE, MAT_MHC_AD = MAT_MHC_AD,
MAT_MTP_RE = MAT_MTP_RE, MAT_MTP_AD = MAT_MTP_AD, MAT_TOTAL = MAT_TOTAL, MAT_HOM_TO = MAT_HOM_TO,
MAT_MUJ_TO = MAT_MUJ_TO, MAT_SI_TOT = MAT_SI_TOT, CUR_SIM_TO = CUR_SIM_TO, CUR_COMB_T =
CUR_COMB_T, FUENTE = NA_character_, FECHA_ACT = NA_character_, LEVANTAMIE = NA_character_,
LATITUD = LATITUD, LONGITUD = LONGITUD, geometry = geometry
) |> st_transform(crs_epsg)
```

```

std_establecimientos_educ$NOM_REG <-
idemop_str(regiones$nombre_completo[match(std_establecimientos_educ$CUT_REG,
str_pad(regiones$region_16r, 2, 'left', '0'))]) std_establecimientos_educ$NOM_PRO <-
idemop_str(provincias$nombre[match(std_establecimientos_educ$CUT_PRO, str_pad(provincias$ine17,
3, 'left', '0'))]) std_establecimientos_educ$NOM_COM <-
idemop_str(comunas$nombre[match(std_establecimientos_educ$CUT_COM, str_pad(comunas$ine17, 5,
'left', '0'))]) std_establecimientos_educ$ID_IDE <- sapply(1:nrow(std_establecimientos_educ),
function(x) gibberish(15, sym = F))

# ----- Estandarización de Establecimientos Salud

std_establecimientos_salud <- shp_establecimientos_salud |> transmute( ID_IDE = NA_character_,
CUT_REG = str_pad(COD_REG, 2, 'left', '0'), NOM_REG = idemop_str(NOM_REG), CUT_PRO =
provincia_desde_comuna(COD_COM), NOM_PRO = idemop_str(PROVINCIA), CUT_COM = str_pad(COD_COM, 5,
'left', '0'), NOM_COM = idemop_str(NOM_COM), VIA = idemop_str(VIA), DIRECCION =
idemop_str(DIRECCION), NUMERO = NUMERO, TELEFONO = FONO, COD_ANT = COD_ANT, COD_VIG = COD_VIG,
COD_M_ANT = COD_M_ANT, COD_M_NUEV = COD_M_NUEV, NOMBRE = idemop_str(NOMBRE), TIPO =
idemop_str(TIPO), AMBITO = idemop_str(AMBITO), PERTENENCI = idemop_str(PERTENENCI), DEPENDENC =
idemop_str(DEPENDENC), CERTIFI = idemop_str(CERTIFI), DEP_ADM = idemop_str(DEP_ADM), NIVEL =
idemop_str(NIVEL), F_INICIO = F_INICIO, URGENCIA = idemop_str(URGENCIA), TIPO_URGE =
idemop_str(TIPO_URGE), CLAS_SAPU = idemop_str(CLAS_SAPU), PRESTADOR = idemop_str(PRESTADOR),
ESTADO = idemop_str(ESTADO), COMPLEJIDA = idemop_str(COMPLEJIDA), TIPO_ATEN =
idemop_str(TIPO_ATEN), F_INCR = F_INCR, FUENTE = NA_character_, FECHA_ACT = NA_character_,
LEVANTAMIE = NA_character_, LATITUD = LATITUD, LONGITUD = LONGITUD, geometry = geometry ) |>
st_transform(crs_epsg)

std_establecimientos_salud$ID_IDE <- sapply(1:nrow(std_establecimientos_salud), function(x)
gibberish(15, sym = F))

# ----- Estandarización de Red Hidrográfica

std_red_hidro <- shp_red_hidro |> transmute( ID_IDE = NA_character_, CUT_COM = NA_character_,
NOMBRE = idemop_str(Nombre), FID_DRENES = fid_drenes, DREN_TIPO = idemop_str(Dren_Tipo), FUENTE =
NA_character_, FECHA_ACT = NA_character_, LEVANTAMIE = NA_character_, geometry = geometry )

std_red_hidro$CUT_COM <- str_pad(comunas$ine17[match(idemop_str(shp_red_hidro$Comuna),
idemop_str(comunas$nombre))], 5, 'left', '0') std_red_hidro$NOM_COM <-
idemop_str(comunas$nombre[match(std_red_hidro$CUT_COM, str_pad(comunas$ine17, 5, 'left', '0'))])
std_red_hidro$CUT_REG <- region_desde_comuna(std_red_hidro$CUT_COM) std_red_hidro$NOM_REG <-
idemop_str(regiones$nombre_completo[match(std_red_hidro$CUT_REG, str_pad(regiones$region_16r, 2,
'left', '0'))]) std_red_hidro$CUT_PRO <- provincia_desde_comuna(std_red_hidro$CUT_COM)
std_red_hidro$NOM_PRO <- idemop_str(provincias$nombre[match(std_red_hidro$CUT_PRO,
str_pad(provincias$ine17, 3, 'left', '0'))])

std_red_hidro %<>% select( ID_IDE, CUT_REG, NOM_REG, CUT_PRO, NOM_PRO, CUT_COM, NOM_COM, NOMBRE,
FID_DRENES, DREN_TIPO, geometry ) %>% st_transform(crs_epsg)

std_red_hidro$ID_IDE <- sapply(1:nrow(std_red_hidro), function(x) gibberish(15, sym = F))

# ----- Estandarización Red vial

std_red_vial_ine <- shp_red_vial_ine |> transmute( ID_IDE = NA_character_, CUT_REG =
region_desde_comuna(CUT), NOM_REG = idemop_str(N_REGION), CUT_PRO = provincia_desde_comuna(CUT),
NOM_PRO = idemop_str(N_PROVINCIA), CUT_COM = str_pad(CUT, 5, 'left', '0'), NOM_COM =
idemop_str(N_COMUNA), NOMBRE = idemop_str(NOMBRE), ALIAS = idemop_str(ALIAS), ROL = ROL,
CLASE_COMU = idemop_str(CLASE_COMU), CLASE_URBA = idemop_str(CLASE_URBA), FUENTE = NA_character_,
FECHA_ACT = NA_character_, LEVANTAMIE = NA_character_, geometry = geometry ) |>
st_transform(crs_epsg) |> st_make_valid()

std_red_vial_ine$ID_IDE <- sapply(1:nrow(std_red_vial_ine), function(x) gibberish(15, sym = F))

# ----- Estandarización de SSR

std_ssr <- shp_servicios_sanitarios_rurales |> transmute( ID_IDE = ID_IDE, CUT_REG =
str_pad(COD_REG, 2, 'left', '0'), NOM_REG = idemop_str(REGION), CUT_PRO = str_pad(COD_PROV, 3,
'left', '0'), NOM_PRO = idemop_str(PROVINCIA), CUT_COM = str_pad(COD_COM, 5, 'left', '0'),
NOM_COM = idemop_str(COMUNA), NOMBRE = idemop_str(NOMBRE), MXSITEID = MXSITEID, FUENTE =
idemop_str(FUENTE), FECHA_ACTU = FECHA_ACTU, LEVANTAMIE = idemop_str(LEVANTAMIE), geometry =
geometry ) |> st_transform(crs_epsg)

```

```
# ----- Estandarización de Territorio Operacional

std_territorio_operacional <- shp_territorio_operacional |> transmute( ID_IDE = NA_character_,
CUT_REG = str_pad(REGION, 2, 'left', '0'), NOM_REG = NA_character_, COD_LOC = COD_LOC, COD_LOCUG =
COD_LOCUG, LOCALIDAD = idemop_str(LOCALIDAD), RUT = RUT, ID_TO = ID_TO, DEM_TO =
idemop_str(DEM_TO), COD_SIS = CODIGO_SIS, NOM_SIS = idemop_str(NOMBRE_SIS), FUENTE =
NA_character_, FECHA_ACTU = NA_character_, LEVANTAMIE = NA_character_, geometry = geometry ) |>
st_transform(crs_epsg) |> st_zm(drop = T, what = 'ZM') |> st_make_valid()

std_territorio_operacional$NOM_REG <-
idemop_str(regiones$nombre_completo[match(std_territorio_operacional$CUT_REG,
str_pad(regiones$region_16r, 2, 'left', '0'))]) std_territorio_operacional$ID_IDE <-
sapply(1:nrow(std_territorio_operacional), function(x) gibberish(15, sym = F))

# ----- Asignación de IDs

Asignación de IDs que simulan los IDs únicos que los datos tendrán en la Geodatabase final, para
efectos del ejemplo. Este ID puede utilizarse posteriormente para asociar los datos con sus
geometrías en la Geodatabase, a modo de variable llave.

std_aldea$ID_IDE <- sapply(1:nrow(std_aldea), function(x) gibberish(15, sym = F))
std_comunal$ID_IDE <- sapply(1:nrow(std_comunal), function(x) gibberish(15, sym = F))
std_distrital$ID_IDE <- sapply(1:nrow(std_distrital), function(x) gibberish(15, sym = F))
std_eje_vial$ID_IDE <- sapply(1:nrow(std_eje_vial), function(x) gibberish(15, sym = F))
std_limite_urbano$ID_IDE <- sapply(1:nrow(std_limite_urbano), function(x) gibberish(15, sym = F))
std_manzanas$ID_IDE <- sapply(1:nrow(std_manzanas), function(x) gibberish(15, sym = F))
std_edificaciones_rurales$ID_IDE <- sapply(1:nrow(std_edificaciones_rurales), function(x)
gibberish(15, sym = F))
```

```
# Wrapping up -----
# APLICACIÓN DE FUNCIÓN idemop_export()

std_df_names <- ls(pattern = '^std_')[ sapply(ls(pattern = '^std_'), function(y)
is.data.frame(get(y))) ]

std_df_names <- std_df_names[std_df_names != 'std_red_vial']

for (i in std_df_names) {

cli_alert_info(rmsg(str_glue('Processing {i}'))) idemop_export(i)
cli_alert_success(rmsg(str_glue('Exported {i}')))

}
```

## 11.2. Anexo 2. Caracterización de los nueve territorios CASA (ODEPA, 2024)

**Metropolitana:** Áreas de alta densidad poblacional (>1000 habitantes/km<sup>2</sup>) y muy cortos tiempos de viaje (<45 minutos) a centros urbanos principales. Predomina el paisaje urbano denso con infraestructura y conectividad bien desarrollada.

**Suburbana:** Áreas con densidad moderada (200-1000 habitantes/km<sup>2</sup>) y tiempos de viaje inferiores a 45 minutos. Estos espacios funcionan como extensiones de áreas metropolitanas con menor concentración urbana.

**Ciudad interconectada:** Áreas urbanas densas (>1000 habitantes/km<sup>2</sup>) con tiempos de viaje moderados (45-90 minutos) hacia grandes ciudades, mostrando una alta conectividad dentro de los espacios urbanos.

**Periurbana:** Áreas mixtas de baja densidad (<200 habitantes/ km<sup>2</sup>), pero con buenos niveles de acceso a grandes ciudades (<45 minutos), sirviendo como transición entre áreas urbanas y rurales.

**Ciudad intermedia:** Espacios mixtos de densidad media (200-1000 habitantes/km<sup>2</sup>) y tiempos de viaje moderados (45-90 minutos), con una configuración dispersa, pero conectada a los principales centros regionales.

**Ciudad independiente:** Áreas mixtas que configuran espacios densos (>1000 habitantes/km<sup>2</sup>) que se encuentran lejanos (>90 minutos) a otros grandes centros urbanos, funcionando como núcleos autónomos en regiones remotas.

**Rural interconectado:** Áreas rurales de baja densidad (<200-1000 habitantes/ km<sup>2</sup>) con tiempos de viaje moderados (45-90 minutos) hacia grandes ciudades, mostrando conectividad con zonas urbanas.

**Rural denso:** Áreas rurales más densas (200-1000 habitantes/km<sup>2</sup>), pero con largos tiempos de viaje (>90 minutos) a grandes centros urbanos, con una mezcla de actividad agrícola y residencial.

**Rural profundo:** Áreas rurales caracterizadas por baja densidad (<200 habitantes/ km<sup>2</sup>) y largos tiempos de viaje (>90 minutos) hacia centros urbanos principales. Predomina el paisaje natural y la desconexión de núcleos urbanos.

## 11.3. Anexo 3. Resultados del algoritmo HDBSCAN a nivel de comuna.

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Arica y Parinacota	Arica	Arica	11950	7306	4644	38.9%
Arica y Parinacota	Arica	Camarones	1296	956	340	26.2%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Arica y Parinacota	Parinacota	General Lagos	953	624	329	34.5%
Arica y Parinacota	Parinacota	Putre	1754	1314	440	25.1%
Tarapacá	Del Tamarugal	Camiña	1322	1099	223	16.9%
Tarapacá	Del Tamarugal	Colchane	2308	1964	344	14.9%
Tarapacá	Del Tamarugal	Huara	3296	2725	571	17.3%
Tarapacá	Del Tamarugal	Pica	1073	593	480	44.7%
Tarapacá	Del Tamarugal	Pozo Almonte	8608	5441	3167	36.8%
Tarapacá	Iquique	Alto Hospicio	1670	1417	253	15.1%
Tarapacá	Iquique	Iquique	3310	2629	681	20.6%
Antofagasta	Antofagasta	Antofagasta	1959	1561	398	20.3%
Antofagasta	Antofagasta	Mejillones	1123	852	271	24.1%
Antofagasta	Antofagasta	Sierra Gorda	877	690	187	21.3%
Antofagasta	Antofagasta	Taltal	1258	849	409	32.5%
Antofagasta	El Loa	Calama	6071	4003	2068	34.1%
Antofagasta	El Loa	Ollagüe	588	517	71	12.1%
Antofagasta	El Loa	San Pedro De Atacama	3250	2812	438	13.5%
Antofagasta	Tocopilla	María Elena	267	211	56	21%
Antofagasta	Tocopilla	Tocopilla	2565	2193	372	14.5%
Atacama	Chañaral	Chañaral	3019	2452	567	18.8%
Atacama	Chañaral	Diego De Almagro	1183	900	283	23.9%
Atacama	Copiapó	Caldera	9633	8362	1271	13.2%
Atacama	Copiapó	Copiapó	6477	4142	2335	36.1%
Atacama	Copiapó	Tierra Amarilla	2140	1517	623	29.1%
Atacama	Huasco	Alto Del Carmen	4031	2837	1194	29.6%
Atacama	Huasco	Freirina	2598	1917	681	26.2%
Atacama	Huasco	Huasco	3310	2626	684	20.7%
Atacama	Huasco	Vallenar	7659	4755	2904	37.9%
Coquimbo	Choapa	Canela	5296	3435	1861	35.1%
Coquimbo	Choapa	Illapel	7092	4353	2739	38.6%
Coquimbo	Choapa	Los Vilos	4225	3051	1174	27.8%
Coquimbo	Choapa	Salamanca	6432	5028	1404	21.8%
Coquimbo	Elqui	Andacollo	1152	849	303	26.3%
Coquimbo	Elqui	Coquimbo	17609	12632	4977	28.3%
Coquimbo	Elqui	La Higuera	3051	2485	566	18.6%
Coquimbo	Elqui	La Serena	16743	12058	4685	28%
Coquimbo	Elqui	Paiguano	3869	2868	1001	25.9%
Coquimbo	Elqui	Vicuña	6122	4581	1541	25.2%
Coquimbo	Limarí	Combarbalá	6481	4454	2027	31.3%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Coquimbo	Limarí	Monte Patria	9916	6980	2936	29.6%
Coquimbo	Limarí	Ovalle	17183	11244	5939	34.6%
Coquimbo	Limarí	Punitaqui	3994	2481	1513	37.9%
Coquimbo	Limarí	Río Hurtado	4030	2792	1238	30.7%
Valparaíso	Los Andes	Calle Larga	2004	1475	529	26.4%
Valparaíso	Los Andes	Los Andes	1919	1291	628	32.7%
Valparaíso	Los Andes	Rinconada	1176	769	407	34.6%
Valparaíso	Los Andes	San Esteban	4019	2766	1253	31.2%
Valparaíso	Marga Marga	Limache	14563	9695	4868	33.4%
Valparaíso	Marga Marga	Olmué	4059	3298	761	18.7%
Valparaíso	Marga Marga	Quilpué	5067	4228	839	16.6%
Valparaíso	Marga Marga	Villa Alemana	1497	1026	471	31.5%
Valparaíso	Petorca	Cabildo	3511	2584	927	26.4%
Valparaíso	Petorca	La Ligua	21914	16554	5360	24.5%
Valparaíso	Petorca	Papudo	1289	921	368	28.5%
Valparaíso	Petorca	Petorca	3421	2442	979	28.6%
Valparaíso	Petorca	Zapallar	2859	2319	540	18.9%
Valparaíso	Quillota	Calera	1279	1145	134	10.5%
Valparaíso	Quillota	Hijuelas	2890	2315	575	19.9%
Valparaíso	Quillota	La Cruz	1346	1103	243	18.1%
Valparaíso	Quillota	Nogales	2325	1245	1080	46.5%
Valparaíso	Quillota	Quillota	5517	3444	2073	37.6%
Valparaíso	San Antonio	Algarrobo	3775	2560	1215	32.2%
Valparaíso	San Antonio	Cartagena	1299	987	312	24%
Valparaíso	San Antonio	El Quisco	1245	835	410	32.9%
Valparaíso	San Antonio	El Tabo	1619	1388	231	14.3%
Valparaíso	San Antonio	San Antonio	4158	3388	770	18.5%
Valparaíso	San Antonio	Santo Domingo	4205	3315	890	21.2%
Valparaíso	San Felipe De Aconcagua	Catemu	2993	2127	866	28.9%
Valparaíso	San Felipe De Aconcagua	Llailay	4593	3335	1258	27.4%
Valparaíso	San Felipe De Aconcagua	Panquehue	1504	1036	468	31.1%
Valparaíso	San Felipe De Aconcagua	Putaendo	4502	3082	1420	31.5%
Valparaíso	San Felipe De Aconcagua	San Felipe	2936	2196	740	25.2%
Valparaíso	San Felipe De Aconcagua	Santa María	2312	1645	667	28.8%
Valparaíso	Valparaíso	Casablanca	6690	4941	1749	26.1%
Valparaíso	Valparaíso	Concón	1740	1366	374	21.5%
Valparaíso	Valparaíso	Juan Fernández	325	178	147	45.2%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Valparaíso	Valparaíso	Puchuncaví	3793	2883	910	24%
Valparaíso	Valparaíso	Quintero	4182	2904	1278	30.6%
Valparaíso	Valparaíso	Valparaíso	3442	2808	634	18.4%
Metropolitana	Chacabuco	Colina	12115	10647	1468	12.1%
Metropolitana	Chacabuco	Lampa	11827	9313	2514	21.3%
Metropolitana	Chacabuco	Tiltil	4239	2981	1258	29.7%
Metropolitana	Cordillera	Pirque	5974	4508	1466	24.5%
Metropolitana	Cordillera	San José De Maipo	4152	3194	958	23.1%
Metropolitana	Maipo	Buín	5431	4250	1181	21.7%
Metropolitana	Maipo	Calera De Tango	4733	3259	1474	31.1%
Metropolitana	Maipo	Paine	17961	12817	5144	28.6%
Metropolitana	Maipo	San Bernardo	2691	1569	1122	41.7%
Metropolitana	Melipilla	Alhué	2420	2041	379	15.7%
Metropolitana	Melipilla	Curacaví	7493	6003	1490	19.9%
Metropolitana	Melipilla	María Pinto	3541	2493	1048	29.6%
Metropolitana	Melipilla	Melipilla	25783	17182	8601	33.4%
Metropolitana	Melipilla	San Pedro	11630	7780	3850	33.1%
Metropolitana	Santiago	Lo Barnechea	1574	1280	294	18.7%
Metropolitana	Santiago	Maipú	1368	926	442	32.3%
Metropolitana	Santiago	Pudahuel	2870	1860	1010	35.2%
Metropolitana	Santiago	Quilicura	279	193	86	30.8%
Metropolitana	Talagante	El Monte	2664	2006	658	24.7%
Metropolitana	Talagante	Isla De Maipo	4058	3331	727	17.9%
Metropolitana	Talagante	Padre Hurtado	3429	2982	447	13%
Metropolitana	Talagante	Peñaflor	3406	2705	701	20.6%
Metropolitana	Talagante	Talagante	7582	6822	760	10%
O'Higgins	Cachapoal	Codegua	2689	1787	902	33.5%
O'Higgins	Cachapoal	Coinco	2303	1505	798	34.7%
O'Higgins	Cachapoal	Coltauco	5312	3924	1388	26.1%
O'Higgins	Cachapoal	Doñihue	2931	2405	526	17.9%
O'Higgins	Cachapoal	Graneros	1983	1428	555	28%
O'Higgins	Cachapoal	Las Cabras	11287	7602	3685	32.6%
O'Higgins	Cachapoal	Machalí	1618	1123	495	30.6%
O'Higgins	Cachapoal	Malloa	4232	2835	1397	33%
O'Higgins	Cachapoal	Mostazal	3272	2365	907	27.7%
O'Higgins	Cachapoal	Olivar	1788	1392	396	22.1%
O'Higgins	Cachapoal	Peumo	1143	837	306	26.8%
O'Higgins	Cachapoal	Pichidegua	4873	3600	1273	26.1%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
O'Higgins	Cachapoal	Quinta De Tilcoco	2562	1760	802	31.3%
O'Higgins	Cachapoal	Rancagua	5545	3992	1553	28%
O'Higgins	Cachapoal	Rengo	6027	4620	1407	23.3%
O'Higgins	Cachapoal	Requinoa	5416	3794	1622	29.9%
O'Higgins	Cachapoal	San Vicente	9973	6719	3254	32.6%
O'Higgins	Cardenal Caro	La Estrella	3054	2155	899	29.4%
O'Higgins	Cardenal Caro	Litueche	4436	3296	1140	25.7%
O'Higgins	Cardenal Caro	Marchihue	3561	2426	1135	31.9%
O'Higgins	Cardenal Caro	Navidad	6658	4103	2555	38.4%
O'Higgins	Cardenal Caro	Paredones	7382	5296	2086	28.3%
O'Higgins	Cardenal Caro	Pichilemu	4642	3273	1369	29.5%
O'Higgins	Colchagua	Chimbarongo	6007	4306	1701	28.3%
O'Higgins	Colchagua	Chépica	2867	2031	836	29.2%
O'Higgins	Colchagua	Lolol	2589	1728	861	33.3%
O'Higgins	Colchagua	Nancagua	3149	2217	932	29.6%
O'Higgins	Colchagua	Palmilla	4169	3263	906	21.7%
O'Higgins	Colchagua	Peralillo	2808	2038	770	27.4%
O'Higgins	Colchagua	Placilla	2927	2128	799	27.3%
O'Higgins	Colchagua	Pumanque	2719	1878	841	30.9%
O'Higgins	Colchagua	San Fernando	5442	3705	1737	31.9%
O'Higgins	Colchagua	Santa Cruz	7268	5008	2260	31.1%
Maule	Cauquenes	Cauquenes	6639	3869	2770	41.7%
Maule	Cauquenes	Chanco	2589	1587	1002	38.7%
Maule	Cauquenes	Pelluhue	3445	2450	995	28.9%
Maule	Curicó	Curicó	8498	5420	3078	36.2%
Maule	Curicó	Hualañé	2729	1839	890	32.6%
Maule	Curicó	Licantén	987	653	334	33.8%
Maule	Curicó	Molina	4749	3610	1139	24%
Maule	Curicó	Rauco	2639	1733	906	34.3%
Maule	Curicó	Romeral	5388	3289	2099	39%
Maule	Curicó	Sagrada Familia	5024	3594	1430	28.5%
Maule	Curicó	Teno	9915	6862	3053	30.8%
Maule	Curicó	Vichuquén	3888	2659	1229	31.6%
Maule	Linares	Colbún	8161	5850	2311	28.3%
Maule	Linares	Linares	11741	7113	4628	39.4%
Maule	Linares	Longaví	10915	6684	4231	38.8%
Maule	Linares	Parral	7425	4923	2502	33.7%
Maule	Linares	Retiro	7668	5065	2603	33.9%
Maule	Linares	San Javier	8720	5917	2803	32.1%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Maule	Linares	Villa Alegre	3934	2522	1412	35.9%
Maule	Linares	Yerbas Buenas	6557	4158	2399	36.6%
Maule	Talca	Constitución	9493	6755	2738	28.8%
Maule	Talca	Curepto	4813	3240	1573	32.7%
Maule	Talca	Empedrado	1067	625	442	41.4%
Maule	Talca	Maule	8145	5023	3122	38.3%
Maule	Talca	Pelarco	3590	2176	1414	39.4%
Maule	Talca	Pencahue	3629	2494	1135	31.3%
Maule	Talca	Río Claro	5015	3386	1629	32.5%
Maule	Talca	San Clemente	15024	9715	5309	35.3%
Maule	Talca	San Rafael	3189	1898	1291	40.5%
Maule	Talca	Talca	6115	4159	1956	32%
Ñuble	Diguillín	Bulnes	6195	3660	2535	40.9%
Ñuble	Diguillín	Chillán	15407	9793	5614	36.4%
Ñuble	Diguillín	Chillán Viejo	3147	2257	890	28.3%
Ñuble	Diguillín	El Carmen	5489	3224	2265	41.3%
Ñuble	Diguillín	Pemuco	2497	1651	846	33.9%
Ñuble	Diguillín	Pinto	5738	4225	1513	26.4%
Ñuble	Diguillín	Quillón	7738	4506	3232	41.8%
Ñuble	Diguillín	San Ignacio	6131	3137	2994	48.8%
Ñuble	Diguillín	Yungay	4926	2938	1988	40.4%
Ñuble	Itata	Cobquecura	3516	2255	1261	35.9%
Ñuble	Itata	Coelemu	4113	2751	1362	33.1%
Ñuble	Itata	Ninhue	2642	1928	714	27%
Ñuble	Itata	Portezuelo	2555	1741	814	31.9%
Ñuble	Itata	Quirihue	2928	1695	1233	42.1%
Ñuble	Itata	Ránquil	3184	2061	1123	35.3%
Ñuble	Itata	Treguaco	2795	1626	1169	41.8%
Ñuble	Punilla	Coihueco	10393	6910	3483	33.5%
Ñuble	Punilla	San Carlos	10650	6286	4364	41%
Ñuble	Punilla	San Fabián	2431	1682	749	30.8%
Ñuble	Punilla	San Nicolás	5288	3239	2049	38.7%
Ñuble	Punilla	Ñiquén	8061	4638	3423	42.5%
Biobío	Arauco	Arauco	5244	3653	1591	30.3%
Biobío	Arauco	Cañete	8291	4797	3494	42.1%
Biobío	Arauco	Contulmo	2016	1217	799	39.6%
Biobío	Arauco	Curanilahue	1833	1528	305	16.6%
Biobío	Arauco	Lebu	1970	1258	712	36.1%
Biobío	Arauco	Los Álamos	1662	1140	522	31.4%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Biobío	Arauco	Tirúa	3306	2432	874	26.4%
Biobío	Biobío	Alto Biobío	7209	4795	2414	33.5%
Biobío	Biobío	Antuco	1706	1081	625	36.6%
Biobío	Biobío	Cabrero	5271	3432	1839	34.9%
Biobío	Biobío	Laja	4956	3112	1844	37.2%
Biobío	Biobío	Los Ángeles	44391	27920	16471	37.1%
Biobío	Biobío	Mulchén	5125	3709	1416	27.6%
Biobío	Biobío	Nacimiento	5316	3109	2207	41.5%
Biobío	Biobío	Negrete	2531	1457	1074	42.4%
Biobío	Biobío	Quilaco	2383	1602	781	32.8%
Biobío	Biobío	Quilleco	3300	2348	952	28.8%
Biobío	Biobío	San Rosendo	959	564	395	41.2%
Biobío	Biobío	Santa Bárbara	6494	4380	2114	32.6%
Biobío	Biobío	Tucapel	3377	2036	1341	39.7%
Biobío	Biobío	Yumbel	14505	8913	5592	38.6%
Biobío	Concepción	Concepción	2925	2044	881	30.1%
Biobío	Concepción	Coronel	2721	2228	493	18.1%
Biobío	Concepción	Florida	5372	3075	2297	42.8%
Biobío	Concepción	Hualpén	417	330	87	20.9%
Biobío	Concepción	Hualqui	3566	2266	1300	36.5%
Biobío	Concepción	Lota	175	129	46	26.3%
Biobío	Concepción	Penco	520	255	265	51%
Biobío	Concepción	San Pedro De La Paz	106	84	22	20.8%
Biobío	Concepción	Santa Juana	4259	2054	2205	51.8%
Biobío	Concepción	Talcahuano	370	280	90	24.3%
Biobío	Concepción	Tomé	4574	3117	1457	31.9%
Araucanía	Cautín	Carahue	9362	4857	4505	48.1%
Araucanía	Cautín	Cholchol	4399	2313	2086	47.4%
Araucanía	Cautín	Cunco	8107	4689	3418	42.2%
Araucanía	Cautín	Curarrehue	5820	3418	2402	41.3%
Araucanía	Cautín	Freire	12899	8111	4788	37.1%
Araucanía	Cautín	Galvarino	6179	3531	2648	42.9%
Araucanía	Cautín	Gorbea	3386	2027	1359	40.1%
Araucanía	Cautín	Lautaro	10310	5998	4312	41.8%
Araucanía	Cautín	Loncoche	9137	5627	3510	38.4%
Araucanía	Cautín	Melipeuco	3898	3028	870	22.3%
Araucanía	Cautín	Nueva Imperial	13957	7912	6045	43.3%
Araucanía	Cautín	Padre Las Casas	18497	16552	1945	10.5%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Araucanía	Cautín	Perquenco	3032	1808	1224	40.4%
Araucanía	Cautín	Pitrufuquén	8774	5403	3371	38.4%
Araucanía	Cautín	Pucón	20438	11762	8676	42.5%
Araucanía	Cautín	Saavedra	9654	5675	3979	41.2%
Araucanía	Cautín	Temuco	12407	6973	5434	43.8%
Araucanía	Cautín	Teodoro Schmidt	9929	5813	4116	41.5%
Araucanía	Cautín	Toltén	4309	2617	1692	39.3%
Araucanía	Cautín	Vilcún	9533	6650	2883	30.2%
Araucanía	Cautín	Villarrica	32980	19031	13949	42.3%
Araucanía	Malleco	Angol	6422	3720	2702	42.1%
Araucanía	Malleco	Collipulli	4404	3056	1348	30.6%
Araucanía	Malleco	Curacautín	5139	3218	1921	37.4%
Araucanía	Malleco	Ercilla	2801	1839	962	34.3%
Araucanía	Malleco	Lonquimay	5206	3606	1600	30.7%
Araucanía	Malleco	Los Sauces	2500	1387	1113	44.5%
Araucanía	Malleco	Lumaco	3317	1833	1484	44.7%
Araucanía	Malleco	Purén	2706	1675	1031	38.1%
Araucanía	Malleco	Renaico	1430	953	477	33.4%
Araucanía	Malleco	Traiguén	2425	1435	990	40.8%
Araucanía	Malleco	Victoria	5308	3868	1440	27.1%
Los Ríos	Ranco	Futrono	5055	3698	1357	26.8%
Los Ríos	Ranco	La Unión	6984	3994	2990	42.8%
Los Ríos	Ranco	Lago Ranco	5227	3285	1942	37.2%
Los Ríos	Ranco	Río Bueno	11370	6296	5074	44.6%
Los Ríos	Valdivia	Corral	2000	1507	493	24.7%
Los Ríos	Valdivia	Lanco	3650	2028	1622	44.4%
Los Ríos	Valdivia	Los Lagos	7160	4710	2450	34.2%
Los Ríos	Valdivia	Mariquina	6862	4669	2193	32%
Los Ríos	Valdivia	Máfil	2272	1396	876	38.6%
Los Ríos	Valdivia	Paillaco	4379	2563	1816	41.5%
Los Ríos	Valdivia	Panguipulli	12725	6823	5902	46.4%
Los Ríos	Valdivia	Valdivia	12024	8323	3701	30.8%
Los Lagos	Chiloé	Ancud	9953	5605	4348	43.7%
Los Lagos	Chiloé	Castro	8161	7344	817	10%
Los Lagos	Chiloé	Chonchi	7953	4243	3710	46.6%
Los Lagos	Chiloé	Curaco De Véz	1896	934	962	50.7%
Los Lagos	Chiloé	Dalcahue	5569	2847	2722	48.9%
Los Lagos	Chiloé	Puqueldón	3134	2817	317	10.1%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Los Lagos	Chiloé	Queilén	3437	2964	473	13.8%
Los Lagos	Chiloé	Quellón	7825	4885	2940	37.6%
Los Lagos	Chiloé	Quemchi	5127	3129	1998	39%
Los Lagos	Chiloé	Quinchao	3863	3306	557	14.4%
Los Lagos	Llanquihue	Calbuco	10649	7047	3602	33.8%
Los Lagos	Llanquihue	Cochamó	3346	2037	1309	39.1%
Los Lagos	Llanquihue	Fresia	3938	2005	1933	49.1%
Los Lagos	Llanquihue	Frutillar	4716	2840	1876	39.8%
Los Lagos	Llanquihue	Llanquihue	2401	1287	1114	46.4%
Los Lagos	Llanquihue	Los Muermos	6627	3219	3408	51.4%
Los Lagos	Llanquihue	Mauñín	6024	3628	2396	39.8%
Los Lagos	Llanquihue	Puerto Montt	29632	19964	9668	32.6%
Los Lagos	Llanquihue	Puerto Varas	14413	10343	4070	28.2%
Los Lagos	Osorno	Osorno	8931	5379	3552	39.8%
Los Lagos	Osorno	Puerto Octay	5213	3084	2129	40.8%
Los Lagos	Osorno	Purranque	4837	2742	2095	43.3%
Los Lagos	Osorno	Puyehue	5821	3721	2100	36.1%
Los Lagos	Osorno	Río Negro	4876	3003	1873	38.4%
Los Lagos	Osorno	San Juan De La Costa	5579	3318	2261	40.5%
Los Lagos	Osorno	San Pablo	4797	2798	1999	41.7%
Los Lagos	Palena	Chaitén	2490	1675	815	32.7%
Los Lagos	Palena	Futaleufú	769	516	253	32.9%
Los Lagos	Palena	Hualaihué	4288	2834	1454	33.9%
Los Lagos	Palena	Palena	856	509	347	40.5%
Aysén	Aysén	Aysén	2777	1872	905	32.6%
Aysén	Aysén	Cisnes	2372	1677	695	29.3%
Aysén	Aysén	Guaitecas	134	97	37	27.6%
Aysén	Capitán Prat	Cochrane	752	424	328	43.6%
Aysén	Capitán Prat	O'higgins	551	408	143	26%
Aysén	Capitán Prat	Tortel	474	323	151	31.9%
Aysén	Coyhaique	Coyhaique	11328	7024	4304	38%
Aysén	Coyhaique	Lago Verde	947	670	277	29.3%
Aysén	General Carrera	Chile Chico	3058	2088	970	31.7%
Aysén	General Carrera	Río Ibáñez	3036	1968	1068	35.2%
Magallanes	Antártica Chilena	Cabo De Hornos	161	95	66	41%
Magallanes	Magallanes	Laguna Blanca	457	276	181	39.6%
Magallanes	Magallanes	Punta Arenas	8161	5720	2441	29.9%
Magallanes	Magallanes	Río Verde	311	92	219	70.4%

Región	Provincia	Comuna	Total de viviendas	Viviendas clusterizadas	Viviendas ruido	Proporción ruido (ruido/total)
Magallanes	Magallanes	San Gregorio	446	154	292	65.5%
Magallanes	Tierra Del Fuego	Porvenir	2022	1185	837	41.4%
Magallanes	Tierra Del Fuego	Primavera	1696	1106	590	34.8%
Magallanes	Tierra Del Fuego	Timaukel	602	374	228	37.9%
Magallanes	Última Esperanza	Natales	3408	2561	847	24.9%
Magallanes	Última Esperanza	Torres Del Paine	634	319	315	49.7%

#### 11.4. Anexo 3. Diccionario de datos.

Capa: viviendas_clusters_hdb			
Campo	Tipo	Origen	Descripción
n_region	integer	INE (APC 2023)	Nombre región
codigo_region	double	INE (APC 2023)	Código identificador por región
n_provincia	character	INE (APC 2023)	Nombre provincia
n_comuna	character	INE (APC 2023)	Nombre comuna
cut	double	INE (APC 2023)	Código Único Territorial
macrozona	character	INE (APC 2023)	Nombre de la macrozona del país
n_distrito	character	INE (APC 2023)	Nombre distrito
id_distrito	character	INE (APC 2023)	Identificador del distrito
tipo_distrito	character	INE (APC 2023)	Detalle del tipo de distrito
fuelle	character	INE (APC 2023)	Fuente interna INE de digitalización de la edificación rural
uso_edificacion	character	INE (APC 2023)	Uso de la edificación rural (Edificación, Otro uso, Vivienda y Vivienda colectiva)
longitud	double	OCUC	Coordenada Longitud
latitud	double	OCUC	Coordenada Latitud
fid_via	double	INE y OCUC	Identificador de red vial
point_x	double	INE (APC 2023)	Coordenada UTM X (Este)
point_y	double	INE (APC 2023)	Coordenada UTM Y (Norte)
point_z	double	OCUC	Altitud de la edificación (metros)
point_d	double	OCUC	Distancia al camino más cerano (metros)
cluster	double	OCUC	Número de cluster original por comuna
id_cluster	character	OCUC	Identificador del Cluster de HDBSCAN (sin ajustar)
membership_prob	double	OCUC	Probabilidad de pertenencia a cluster 0 - 1 (1 indica máxima pertenencia)
outlier_score	double	OCUC	Puntaje de outlier (puntajes más altos indican valores atípicos)
lambda	double	OCUC	Distancia de alcance mutua (valor original)
mrd	double	OCUC	Distancia de alcance mutua (en metros)
id_cluster_final	character	OCUC	Identificador del Cluster Final de HDBSCAN (remueve puntos con mrd < mrd0.9)
es_ruido	double	OCUC	Indica si edificación pertenece es ruido (1) o pertenece a un cluster (0)
tipo_ruido	character	OCUC	Indica si edificación es ruido permanente o ruido satélite
Capa: poligonos_clusters_hdb			
Campo	Tipo	Descripción	

n_region	integer	Nombre región del cluster
n_provincia	character	Nombre provincia
n_comuna	character	Nombre comuna
cut	integer	Código Único Territorial
macrozona	character	Nombre macrozona del país
id_cluster_final	character	Identificador del Cluster Final de HDBSCAN (remueve puntos con mrd > mrd0.9)
n_viviendas	double	Número total de viviendas dentro del cluster
membership_min	double	Probabilidad de pertenencia (mínima) de todas las viviendas dentro de un cluster
membership_max	double	Probabilidad de pertenencia (máxima) de todas las viviendas dentro de un cluster
membership_mean	double	Probabilidad de pertenencia (promedio) de todas las viviendas dentro de un cluster
membership_std	double	Probabilidad de pertenencia (desv. estándar) de todas las viviendas dentro de un cluster
outlier_min	double	Puntaje outlier (mínima) de todas las viviendas dentro de un cluster
outlier_max	double	Puntaje outlier (máxima) de todas las viviendas dentro de un cluster
outlier_mean	double	Puntaje outlier (promedio) de todas las viviendas dentro de un cluster
outlier_std	double	Puntaje outlier (desv. estándar) de todas las viviendas dentro de un cluster
mrd_min	double	Mínimo de distancia de alcance mutua (en metros)
mrd_max	double	Máximo de distancia de alcance mutua (en metros)
mrd_mean	double	Promedio de distancia de alcance mutua (en metros)
mrd_std	double	Desviación estándar de distancia de alcance mutua (en metros)
mrd_range	double	Rango de distancia de alcance mutua (en metros)
class_arranque	character	Clasificación del tamaño del cluster según número de arranques
consistencia_cluster	character	Tipo de consistencia de los clúster

Capa: mapas_region		
Campo	Tipo	Descripción
codigo_region	double	Código por región
macrozona	character	Nombre de la macrozona
n_region	character	Nombre región del cluster
n_cluster	integer	Nombre cluster
n_viv	integer	Número de viviendas
min_viv	integer	Mínimo de viviendas
max_viv	integer	Máximo de viviendas
mean_viv	double	Promedio de viviendas
dif	integer	Diferencia entre el máximo y mínimo de viviendas
viv_por_cluster	double	Número de viviendas por clúster
cluster_por_viv	double	Número de clusters por viviendas

Capa: mapas_comuna		
Campo	Tipo	Descripción
codigo_region	double	Código por región
macrozona	character	Nombre de la macrozona
n_region	character	Nombre región del cluster
n_cluster	integer	Nombre cluster
n_viv	integer	Número de viviendas
min_viv	integer	Mínimo de viviendas
max_viv	integer	Máximo de viviendas
mean_viv	double	Promedio de viviendas
dif	integer	Diferencia entre el máximo y mínimo de viviendas
viv_por_cluster	double	Número de viviendas por clúster
cluster_por_viv	double	Número de clusters por viviendas

Capa: mapas_distrito		
Campo	Tipo	Descripción
n_region	character	Nombre región del cluster
cut	integer	Nombre de la macrozona
n_provincia	character	Nombre provincia
n_comuna	character	Nombre comuna
n_distrito	character	Nombre distrito
id_distrito	character	Identificador del distrito
viviendas_cluster	integer	Número de viviendas clusterizadas
viviendas_ruido	integer	Número de viviendas clasificadas como ruido
total_viviendas	integer	Número total de viviendas

prop_viviendas_cluster	double	Proporción de viviendas clusterizadas y el total de viviendas
prop_viviendas_ruido	double	Proporción de viviendas ruido y el total de viviendas

### 11.5. Anexo 4. Resultados del algoritmo HDBSCAN en clústeres a nivel de comuna (MB: Muy Bajo, B: Bajo, M: Medio, A: Alto, MA: Muy Alto, C: Consistente y I: Inconsistente).

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters					Nivel de consistencia		Tamaño de los clusters (%)					Nivel de consistencia	
				MB	B	M	A	MA	C	I	MB	B	M	A	MA	C	I
Arica y Parinacota	Arica	Arica	83	41	27	12	1	2	78	5	49,4	32,5	14,5	1,2	2,4	94	6
Arica y Parinacota	Arica	Camarones	17	9	7	1	0	0	16	1	52,9	41,2	5,9	0	0	94,1	5,9
Arica y Parinacota	Parinacota	General Lagos	12	8	3	1	0	0	11	1	66,7	25	8,3	0	0	91,7	8,3
Arica y Parinacota	Parinacota	Putre	22	12	5	5	0	0	20	2	54,5	22,7	22,7	0	0	90,9	9,1
Tarapacá	Del Tamarugal	Camiña	12	4	5	2	1	0	11	1	33,3	41,7	16,7	8,3	0	91,7	8,3
Tarapacá	Del Tamarugal	Colchane	21	10	7	2	2	0	20	1	47,6	33,3	9,5	9,5	0	95,2	4,8
Tarapacá	Del Tamarugal	Huara	36	17	10	9	0	0	34	2	47,2	27,8	25	0	0	94,4	5,6
Tarapacá	Del Tamarugal	Pica	10	6	3	1	0	0	9	1	60	30	10	0	0	90	10
Tarapacá	Del Tamarugal	Pozo Almonte	43	20	7	12	3	1	40	3	46,5	16,3	27,9	7	2,3	93	7
Tarapacá	Iquique	Alto Hospicio	8	3	4	0	0	1	7	1	37,5	50	0	0	12,5	87,5	12,5
Tarapacá	Iquique	Iquique	22	8	6	7	1	0	20	2	36,4	27,3	31,8	4,5	0	90,9	9,1
Antofagasta	Antofagasta	Antofagasta	14	8	3	1	2	0	13	1	57,1	21,4	7,1	14,3	0	92,9	7,1
Antofagasta	Antofagasta	Mejillones	7	3	1	2	1	0	6	1	42,9	14,3	28,6	14,3	0	85,7	14,3
Antofagasta	Antofagasta	Sierra Gorda	4	2	0	0	2	0	3	1	50	0	0	50	0	75	25
Antofagasta	Antofagasta	Taltal	13	9	2	2	0	0	12	1	69,2	15,4	15,4	0	0	92,3	7,7
Antofagasta	El Loa	Calama	32	15	5	8	3	1	30	2	46,9	15,6	25	9,4	3,1	93,8	6,2
Antofagasta	El Loa	Ollagüe	5	2	1	2	0	0	4	1	40	20	40	0	0	80	20
Antofagasta	El Loa	San Pedro De Atacama	16	6	3	4	2	1	15	1	37,5	18,8	25	12,5	6,2	93,8	6,2
Antofagasta	Tocopilla	María Elena	1	0	0	1	0	0	1	0	0	0	100	0	0	100	0
Antofagasta	Tocopilla	Tocopilla	17	4	6	5	2	0	16	1	23,5	35,3	29,4	11,8	0	94,1	5,9
Atacama	Chañaral	Chañaral	19	6	5	6	2	0	18	1	31,6	26,3	31,6	10,5	0	94,7	5,3
Atacama	Chañaral	Diego De Almagro	8	2	3	2	1	0	7	1	25	37,5	25	12,5	0	87,5	12,5
Atacama	Copiapó	Caldera	25	13	1	4	3	4	23	2	52	4	16	12	16	92	8
Atacama	Copiapó	Copiapó	43	20	14	8	0	1	40	3	46,5	32,6	18,6	0	2,3	93	7
Atacama	Copiapó	Tierra Amarilla	19	11	4	3	1	0	18	1	57,9	21,1	15,8	5,3	0	94,7	5,3
Atacama	Huasco	Alto Del Carmen	46	25	17	3	1	0	43	3	54,3	37	6,5	2,2	0	93,5	6,5
Atacama	Huasco	Freirina	22	11	4	7	0	0	20	2	50	18,2	31,8	0	0	90,9	9,1
Atacama	Huasco	Huasco	19	7	6	3	2	1	18	1	36,8	31,6	15,8	10,5	5,3	94,7	5,3
Atacama	Huasco	Vallenar	45	23	14	5	2	1	42	3	51,1	31,1	11,1	4,4	2,2	93,3	6,7
Coquimbo	Choapa	Canela	36	20	6	9	0	1	34	2	55,6	16,7	25	0	2,8	94,4	5,6
Coquimbo	Choapa	Illapel	65	34	20	10	1	0	61	4	52,3	30,8	15,4	1,5	0	93,8	6,2
Coquimbo	Choapa	Los Vilos	29	14	4	9	2	0	27	2	48,3	13,8	31	6,9	0	93,1	6,9
Coquimbo	Choapa	Salamanca	43	14	14	11	4	0	40	3	32,6	32,6	25,6	9,3	0	93	7
Coquimbo	Elqui	Andacollo	9	4	4	0	1	0	8	1	44,4	44,4	0	11,1	0	88,9	11,1
Coquimbo	Elqui	Coquimbo	79	30	21	18	7	3	75	4	38	26,6	22,8	8,9	3,8	94,9	5,1
Coquimbo	Elqui	La Higuera	18	9	2	4	3	0	17	1	50	11,1	22,2	16,7	0	94,4	5,6
Coquimbo	Elqui	La Serena	84	31	12	29	11	1	79	5	36,9	14,3	34,5	13,1	1,2	94	6
Coquimbo	Elqui	Paiguano	28	12	6	9	0	1	26	2	42,9	21,4	32,1	0	3,6	92,9	7,1
Coquimbo	Elqui	Vicuña	39	13	12	11	3	0	37	2	33,3	30,8	28,2	7,7	0	94,9	5,1
Coquimbo	Limarí	Combarbalá	43	17	11	13	2	0	40	3	39,5	25,6	30,2	4,7	0	93	7
Coquimbo	Limarí	Monte Patria	77	34	21	19	3	0	73	4	44,2	27,3	24,7	3,9	0	94,8	5,2
Coquimbo	Limarí	Ovalle	161	83	54	22	1	1	153	8	51,6	33,5	13,7	0,6	0,6	95	5
Coquimbo	Limarí	Punitaqui	40	21	12	7	0	0	38	2	52,5	30	17,5	0	0	95	5
Coquimbo	Limarí	Río Hurtado	33	13	14	5	1	0	31	2	39,4	42,4	15,2	3	0	93,9	6,1
Valparaíso	Los Andes	Calle Larga	18	8	5	4	1	0	17	1	44,4	27,8	22,2	5,6	0	94,4	5,6
Valparaíso	Los Andes	Los Andes	19	12	4	2	1	0	18	1	63,2	21,1	10,5	5,3	0	94,7	5,3
Valparaíso	Los Andes	Rinconada	9	4	2	3	0	0	8	1	44,4	22,2	33,3	0	0	88,9	11,1

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters					Nivel de consistencia		Tamaño de los clusters					Nivel de consistencia	
				MB	B	M	A	MA	C	I	Tamaño de los clusters					Nivel de consistencia	
											Tamaño de los clusters (%)					C	I
Valparaíso	Los Andes	San Esteban	30	10	8	12	0	0	28	2	33,3	26,7	40	0	0		
Valparaíso	Marga Marga	Limache	73	30	25	11	5	2	69	4	41,1	34,2	15,1	6,8	2,7	94,5	5,5
Valparaíso	Marga Marga	Olmué	12	6	4	1	0	1	11	1	50	33,3	8,3	0	8,3	91,7	8,3
Valparaíso	Marga Marga	Quilpué	15	6	4	3	1	1	14	1	40	26,7	20	6,7	6,7	93,3	6,7
Valparaíso	Marga Marga	Villa Alemana	11	4	2	5	0	0	10	1	36,4	18,2	45,5	0	0	90,9	9,1
Valparaíso	Petorca	Cabildo	29	12	7	9	1	0	27	2	41,4	24,1	31	3,4	0	93,1	6,9
Valparaíso	Petorca	La Ligua	98	36	25	22	10	5	93	5	36,7	25,5	22,4	10,2	5,1	94,9	5,1
Valparaíso	Petorca	Papudo	9	2	5	2	0	0	8	1	22,2	55,6	22,2	0	0	88,9	11,1
Valparaíso	Petorca	Petorca	24	9	8	6	1	0	22	2	37,5	33,3	25	4,2	0	91,7	8,3
Valparaíso	Petorca	Zapallar	11	4	2	3	0	2	10	1	36,4	18,2	27,3	0	18,2	90,9	9,1
Valparaíso	Quillota	Calera	3	0	0	1	2	0	2	1	0	0	33,3	66,7	0	66,7	33,3
Valparaíso	Quillota	Hijuelas	15	5	4	4	2	0	14	1	33,3	26,7	26,7	13,3	0	93,3	6,7
Valparaíso	Quillota	La Cruz	7	4	2	0	0	1	6	1	57,1	28,6	0	0	14,3	85,7	14,3
Valparaíso	Quillota	Nogales	24	14	7	3	0	0	22	2	58,3	29,2	12,5	0	0	91,7	8,3
Valparaíso	Quillota	Quillota	36	13	11	12	0	0	34	2	36,1	30,6	33,3	0	0	94,4	5,6
Valparaíso	San Antonio	Algarrobo	17	5	7	2	2	1	16	1	29,4	41,2	11,8	11,8	5,9	94,1	5,9
Valparaíso	San Antonio	Cartagena	8	1	3	4	0	0	7	1	12,5	37,5	50	0	0	87,5	12,5
Valparaíso	San Antonio	El Quisco	13	6	4	3	0	0	12	1	46,2	30,8	23,1	0	0	92,3	7,7
Valparaíso	San Antonio	El Tabo	6	1	1	3	0	1	5	1	16,7	16,7	50	0	16,7	83,3	16,7
Valparaíso	San Antonio	San Antonio	24	8	9	5	1	1	22	2	33,3	37,5	20,8	4,2	4,2	91,7	8,3
Valparaíso	San Antonio	Santo Domingo	22	9	6	3	4	0	20	2	40,9	27,3	13,6	18,2	0	90,9	9,1
Valparaíso	San Felipe De Aconcagua	Catemu	22	9	6	6	1	0	20	2	40,9	27,3	27,3	4,5	0	90,9	9,1
Valparaíso	San Felipe De Aconcagua	Llailay	25	9	9	4	3	0	23	2	36	36	16	12	0	92	8
Valparaíso	San Felipe De Aconcagua	Panquehue	9	4	1	4	0	0	8	1	44,4	11,1	44,4	0	0	88,9	11,1
Valparaíso	San Felipe De Aconcagua	Putendo	28	10	10	6	2	0	26	2	35,7	35,7	21,4	7,1	0	92,9	7,1
Valparaíso	San Felipe De Aconcagua	San Felipe	20	7	6	6	1	0	19	1	35	30	30	5	0	95	5
Valparaíso	San Felipe De Aconcagua	Santa María	15	4	6	4	1	0	14	1	26,7	40	26,7	6,7	0	93,3	6,7
Valparaíso	Valparaíso	Casablanca	45	15	12	14	4	0	42	3	33,3	26,7	31,1	8,9	0	93,3	6,7
Valparaíso	Valparaíso	Concón	9	1	2	4	2	0	8	1	11,1	22,2	44,4	22,2	0	88,9	11,1
Valparaíso	Valparaíso	Juan Fernández	2	0	1	1	0	0	1	1	0	50	50	0	0	50	50
Valparaíso	Valparaíso	Puchuncaví	28	15	6	5	1	1	26	2	53,6	21,4	17,9	3,6	3,6	92,9	7,1
Valparaíso	Valparaíso	Quintero	28	8	10	8	2	0	26	2	28,6	35,7	28,6	7,1	0	92,9	7,1
Valparaíso	Valparaíso	Valparaíso	15	8	4	1	0	2	14	1	53,3	26,7	6,7	0	13,3	93,3	6,7
Metropolitana	Chacabuco	Colina	19	5	7	3	2	2	18	1	26,3	36,8	15,8	10,5	10,5	94,7	5,3
Metropolitana	Chacabuco	Lampa	46	17	11	13	4	1	43	3	37	23,9	28,3	8,7	2,2	93,5	6,5
Metropolitana	Chacabuco	Tiltil	33	14	11	7	1	0	31	2	42,4	33,3	21,2	3	0	93,9	6,1
Metropolitana	Cordillera	Pirque	31	10	10	7	2	2	29	2	32,3	32,3	22,6	6,5	6,5	93,5	6,5
Metropolitana	Cordillera	San José De Maipo	33	14	10	7	1	1	31	2	42,4	30,3	21,2	3	3	93,9	6,1
Metropolitana	Maipo	Buín	45	14	16	14	1	0	42	3	31,1	35,6	31,1	2,2	0	93,3	6,7
Metropolitana	Maipo	Calera De Tango	30	14	6	9	0	1	28	2	46,7	20	30	0	3,3	93,3	6,7
Metropolitana	Maipo	Paine	120	49	31	34	3	3	114	6	40,8	25,8	28,3	2,5	2,5	95	5
Metropolitana	Maipo	San Bernardo	25	13	7	5	0	0	23	2	52	28	20	0	0	92	8

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters					Nivel de consistencia		Tamaño de los clusters					Nivel de consistencia	
				MB	B	M	A	MA	C	I	Tamaño de los clusters					Nivel de consistencia	
											Tamaño de los clusters (%)					C	I
Metropolitana	Melipilla	Alhué	13	5	1	5	2	0	12	1	38,5	7,7	38,5	15,4	0		
Metropolitana	Melipilla	Curacaví	30	9	7	6	4	4	28	2	30	23,3	20	13,3	13,3	93,3	6,7
Metropolitana	Melipilla	María Pinto	32	15	9	7	1	0	30	2	46,9	28,1	21,9	3,1	0	93,8	6,2
Metropolitana	Melipilla	Melipilla	183	78	55	41	7	2	173	10	42,6	30,1	22,4	3,8	1,1	94,5	5,5
Metropolitana	Melipilla	San Pedro	80	35	26	13	5	1	76	4	43,8	32,5	16,2	6,2	1,2	95	5
Metropolitana	Santiago	Lo Barnechea	14	5	4	5	0	0	13	1	35,7	28,6	35,7	0	0	92,9	7,1
Metropolitana	Santiago	Maipú	9	4	2	2	1	0	8	1	44,4	22,2	22,2	11,1	0	88,9	11,1
Metropolitana	Santiago	Pudahuel	28	16	7	5	0	0	26	2	57,1	25	17,9	0	0	92,9	7,1
Metropolitana	Santiago	Quilicura	3	1	2	0	0	0	2	1	33,3	66,7	0	0	0	66,7	33,3
Metropolitana	Talagante	El Monte	23	8	8	7	0	0	21	2	34,8	34,8	30,4	0	0	91,3	8,7
Metropolitana	Talagante	Isla De Maipo	18	4	5	5	3	1	17	1	22,2	27,8	27,8	16,7	5,6	94,4	5,6
Metropolitana	Talagante	Padre Hurtado	15	2	3	7	3	0	14	1	13,3	20	46,7	20	0	93,3	6,7
Metropolitana	Talagante	Peñaflor	15	1	8	4	1	1	14	1	6,7	53,3	26,7	6,7	6,7	93,3	6,7
Metropolitana	Talagante	Talagante	3	1	0	1	0	1	2	1	33,3	0	33,3	0	33,3	66,7	33,3
O'Higgins	Cachapoal	Codegua	28	14	10	4	0	0	26	2	50	35,7	14,3	0	0	92,9	7,1
O'Higgins	Cachapoal	Coinco	17	5	8	4	0	0	16	1	29,4	47,1	23,5	0	0	94,1	5,9
O'Higgins	Cachapoal	Coltauco	37	17	6	13	1	0	35	2	45,9	16,2	35,1	2,7	0	94,6	5,4
O'Higgins	Cachapoal	Doñihue	9	1	4	2	0	2	8	1	11,1	44,4	22,2	0	22,2	88,9	11,1
O'Higgins	Cachapoal	Graneros	17	6	3	8	0	0	16	1	35,3	17,6	47,1	0	0	94,1	5,9
O'Higgins	Cachapoal	Las Cabras	92	46	27	16	2	1	87	5	50	29,3	17,4	2,2	1,1	94,6	5,4
O'Higgins	Cachapoal	Machalí	15	7	4	4	0	0	14	1	46,7	26,7	26,7	0	0	93,3	6,7
O'Higgins	Cachapoal	Malloa	35	12	13	10	0	0	33	2	34,3	37,1	28,6	0	0	94,3	5,7
O'Higgins	Cachapoal	Mostazal	23	8	6	9	0	0	21	2	34,8	26,1	39,1	0	0	91,3	8,7
O'Higgins	Cachapoal	Olivar	12	3	7	1	0	1	11	1	25	58,3	8,3	0	8,3	91,7	8,3
O'Higgins	Cachapoal	Peumo	12	5	5	2	0	0	11	1	41,7	41,7	16,7	0	0	91,7	8,3
O'Higgins	Cachapoal	Pichidegua	40	19	10	11	0	0	38	2	47,5	25	27,5	0	0	95	5
O'Higgins	Cachapoal	Quinta De Tilcoco	18	7	5	5	1	0	17	1	38,9	27,8	27,8	5,6	0	94,4	5,6
O'Higgins	Cachapoal	Rancagua	42	21	10	9	1	1	39	3	50	23,8	21,4	2,4	2,4	92,9	7,1
O'Higgins	Cachapoal	Rengo	43	19	13	8	2	1	40	3	44,2	30,2	18,6	4,7	2,3	93	7
O'Higgins	Cachapoal	Requinoa	46	23	15	7	0	1	43	3	50	32,6	15,2	0	2,2	93,5	6,5
O'Higgins	Cachapoal	San Vicente	80	34	26	18	2	0	76	4	42,5	32,5	22,5	2,5	0	95	5
O'Higgins	Cardenal Caro	La Estrella	25	13	6	4	2	0	23	2	52	24	16	8	0	92	8
O'Higgins	Cardenal Caro	Litueche	30	7	12	10	1	0	28	2	23,3	40	33,3	3,3	0	93,3	6,7
O'Higgins	Cardenal Caro	Marchihue	27	11	10	5	1	0	25	2	40,7	37	18,5	3,7	0	92,6	7,4
O'Higgins	Cardenal Caro	Navidad	42	25	8	6	1	2	39	3	59,5	19	14,3	2,4	4,8	92,9	7,1
O'Higgins	Cardenal Caro	Paredones	46	22	9	12	1	2	43	3	47,8	19,6	26,1	2,2	4,3	93,5	6,5
O'Higgins	Cardenal Caro	Pichilemu	36	16	9	9	2	0	34	2	44,4	25	25	5,6	0	94,4	5,6
O'Higgins	Colchagua	Chimbarongo	44	14	19	10	1	0	41	3	31,8	43,2	22,7	2,3	0	93,2	6,8
O'Higgins	Colchagua	Chépica	30	13	11	6	0	0	28	2	43,3	36,7	20	0	0	93,3	6,7
O'Higgins	Colchagua	Lolol	28	14	10	4	0	0	26	2	50	35,7	14,3	0	0	92,9	7,1
O'Higgins	Colchagua	Nancagua	25	9	10	4	2	0	23	2	36	40	16	8	0	92	8
O'Higgins	Colchagua	Palmilla	32	9	10	12	1	0	30	2	28,1	31,2	37,5	3,1	0	93,8	6,2
O'Higgins	Colchagua	Peralillo	23	9	9	4	1	0	21	2	39,1	39,1	17,4	4,3	0	91,3	8,7
O'Higgins	Colchagua	Placilla	22	9	7	4	2	0	20	2	40,9	31,8	18,2	9,1	0	90,9	9,1

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters					Nivel de consistencia		Tamaño de los clusters					Nivel de consistencia	
				MB	B	M	A	MA	C	I	Tamaño de los clusters					Nivel de consistencia	
											Tamaño de los clusters (%)					C	I
O'Higgins	Colchagua	Pumanque	23	9	10	3	1	0	21	2	39,1	43,5	13	4,3	0		
O'Higgins	Colchagua	San Fernando	45	23	10	11	1	0	42	3	51,1	22,2	24,4	2,2	0	93,3	6,7
O'Higgins	Colchagua	Santa Cruz	59	26	16	16	1	0	56	3	44,1	27,1	27,1	1,7	0	94,9	5,1
Maule	Cauquenes	Cauquenes	54	29	13	12	0	0	51	3	53,7	24,1	22,2	0	0	94,4	5,6
Maule	Cauquenes	Chanco	19	8	8	2	1	0	18	1	42,1	42,1	10,5	5,3	0	94,7	5,3
Maule	Cauquenes	Pelluhue	28	18	6	2	2	0	26	2	64,3	21,4	7,1	7,1	0	92,9	7,1
Maule	Curicó	Curicó	71	39	19	11	1	1	67	4	54,9	26,8	15,5	1,4	1,4	94,4	5,6
Maule	Curicó	Hualañé	25	13	7	4	1	0	23	2	52	28	16	4	0	92	8
Maule	Curicó	Licantén	10	5	4	1	0	0	9	1	50	40	10	0	0	90	10
Maule	Curicó	Molina	33	18	6	7	0	2	31	2	54,5	18,2	21,2	0	6,1	93,9	6,1
Maule	Curicó	Rauco	24	12	7	5	0	0	22	2	50	29,2	20,8	0	0	91,7	8,3
Maule	Curicó	Romerol	49	28	15	3	3	0	46	3	57,1	30,6	6,1	6,1	0	93,9	6,1
Maule	Curicó	Sagrada Familia	40	13	16	10	1	0	38	2	32,5	40	25	2,5	0	95	5
Maule	Curicó	Teno	80	33	21	25	1	0	76	4	41,2	26,2	31,2	1,2	0	95	5
Maule	Curicó	Vichuquén	35	15	13	7	0	0	33	2	42,9	37,1	20	0	0	94,3	5,7
Maule	Linares	Colbún	57	27	14	14	1	1	54	3	47,4	24,6	24,6	1,8	1,8	94,7	5,3
Maule	Linares	Linares	116	64	38	12	2	0	110	6	55,2	32,8	10,3	1,7	0	94,8	5,2
Maule	Linares	Longaví	89	48	19	21	1	0	84	5	53,9	21,3	23,6	1,1	0	94,4	5,6
Maule	Linares	Parral	58	26	16	14	1	1	55	3	44,8	27,6	24,1	1,7	1,7	94,8	5,2
Maule	Linares	Retiro	61	26	21	13	1	0	58	3	42,6	34,4	21,3	1,6	0	95,1	4,9
Maule	Linares	San Javier	74	35	22	16	1	0	70	4	47,3	29,7	21,6	1,4	0	94,6	5,4
Maule	Linares	Villa Alegre	36	21	6	9	0	0	34	2	58,3	16,7	25	0	0	94,4	5,6
Maule	Linares	Yerbas Buenas	61	34	18	8	1	0	58	3	55,7	29,5	13,1	1,6	0	95,1	4,9
Maule	Talca	Constitución	65	35	15	9	5	1	61	4	53,8	23,1	13,8	7,7	1,5	93,8	6,2
Maule	Talca	Curepto	37	13	15	8	1	0	35	2	35,1	40,5	21,6	2,7	0	94,6	5,4
Maule	Talca	Empedrado	12	9	1	2	0	0	11	1	75	8,3	16,7	0	0	91,7	8,3
Maule	Talca	Maule	55	21	14	19	1	0	52	3	38,2	25,5	34,5	1,8	0	94,5	5,5
Maule	Talca	Pelarco	25	9	10	5	1	0	23	2	36	40	20	4	0	92	8
Maule	Talca	Pencahue	35	17	12	5	1	0	33	2	48,6	34,3	14,3	2,9	0	94,3	5,7
Maule	Talca	Río Claro	43	21	10	12	0	0	40	3	48,8	23,3	27,9	0	0	93	7
Maule	Talca	San Clemente	119	55	39	23	1	1	113	6	46,2	32,8	19,3	0,8	0,8	95	5
Maule	Talca	San Rafael	31	18	9	4	0	0	29	2	58,1	29	12,9	0	0	93,5	6,5
Maule	Talca	Talca	51	20	19	11	1	0	48	3	39,2	37,3	21,6	2	0	94,1	5,9
Nuble	Diguillín	Bulnes	53	26	15	11	1	0	50	3	49,1	28,3	20,8	1,9	0	94,3	5,7
Nuble	Diguillín	Chillán	94	37	29	20	7	1	89	5	39,4	30,9	21,3	7,4	1,1	94,7	5,3
Nuble	Diguillín	Chillán Viejo	22	12	4	4	1	1	20	2	54,5	18,2	18,2	4,5	4,5	90,9	9,1
Nuble	Diguillín	El Carmen	53	30	17	5	1	0	50	3	56,6	32,1	9,4	1,9	0	94,3	5,7
Nuble	Diguillín	Pemuco	18	8	6	3	1	0	17	1	44,4	33,3	16,7	5,6	0	94,4	5,6
Nuble	Diguillín	Pinto	30	17	3	7	1	2	28	2	56,7	10	23,3	3,3	6,7	93,3	6,7
Nuble	Diguillín	Quillón	56	24	20	11	1	0	53	3	42,9	35,7	19,6	1,8	0	94,6	5,4
Nuble	Diguillín	San Ignacio	51	27	19	4	1	0	48	3	52,9	37,3	7,8	2	0	94,1	5,9
Nuble	Diguillín	Yungay	37	16	11	9	1	0	35	2	43,2	29,7	24,3	2,7	0	94,6	5,4
Nuble	Itata	Cobquecura	32	22	3	6	1	0	30	2	68,8	9,4	18,8	3,1	0	93,8	6,2
Nuble	Itata	Coelemu	36	20	9	6	1	0	34	2	55,6	25	16,7	2,8	0	94,4	5,6
Nuble	Itata	Ninhue	15	8	4	2	0	1	14	1	53,3	26,7	13,3	0	6,7	93,3	6,7
Nuble	Itata	Portezuelo	24	8	11	5	0	0	22	2	33,3	45,8	20,8	0	0	91,7	8,3
Nuble	Itata	Quirihue	36	24	10	2	0	0	34	2	66,7	27,8	5,6	0	0	94,4	5,6
Nuble	Itata	Ránquil	29	15	7	7	0	0	27	2	51,7	24,1	24,1	0	0	93,1	6,9
Nuble	Itata	Treguaco	20	8	9	2	1	0	19	1	40	45	10	5	0	95	5
Nuble	Punilla	Coihueco	74	28	27	17	1	1	70	4	37,8	36,5	23	1,4	1,4	94,6	5,4
Nuble	Punilla	San Carlos	88	44	22	22	0	0	83	5	50	25	25	0	0	94,3	5,7
Nuble	Punilla	San Fabián	23	11	6	6	0	0	21	2	47,8	26,1	26,1	0	0	91,3	8,7
Nuble	Punilla	San Nicolás	55	33	15	7	0	0	52	3	60	27,3	12,7	0	0	94,5	5,5
Nuble	Punilla	Niquén	69	34	22	13	0	0	65	4	49,3	31,9	18,8	0	0	94,2	5,8
Biobío	Arauco	Arauco	40	21	9	6	4	0	38	2	52,5	22,5	15	10	0	95	5
Biobío	Arauco	Cañete	60	34	10	14	2	0	57	3	56,7	16,7	23,3	3,3	0	95	5
Biobío	Arauco	Contulmo	19	12	3	4	0	0	18	1	63,2	15,8	21,1	0	0	94,7	5,3

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters					Nivel de consistencia		Tamaño de los clusters					Nivel de consistencia	
				MB	B	M	A	MA	C	I	Tamaño de los clusters					Nivel de consistencia	
											Tamaño de los clusters (%)					Nivel de consistencia (%)	
										MB	B	M	A	MA	C	I	
Biobío	Arauco	Curanilahue	12	5	2	3	2	0	11	1	41,7	16,7	25	16,7	0	91,7	8,3
Biobío	Arauco	Lebu	24	16	5	3	0	0	22	2	66,7	20,8	12,5	0	0	91,7	8,3
Biobío	Arauco	Los Álamos	16	10	3	3	0	0	15	1	62,5	18,8	18,8	0	0	93,8	6,2
Biobío	Arauco	Tirúa	15	6	4	2	2	1	14	1	40	26,7	13,3	13,3	6,7	93,3	6,7
Biobío	Biobío	Alto Biobío	72	35	25	11	1	0	68	4	48,6	34,7	15,3	1,4	0	94,4	5,6
Biobío	Biobío	Antuco	16	7	6	3	0	0	15	1	43,8	37,5	18,8	0	0	93,8	6,2
Biobío	Biobío	Cabrero	48	21	20	6	1	0	45	3	43,8	41,7	12,5	2,1	0	93,8	6,2
Biobío	Biobío	Laja	45	24	12	9	0	0	42	3	53,3	26,7	20	0	0	93,3	6,7
Biobío	Biobío	Los Ángeles	283	117	82	69	14	1	268	15	41,3	29	24,4	4,9	0,4	94,7	5,3
Biobío	Biobío	Mulchén	50	25	15	8	2	0	47	3	50	30	16	4	0	94	6
Biobío	Biobío	Nacimiento	45	23	12	10	0	0	42	3	51,1	26,7	22,2	0	0	93,3	6,7
Biobío	Biobío	Negrete	21	9	6	6	0	0	20	1	42,9	28,6	28,6	0	0	95,2	4,8
Biobío	Biobío	Quilaco	25	16	5	3	1	0	23	2	64	20	12	4	0	92	8
Biobío	Biobío	Quilleco	26	10	9	7	0	0	24	2	38,5	34,6	26,9	0	0	92,3	7,7
Biobío	Biobío	San Rosendo	10	5	4	1	0	0	9	1	50	40	10	0	0	90	10
Biobío	Biobío	Santa Bárbara	63	32	19	12	0	0	59	4	50,8	30,2	19	0	0	93,7	6,3
Biobío	Biobío	Tucapel	31	14	14	3	0	0	29	2	45,2	45,2	9,7	0	0	93,5	6,5
Biobío	Biobío	Yumbel	130	69	44	13	4	0	123	7	53,1	33,8	10	3,1	0	94,6	5,4
Biobío	Concepción	Concepción	18	7	4	7	0	0	17	1	38,9	22,2	38,9	0	0	94,4	5,6
Biobío	Concepción	Coronel	18	7	4	6	1	0	17	1	38,9	22,2	33,3	5,6	0	94,4	5,6
Biobío	Concepción	Florida	43	21	13	9	0	0	40	3	48,8	30,2	20,9	0	0	93	7
Biobío	Concepción	Hualpén	5	2	2	1	0	0	4	1	40	40	20	0	0	80	20
Biobío	Concepción	Hualqui	26	14	5	5	2	0	24	2	53,8	19,2	19,2	7,7	0	92,3	7,7
Biobío	Concepción	Lota	3	2	1	0	0	0	2	1	66,7	33,3	0	0	0	66,7	33,3
Biobío	Concepción	Penco	6	4	2	0	0	0	5	1	66,7	33,3	0	0	0	83,3	16,7
Biobío	Concepción	San Pedro De La Paz	3	3	0	0	0	0	2	1	100	0	0	0	0	66,7	33,3
Biobío	Concepción	Santa Juana	33	20	8	5	0	0	31	2	60,6	24,2	15,2	0	0	93,9	6,1
Biobío	Concepción	Talcahuano	5	2	3	0	0	0	4	1	40	60	0	0	0	80	20
Biobío	Concepción	Tomé	40	17	15	6	2	0	38	2	42,5	37,5	15	5	0	95	5
Araucanía	Cautín	Carahue	71	42	14	14	1	0	67	4	59,2	19,7	19,7	1,4	0	94,4	5,6
Araucanía	Cautín	Cholchol	35	19	11	5	0	0	33	2	54,3	31,4	14,3	0	0	94,3	5,7
Araucanía	Cautín	Cunco	60	29	16	14	1	0	57	3	48,3	26,7	23,3	1,7	0	95	5
Araucanía	Cautín	Curarrehue	57	41	7	9	0	0	54	3	71,9	12,3	15,8	0	0	94,7	5,3
Araucanía	Cautín	Freire	62	26	18	13	2	3	58	4	41,9	29	21	3,2	4,8	93,5	6,5
Araucanía	Cautín	Galvarino	61	37	18	6	0	0	58	3	60,7	29,5	9,8	0	0	95,1	4,9
Araucanía	Cautín	Gorbea	26	15	7	3	1	0	24	2	57,7	26,9	11,5	3,8	0	92,3	7,7
Araucanía	Cautín	Lautaro	89	52	24	12	1	0	84	5	58,4	27	13,5	1,1	0	94,4	5,6
Araucanía	Cautín	Loncoche	57	32	15	6	3	1	54	3	56,1	26,3	10,5	5,3	1,8	94,7	5,3
Araucanía	Cautín	Melipeuco	16	6	4	2	3	1	15	1	37,5	25	12,5	18,8	6,2	93,8	6,2
Araucanía	Cautín	Nueva Imperial	101	50	36	12	1	2	96	5	49,5	35,6	11,9	1	2	95	5
Araucanía	Cautín	Padre Las Casas	4	1	1	0	0	2	3	1	25	25	0	0	50	75	25
Araucanía	Cautín	Perquenco	25	10	9	6	0	0	23	2	40	36	24	0	0	92	8
Araucanía	Cautín	Pitrufquén	59	23	22	11	2	1	56	3	39	37,3	18,6	3,4	1,7	94,9	5,1
Araucanía	Cautín	Pucón	144	74	41	23	5	1	136	8	51,4	28,5	16	3,5	0,7	94,4	5,6
Araucanía	Cautín	Saavedra	47	20	15	9	0	3	44	3	42,6	31,9	19,1	0	6,4	93,6	6,4
Araucanía	Cautín	Temuco	88	53	21	11	2	1	83	5	60,2	23,9	12,5	2,3	1,1	94,3	5,7
Araucanía	Cautín	Teodoro Schmidt	62	29	16	12	5	0	58	4	46,8	25,8	19,4	8,1	0	93,5	6,5
Araucanía	Cautín	Toltén	32	17	10	2	3	0	30	2	53,1	31,2	6,2	9,4	0	93,8	6,2
Araucanía	Cautín	Vilcún	43	20	9	10	1	3	40	3	46,5	20,9	23,3	2,3	7	93	7
Araucanía	Cautín	Villarrica	246	133	64	42	6	1	233	13	54,1	26	17,1	2,4	0,4	94,7	5,3
Araucanía	Malleco	Angol	64	40	14	9	1	0	60	4	62,5	21,9	14,1	1,6	0	93,8	6,2
Araucanía	Malleco	Collipulli	39	24	6	8	0	1	37	2	61,5	15,4	20,5	0	2,6	94,9	5,1
Araucanía	Malleco	Curacautín	46	26	13	6	1	0	43	3	56,5	28,3	13	2,2	0	93,5	6,5
Araucanía	Malleco	Ercilla	11	1	5	3	1	1	10	1	9,1	45,5	27,3	9,1	9,1	90,9	9,1
Araucanía	Malleco	Lonquimay	32	12	11	6	2	1	30	2	37,5	34,4	18,8	6,2	3,1	93,8	6,2
Araucanía	Malleco	Los Sauces	27	17	8	2	0	0	25	2	63	29,6	7,4	0	0	92,6	7,4

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters					Nivel de consistencia		Tamaño de los clusters					Nivel de consistencia	
				MB	B	M	A	MA	C	I	Tamaño de los clusters					Nivel de consistencia	
											Tamaño de los clusters (%)					C	I
Araucanía	Malleco	Lumaco	27	15	6	6	0	0	25	2	55,6	22,2	22,2	0	0		
Araucanía	Malleco	Purén	23	13	6	3	1	0	21	2	56,5	26,1	13	4,3	0	91,3	8,7
Araucanía	Malleco	Renaico	13	5	5	3	0	0	12	1	38,5	38,5	23,1	0	0	92,3	7,7
Araucanía	Malleco	Traiguén	24	11	12	1	0	0	22	2	45,8	50	4,2	0	0	91,7	8,3
Araucanía	Malleco	Victoria	43	26	6	9	1	1	40	3	60,5	14	20,9	2,3	2,3	93	7
Los Ríos	Ranco	Futroneo	35	18	7	7	2	1	33	2	51,4	20	20	5,7	2,9	94,3	5,7
Los Ríos	Ranco	La Unión	50	23	14	11	2	0	47	3	46	28	22	4	0	94	6
Los Ríos	Ranco	Lago Ranco	34	15	11	5	3	0	32	2	44,1	32,4	14,7	8,8	0	94,1	5,9
Los Ríos	Ranco	Río Bueno	84	49	19	14	2	0	79	5	58,3	22,6	16,7	2,4	0	94	6
Los Ríos	Valdivia	Corral	18	7	7	4	0	0	17	1	38,9	38,9	22,2	0	0	94,4	5,6
Los Ríos	Valdivia	Lanco	32	14	15	3	0	0	30	2	43,8	46,9	9,4	0	0	93,8	6,2
Los Ríos	Valdivia	Los Lagos	69	40	16	12	1	0	65	4	58	23,2	17,4	1,4	0	94,2	5,8
Los Ríos	Valdivia	Mariquina	48	25	9	13	0	1	45	3	52,1	18,8	27,1	0	2,1	93,8	6,2
Los Ríos	Valdivia	Máfil	23	13	6	4	0	0	21	2	56,5	26,1	17,4	0	0	91,3	8,7
Los Ríos	Valdivia	Paillaco	32	13	12	6	1	0	30	2	40,6	37,5	18,8	3,1	0	93,8	6,2
Los Ríos	Valdivia	Panguipulli	92	57	19	14	2	0	87	5	62	20,7	15,2	2,2	0	94,6	5,4
Los Ríos	Valdivia	Valdivia	84	37	19	24	3	1	79	5	44	22,6	28,6	3,6	1,2	94	6
Los Lagos	Chiloé	Ancud	86	41	31	14	0	0	81	5	47,7	36	16,3	0	0	94,2	5,8
Los Lagos	Chiloé	Castro	2	0	0	0	0	2	1	1	0	0	0	0	100	50	50
Los Lagos	Chiloé	Chonchi	62	31	22	8	1	0	58	4	50	35,5	12,9	1,6	0	93,5	6,5
Los Lagos	Chiloé	Curaco De Vélaz	11	6	3	1	1	0	10	1	54,5	27,3	9,1	9,1	0	90,9	9,1
Los Lagos	Chiloé	Dalcahue	47	26	13	8	0	0	44	3	55,3	27,7	17	0	0	93,6	6,4
Los Lagos	Chiloé	Puqueldón	2	0	0	1	0	1	1	1	0	0	50	0	50	50	50
Los Lagos	Chiloé	Queilén	9	6	2	0	0	1	8	1	66,7	22,2	0	0	11,1	88,9	11,1
Los Lagos	Chiloé	Quellón	76	48	12	16	0	0	72	4	63,2	15,8	21,1	0	0	94,7	5,3
Los Lagos	Chiloé	Quemchi	32	17	9	2	3	1	30	2	53,1	28,1	6,2	9,4	3,1	93,8	6,2
Los Lagos	Chiloé	Quinchao	10	1	1	4	3	1	9	1	10	10	40	30	10	90	10
Los Lagos	Llanquihue	Calbuco	62	29	16	13	2	2	58	4	46,8	25,8	21	3,2	3,2	93,5	6,5
Los Lagos	Llanquihue	Cochamó	26	15	6	4	1	0	24	2	57,7	23,1	15,4	3,8	0	92,3	7,7
Los Lagos	Llanquihue	Fresia	33	21	8	4	0	0	31	2	63,6	24,2	12,1	0	0	93,9	6,1
Los Lagos	Llanquihue	Fruillillar	36	19	9	7	1	0	34	2	52,8	25	19,4	2,8	0	94,4	5,6
Los Lagos	Llanquihue	Llanquihue	23	13	7	3	0	0	21	2	56,5	30,4	13	0	0	91,3	8,7
Los Lagos	Llanquihue	Los Muermos	51	29	15	6	1	0	48	3	56,9	29,4	11,8	2	0	94,1	5,9
Los Lagos	Llanquihue	Mauñín	52	29	15	7	1	0	49	3	55,8	28,8	13,5	1,9	0	94,2	5,8
Los Lagos	Llanquihue	Puerto Montt	184	91	46	31	13	3	174	10	49,5	25	16,8	7,1	1,6	94,6	5,4
Los Lagos	Llanquihue	Puerto Varas	74	30	23	17	3	1	70	4	40,5	31,1	23	4,1	1,4	94,6	5,4
Los Lagos	Osorno	Osorno	75	37	23	13	2	0	71	4	49,3	30,7	17,3	2,7	0	94,7	5,3
Los Lagos	Osorno	Puerto Octay	49	34	10	3	2	0	46	3	69,4	20,4	6,1	4,1	0	93,9	6,1
Los Lagos	Osorno	Purranque	43	20	20	3	0	0	40	3	46,5	46,5	7	0	0	93	7
Los Lagos	Osorno	Puyehue	49	32	11	4	1	1	46	3	65,3	22,4	8,2	2	2	93,9	6,1
Los Lagos	Osorno	Río Negro	45	22	18	4	1	0	42	3	48,9	40	8,9	2,2	0	93,3	6,7
Los Lagos	Osorno	San Juan De La Costa	43	22	11	9	1	0	40	3	51,2	25,6	20,9	2,3	0	93	7
Los Lagos	Osorno	San Pablo	50	32	12	6	0	0	47	3	64	24	12	0	0	94	6
Los Lagos	Palena	Chaitén	28	15	9	4	0	0	26	2	53,6	32,1	14,3	0	0	92,9	7,1
Los Lagos	Palena	Futaleufú	9	5	3	1	0	0	8	1	55,6	33,3	11,1	0	0	88,9	11,1
Los Lagos	Palena	Hualaihué	45	27	12	5	1	0	42	3	60	26,7	11,1	2,2	0	93,3	6,7
Los Lagos	Palena	Palena	6	2	2	2	0	0	5	1	33,3	33,3	33,3	0	0	83,3	16,7
Aysén	Aysén	Aysén	25	13	7	4	1	0	23	2	52	28	16	4	0	92	8
Aysén	Aysén	Cisnes	19	8	5	6	0	0	18	1	42,1	26,3	31,6	0	0	94,7	5,3
Aysén	Aysén	Guaitecas	3	3	0	0	0	0	2	1	100	0	0	0	0	66,7	33,3
Aysén	Capitán Prat	Cochrane	6	4	1	1	0	0	5	1	66,7	16,7	16,7	0	0	83,3	16,7
Aysén	Capitán Prat	O'higgins	4	2	1	1	0	0	3	1	50	25	25	0	0	75	25
Aysén	Capitán Prat	Tortel	3	1	0	2	0	0	2	1	33,3	0	66,7	0	0	66,7	33,3
Aysén	Coyhaique	Coyhaique	88	41	27	17	3	0	83	5	46,6	30,7	19,3	3,4	0	94,3	5,7
Aysén	Coyhaique	Lago Verde	7	3	1	3	0	0	6	1	42,9	14,3	42,9	0	0	85,7	14,3
Aysén	General Carrera	Chile Chico	22	14	4	2	1	1	20	2	63,6	18,2	9,1	4,5	4,5	90,9	9,1

Región	Provincia	Comuna	Total cluster	Tamaño de los clusters					Nivel de consistencia		Tamaño de los clusters					Nivel de consistencia	
				MB	B	M	A	MA	C	I	Tamaño de los clusters					Nivel de consistencia	
											Tamaño de los clusters (%)					C	I
MB	B	M	A	MA	C	I	MB	B	M	A	MA	C	I				
Aysén	General Carrera	Río Ibáñez	19	12	3	2	1	1	18	1	63,2	15,8	10,5	5,3	5,3	94,7	5,3
Magallanes	Antártica Chilena	Cabo De Hornos	2	1	1	0	0	0	1	1	50	50	0	0	0	50	50
Magallanes	Magallanes	Laguna Blanca	4	1	2	1	0	0	3	1	25	50	25	0	0	75	25
Magallanes	Magallanes	Punta Arenas	54	16	22	14	1	1	51	3	29,6	40,7	25,9	1,9	1,9	94,4	5,6
Magallanes	Magallanes	Río Verde	2	1	1	0	0	0	1	1	50	50	0	0	0	50	50
Magallanes	Magallanes	San Gregorio	2	0	2	0	0	0	1	1	0	100	0	0	0	50	50
Magallanes	Tierra Del Fuego	Porvenir	20	14	3	2	1	0	19	1	70	15	10	5	0	95	5
Magallanes	Tierra Del Fuego	Primavera	23	18	4	1	0	0	21	2	78,3	17,4	4,3	0	0	91,3	8,7
Magallanes	Tierra Del Fuego	Timaukel	8	5	3	0	0	0	7	1	62,5	37,5	0	0	0	87,5	12,5
Magallanes	Última Esperanza	Natales	19	11	7	0	0	1	18	1	57,9	36,8	0	0	5,3	94,7	5,3
Magallanes	Última Esperanza	Torres Del Paine	8	6	2	0	0	0	7	1	75	25	0	0	0	87,5	12,5