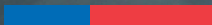




# Asentamientos Humanos Rurales en Chile, Clasificación Comunal

Una aproximación desde el análisis espacial de  
la concentración y dispersión de la población

División de Políticas y Desarrollo Territorial  
Departamento de Estudios y Análisis Territorial  
Septiembre de 2020



©Subsecretaría de Desarrollo Regional y Administrativo

Número de Inscripción: 2020-A-6498

Fecha de Inscripción: 18 de agosto 2020

# Asentamientos Humanos Rurales en Chile, Clasificación Comunal

## **Jefa de División de Políticas y Desarrollo Territorial**

Eleonora Espinoza Hernández

## **Coordinación**

Departamento de Estudios y Análisis Territorial

Andrés Barrientos Cárdenas, Jefe de Departamento de Estudios y Análisis Territorial

## **Elaboración**

Matías Poch Clavero, Profesional de la Unidad de Investigación y Estudios Territoriales

## **Revisión**

Profesionales de la División de Políticas y Desarrollo Territorial

## **Diseño**

Natalia Quintela

## **Fotografía**

Patricio Cortés

## **SUBDERE, septiembre 2020**

La información territorial en el presente documento, que se refiera o relacione con los límites y fronteras de Chile, no compromete en modo alguno al Estado de Chile de acuerdo al artículo 2do, letra g) del DFL N°83 de 1979 del Ministerio de Relaciones Exteriores. Además esta información es de carácter referencial e ilustrativo, para la comprensión de los fenómenos y dinámicas espaciales analizadas cuantitativamente.

## PRESENTACIÓN

Para impulsar el desarrollo territorial y el proceso de descentralización, la Subsecretaría de Desarrollo Regional y Administrativo ha realizado un gran esfuerzo en reconocer y comprender la heterogeneidad de los territorios del país, yendo más allá de las estadísticas a escala comunal. Prueba de ello, han sido los diversos estudios referidos a las condiciones de aislamiento de las localidades del país.

Al hablar de complejidad, se suele asociar a una característica desfavorable. Pero desde la perspectiva del trabajo de la SUBDERE desarrolla, representa una oportunidad para revertir el centralismo y movilizar potencialidades endógenas de cada territorio, focalizar de mejor manera los esfuerzos administrativos y financieros, mejorando la calidad de vida de nuestros habitantes, sobre todo, una oportunidad para ponernos al servicio de las personas que habitan en estas localidades.

Esta experiencia de estudios a escalas subcomunales, implica conocer la complejidad de los territorios, más allá de datos estadísticos consolidados a escala

comunal. Este es un gran desafío ya que se debe a cuantificar, medir y sistematizar la información espacial a escalas más precisas. Tener una medición de la complejidad, en términos espaciales y de localización, permite diferenciar comunas que tienen condiciones territoriales más complejas, lo que implicaría mayores dificultades para atender a estas localidades por parte del municipio.

El presente estudio incorpora, al proceso de cálculo los últimos datos oficiales disponibles, del Instituto Nacional de Estadísticas INE. Pero que, a diferencia de los estudios desarrollados en la línea de la medición de las condiciones de aislamiento, los datos de tiempos y distancias, ya no se realizan a un punto en el mapa, que representa una localidad, el cálculo está hecho en base al levantamiento de vivienda rural del pre Censo 2016. Esto implica un enorme desafío en términos de cálculo, pues son más de 950.000 viviendas distribuidas en el territorio nacional, esto nos obliga a contar con sistemas de información altamente especializados, que puedan ser capaz de manejar bases de datos de más de 350 millones de registros.

La voluntad que nos congrega, apunta a facilitar a todos los ciudadanos del país, a una mejor intervención por parte de las instituciones subnacionales, para que así puedan desarrollarse de manera integral. Como bien lo expresa el Programa de Gobierno del Presidente Sebastián Piñera Echenique es ineludible “fomentar un desarrollo armónico que aproveche todo el potencial de nuestro país y otorgue iguales oportunidades a todos”.

**Juan Manuel Masferrer Vidal**  
Subsecretario Desarrollo Regional y Administrativo





# ÍNDICE

<b>I. ANTECEDENTES GENERALES</b> .....	8
<b>II. OBJETIVOS</b> .....	15
2.1 OBJETIVO GENERAL .....	15
2.2 OBJETIVOS ESPECÍFICOS .....	15
<b>III. SÍNTESIS METODOLÓGICA</b> .....	17
3.1 PROPUESTA METODOLÓGICA .....	22
3.1.1 Datos de entrada .....	23
3.1.2 Construcción de indicadores (variables del método de clasificación) .....	23
3.1.3 Métodos de Clasificación .....	24
<b>IV. DESARROLLO</b> .....	26
4.1 ANÁLISIS DE VARIABLES .....	26
4.2 ELECCIÓN DEL MÉTODO DE CLASIFICACIÓN .....	29
4.3 COMUNAS SEGÚN CLASIFICACIÓN .....	32
4.3.1 Resultados de Clasificación .....	36
<b>V. CONCLUSIONES</b> .....	38
<b>VI. ANEXOS</b> .....	39
6.1 LISTADO DE COMUNAS SEGÚN CLASIFICACIÓN .....	39
6.2 MÉTODOS DE CLASIFICACIÓN UTILIZADOS .....	55
6.2.1 Agrupamiento por K medias .....	55
6.2.2 Agrupamiento Jerárquico .....	58
6.2.3 Elección del método de agrupamiento .....	61
6.3 CONCEPTOS TEÓRICOS .....	63
6.3.1 Conceptos de econometría espacial .....	63
6.3.2 Indicador de Concentración/Dispersión .....	69
6.4 INFORMACIÓN UTILIZADA .....	70
6.4.1 Base de datos de tiempos de desplazamiento y distancias .....	70
<b>BIBLIOGRAFÍA</b> .....	72

## I. ANTECEDENTES GENERALES

La densidad de población, medida en hab/km<sup>2</sup> o hab/hectáreas, ha sido utilizada frecuentemente como una aproximación a la dispersión o concentración de la población en un territorio (usualmente la comuna). Uno de los problemas de este indicador es que, en sectores rurales, la base de cálculo es variable para cada comuna, excluyendo información relacionada con la distribución espacial de la población.

En Chile, los municipios están definidos como “corporaciones autónomas de derecho público, con personalidad jurídica y patrimonio propio, cuya finalidad es satisfacer las necesidades de la comunidad local y asegurar su participación en el progreso económico, social y cultural de las respectivas comunas”<sup>1</sup>. Por lo tanto, en su accionar territorial, el municipio tiene la “capacidad de gobernar, ejercer autoridad y realizar acciones que ordenen, dispongan y organicen el territorio contenido en su jurisdicción con el fin de satisfacer las demandas de su población” en términos

de espacio público, equipamiento, tránsito, aseo y ornato y regulación de las áreas urbanas<sup>2</sup>.

La División de Municipalidades de SUBDERE desarrolló una tipología municipal para la provisión de servicios<sup>3</sup>. En dicho documento, se utiliza como aproximación para medir la dispersión, el principio físico de la entropía (proveniente de las leyes de la termodinámica). La entropía<sup>4</sup> se considera como medida de desorden de un sistema de partículas que tienen cierta función de probabilidad (en este caso, las localidades y su población).

2 El Municipio y la Gobernabilidad del Territorio Comunal. Pág 47, Fuente L, Allard P, Orellana A.

3 Tipología comunal para la provisión de servicios municipales, División de Municipalidades, Departamento de Finanzas Municipales. SUBDERE, Ministerio del Interior.

4 La entropía puede interpretarse como una medida de la distribución aleatoria de un sistema. Se dice que un sistema altamente distribuido al azar, tiene alta entropía. Un sistema en una condición improbable, tendrá una tendencia natural a reorganizarse a una condición más probable (similar a una distribución al azar), reorganización que dará como resultado un aumento de la entropía. La entropía alcanzará un máximo, cuando el sistema se acerque al equilibrio y entonces se alcanzará la configuración de mayor probabilidad (distribución al azar).

1 Ley N°18.695 Ley Orgánica Constitucional de Municipalidades. Art 1.







Los asentamientos humanos “normalmente se localizan cerca o en medio de zonas con recursos naturales suficientes o abundantes, como el agua dulce. Muchos también se ubican en sitios con recursos naturales económicamente importantes debido a la conveniencia de su comercio o desarrollo industrial”<sup>5</sup>.

En el Atlas Nacional de España, el hábitat rural hace referencia al asentamiento de la población en el territorio, teniendo en cuenta sus variables naturales. Así, el hábitat rural se refiere a las aldeas, los lugares, los pueblos o los barrios existentes en España, en municipios que no alcanzan un efectivo superior a 10.000 habitantes<sup>6</sup>.

Una solución, para el problema de la base métrica (superficie) de la densidad de la población, es la planteada en la publicación “Estudios de Microrregiones y Corredores Rurales”<sup>7</sup>, donde la superficie utilizada en el cálculo está referida al espacio ocupado habitacionalmente por localidades rurales. En esta publicación se realiza un análisis de la dispersión de la población, basado en la definición de “zonas rurales en el espacio”. El estudio considera zonas de aproximadamente 40 km<sup>2</sup> (cuadrados de 6,5 x 6,5 km.), dado que se considera dentro del límite máximo dis-

gregación y de homogenización del espacio<sup>8</sup>. Se hace referencia a que, además del tamaño de la población, es importante medir los niveles de accesibilidad<sup>9</sup> de cada una de las localidades rurales, hacia los principales centros prestadores de servicios básicos.

Otros de los indicadores de concentración y dispersión de población, poseen fuertes supuestos de orden jerárquicos (ej: índice de primacía), regla rango tamaño y que la condición de ruralidad está asociada a dispersión<sup>10</sup> (ej: índice de Demangeon).

En el Atlas Nacional de España, se refieren a la distribución de la población en función de la agrupación de viviendas. Sin embargo, no definen medidas de agrupación como distancias entre cada una de las viviendas. Sin embargo, realizan las siguientes definiciones:

- Concentrado, la población se agrupa en aldeas grandes.

5 <https://www.geoenciclopedia.com/asentamientos-humanos/>

6 Atlas Nacional de España, Cap. 8 pág. 204.

7 Nº 4 junio 2008 Revista Electrónica de Ciencias Sociales. ISSN: 1886-8452.

8 Esta distancia es adecuada para efectuar desplazamientos cotidianos en los modos de transporte disponibles en las áreas rurales. En un cuadrado de aproximadamente 6,5Km de lado la distancia promedio a recorrer entre cualquier punto es de 6,5 Km., lo que implicaría los siguientes tiempos: entre 01h:0m y 01h:30m en caminata; entre 00h:30m y 01h:00m en tracción animal; entre 00h:10 y 00h:20 min. en automotores.

9 Se entiende como accesibilidad a las facilidades que permiten que un lugar determinado pueda ser visitado. Las facilidades de acceder están determinadas por los costos de trasladarse hacia un punto de interés.

10 Atlas Nacional de España, Cap. 8 pág. 206,

- Disperso, en los casos donde los lugares son minúsculos, de 2 a 10 casas.
- Mixto, entre las 2 situaciones anteriores.
- Diseminado, para referirse a un asentamiento de la población del campo en entidades de una sola casa.

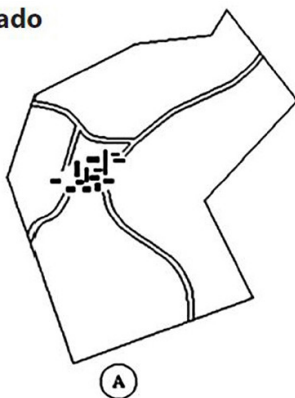
En el Atlas Nacional de España, se define el tipo de asentamiento rural, denominado mixto o intercalar, como una tipología intermedia entre el hábitat

concentrado y el disperso. En ella, al agrupamiento histórico de la población rural en aldeas se le une una dispersión o diseminación reciente de nuevas casas,, localizadas en las proximidades de una vía de comunicación, cerca de un paraje atractivo o en una finca amplia, que al generalizarse rompen con la tendencia a la agrupación del hábitat en núcleos mayores.

En las siguientes ilustraciones se muestran ejemplos de las características de distribución de los asentamientos humanos en el hábitat rural.

Ilustración 1

**(A) Concentrado**

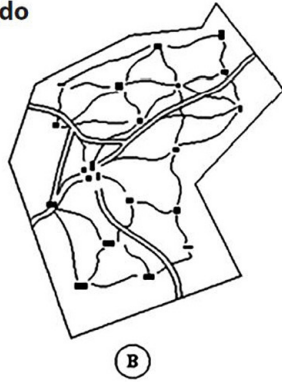


*Sasamón, Burgos*

Fuente: Atlas Nacional de España (2019)

Ilustración 2

**(B) Diseminado**



Vega de Pas, Cantabria

Fuente: Atlas Nacional de España (2019)

Ilustración 3

**(C) Intercalar**



Xove, Lugo

Fuente: Atlas Nacional de España (2019)

Otros investigadores se han acercado a través de la comparación de datos estadísticos de la densidad de población por unidades territoriales analizadas, comparándolas con la curva de Lorenz (mide la desigualdad de los datos).

Por otro lado, las teorías de localización explican la forma en que se distribuyen los asentamientos humanos y cómo ellos utilizan el espacio que los rodea y las interacciones que generan. Estas teorías son modelos simplificados de una realidad irregular y compleja. Las principales son las propuestas por los autores Von Thünen, Weber y Christaller.

En lo que respecta a la teoría de localización (lugares centrales) de Walter Christaller (1933 -1969), esta postula que la población se concentra en torno a los "lugares centrales", con el fin de acortar los tiempos de desplazamientos. De lo anterior se deduce el concepto de jerarquía, donde los servicios más básicos ocupan el primer lugar, luego los servicios más especializados.

Lo anterior se sustenta en la medición de la accesibilidad de las localidades hacia su municipio, como se enuncia en la primera ley de la geografía o principio de la autocorrelación espacial de Tobler (1979):

*"Todas las cosas están relacionadas entre sí, pero las cosas más próximas en el espacio tienen una relación mayor que las distantes."*

De los casos revisados, se tienen distintas formas de clasificación:

1. Interpretación de fotografías aéreas e imágenes satelitales, como es el caso Español.
2. Basada en supuestos de jerarquía, en función de la cantidad de población.
3. Basada en principios físicos de distribución aleatoria de partículas en sistema.
4. Basada en técnicas de desigualdad estadísticas de datos.

Todas las formas de clasificación revisadas tienen asociado el problema de definición de la superficie desde la cual se toman los datos. Otro problema, es que no asocian algún tipo de relación espacial y funcional entre las localidades y su centro administrativo local (municipio).

La propuesta de clasificación que se desarrollará en el presente estudio, busca dar respuesta a eliminar la arbitrariedad de la delimitación de la superficie e incorporar elementos de funcionalidad e interrelación espacial, entre donde se localiza el municipio y las localidades que este debe atender.

## II. OBJETIVOS

### 2.1 OBJETIVO GENERAL

Determinar grupos o clases de comunas, según su sistema de asentamientos humanos rurales en torno a su municipio, en base a los niveles de accesibilidad, costos de traslado y la concentración o dispersión de la población.

### 2.2 OBJETIVOS ESPECÍFICOS

- a) Calcular los niveles de accesibilidad física o ponderación espacial, de cada localidad a su municipio.
- b) Medición de concentración/dispersión, del sistema de localidades al municipio.
- c) Estimación de costos de traslado por concepto de combustible, desde el municipio a sus localidades.
- d) Análisis de conglomerados de comunas, mediante técnicas de estadística exploratoria.







### III. SÍNTESIS METODOLÓGICA

Este trabajo busca determinar una medida de concentración/dispersión de la población, en torno a sus sedes comunales, donde se encuentra localizado el municipio. Sin embargo, para entender la complejidad del accionar del municipio en torno a sus sistemas de asentamiento, es necesario definir los elementos.

La complejidad del accionar del municipio en torno a su sistema de asentamientos humanos, está referida a:

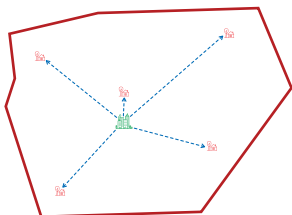
1. La jerarquía del municipio, cantidad de servicios que presta la localidad donde se encuentra el municipio.
2. Las relaciones que establecen los asentamientos con su municipio.
3. Morfología del sistema, patrones de distribución de población en el territorio.

Si se asume que todos los municipios tienen la misma jerarquía asociada a la cantidad de servicios, que se encuentran presentes en la localidad donde se localiza y todas las localidades de la comuna tienen la misma cantidad de población, pero con distinta distribución. Parece intuitivo determinar que la comuna que tenga menos localidades y que se encuentren más próximas al municipio, son más fáciles de intervenir en términos espaciales.

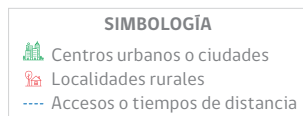
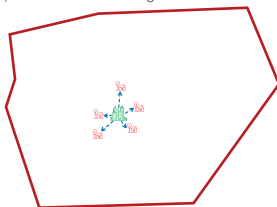
La Ilustración 4 muestra esquemáticamente qué comuna sería la menos compleja de intervenir, asumiendo que todas sus localidades tienen la misma población y sus municipios la misma jerarquía de servicios, pero con distintos patrones de distribución.

#### Ilustración 4. Esquema de distribución de localidades

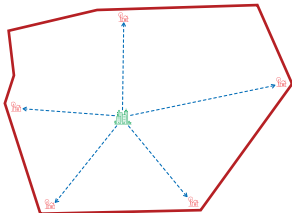
**Caso 1.** Todas las localidades tienen la misma población



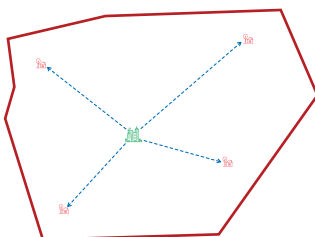
**Caso 2.** Todas las localidades tienen la misma población, pero cambió su configuración



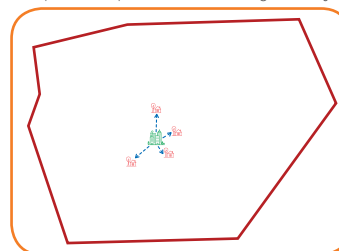
**Caso 3.** Todas las localidades tienen la misma población, pero cambió su configuración



**Caso 4.** Todas las localidades tienen la misma población, pero ahora solo hay 4



**Caso 5.** Todas las localidades tienen la misma población, pero cambió su configuración y son 4



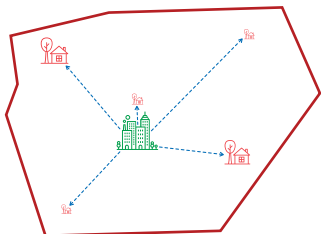
Fuente: Elaboración propia

Lo intuitivo del esquema anterior, se complejiza al incorporar el hecho de que no todas las localidades donde se encuentra el municipio tienen la misma cantidad de servicios, y que no todas las localidades tienen la misma cantidad de población.

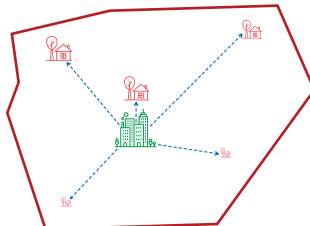
En la Ilustración 5, se muestra esquemáticamente qué comuna sería la menos compleja de intervenir, asumiendo que todas sus localidades tienen distinta población y sus municipios tienen distinta jerarquía de servicios, pero con distintos patrones de distribución. La comuna que sería más fácil de intervenir, es la que posea la menor cantidad de localidades, y que la mayor parte de su población se encuentre lo más próxima a su municipio.

### Ilustración 5. Esquema de distribución de localidades con distinta población

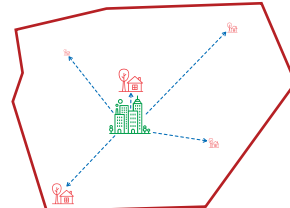
Caso 6. Todas las localidades tienen distinta población



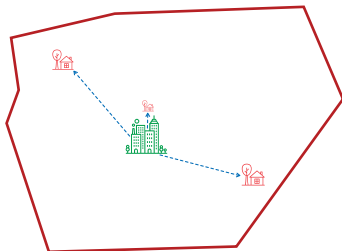
Caso 7. Todas las localidades tienen distinta población



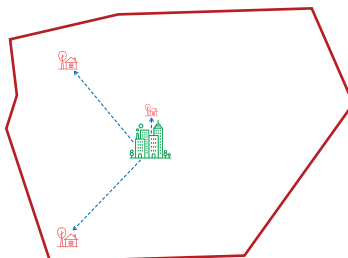
Caso 8. Todas las localidades tienen distinta población



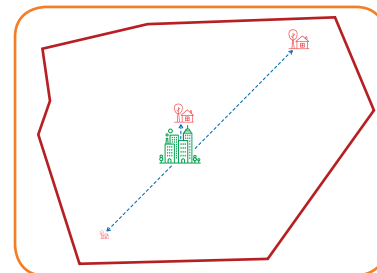
Caso 10. Todas las localidades tienen distinta población



Caso 11. Todas las localidades tienen distinta población

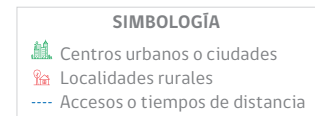


Caso 12. Todas las localidades tienen distinta población



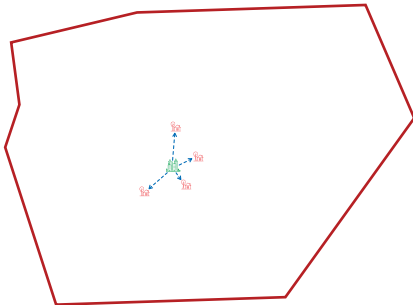
Fuente: Elaboración propia

En la Ilustración 6, se muestra esquemáticamente qué comuna sería la menos compleja de intervenir, con patrones de localización, distribución de población y jerarquía diferentes.

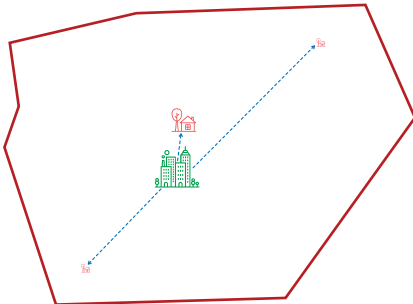


## Ilustración 6. Esquema de complejidad de sistemas




**Caso 5.** Todas las localidades tienen la misma población, pero cambió su configuración y son 4



**Caso 12.** Todas las localidades tienen distinta población



### SIMBOLOGÍA

-  Centros urbanos o ciudades
-  Localidades rurales
-  Accesos o tiempos de distancia

Ya no resulta tan intuitivo, pues depende de:

1. Cantidad de localidades conforman el sistema dentro la comuna.
2. Los habitantes que tienen cada una de las localidades.
3. El grado de accesibilidad desde el municipio a cada localidad.
4. Costo de traslado en transporte del personal a cada localidad.

**Fuente:** Elaboración propia

Los ejemplos presentados dan una noción de la complejidad de clasificar, pues se observa que las comunas tienen una distribución de la población distinta una de otra, tanto en cuanto a la cantidad de asentamientos, su número de habitantes, el grado de accesibilidad del municipio a cada una de las localidades y el costo de traslado que debe asumir un municipio.

La clasificación del caso español tiene un fuerte componente de fotointerpretación, donde queda a criterio del analista qué tipo de patrón de distribución es el que mejor representa la configuración de la localización de la población.

Actualmente, a partir del trabajo de actualización de Localidades en Condición de aislamiento, SUBDERE dispone de información georreferenciada de las viviendas rurales (provenientes del pre censo 2016), acceso a servicios de la localidad donde se encuentra el municipio, distancias y tiempos de desplazamiento de las viviendas del pre censo 2016 al municipio. Estas bases de datos permiten realizar análisis estadísticos de similitud de las variables por comuna, lo que a su vez permite clasificar tipos de comuna en función de datos y no de una interpretación visual realizada por un analista.

Para la clasificación de la complejidad del sistema de asentamientos humanos se considerarán las siguientes variables:

- a) Índice de integración de la localidad donde se encuentra el municipio (capital comunal). Que



corresponde al valor del índice de integración de la localidad donde se encuentra el municipio<sup>11</sup>.

- b) Indicador de concentración/dispersión de localidades rurales en torno a la localidad donde se localiza el municipio<sup>12</sup>.
- c) Costo en combustible para recorrer todas las localidades de una comuna desde su municipio. Corresponde al consumo de litros de combustible, por total de kilómetros que debe recorrer un vehículo

desde el municipio, utilizando un rendimiento de 12 km/lit y el valor promedio del diésel en la comuna según: Sistema de Información en Línea de Precios De Combustibles en Estaciones de Servicio<sup>13</sup>.

El universo del presente análisis corresponde a todas comunas de Chile que presentan una localidad clasificada por el INE como manzana entidad rural, según los resultados del Censo realizado el año 2017.

Las comunas de los territorios insulares de Rapa Nui, Juan Fernández y la Antártica Chilena, debido a sus condiciones de aislamiento serán incluidas dentro del rango de mayor complejidad.

11 El valor del índice corresponde a un cálculo preliminar, basado en Datos obtenidos de cálculos preliminares basados en levantamiento de viviendas de Pre Censo 2016 (INE) y metodología 2012 considerada en la actualización de variables del Estudio de "Identificación de Localidades de Condición de Aislamiento" (2018) SUBDERE. [http://geoportal.subdere.gov.cl/descargas\\_locais\\_2018](http://geoportal.subdere.gov.cl/descargas_locais_2018).

12 Ver en ANEXO Capitulo V.1.2 Indicador de Concentración/Dispersión.

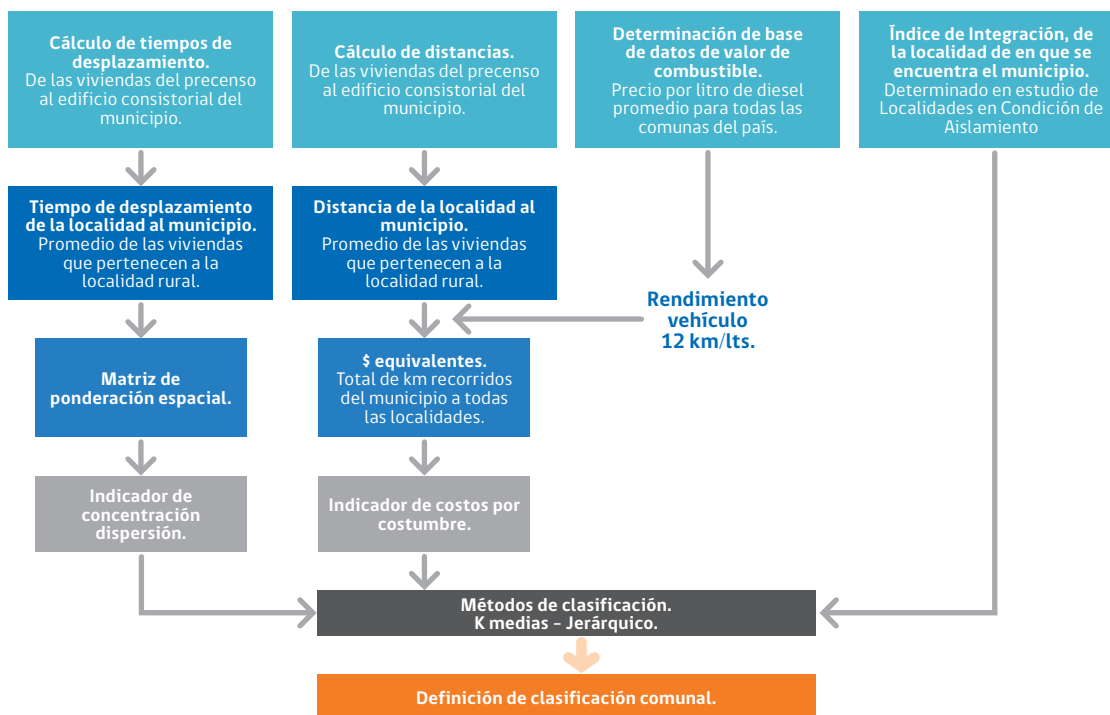
13 <http://www.bencinaenlinea.cl/web2/>

### 3.1 PROPUESTA METODOLÓGICA

La propuesta metodológica se desglosa esquemáticamente en la Ilustración 7, con cada uno de los elementos involucrados y su funcionalidad, para el logro de los ob-

jetivos planteados. En los capítulos 6.2 y 6.3 se detallarán aspectos teóricos de los cálculos involucrados métodos y técnicas de clasificación por conglomerados utilizadas.

Ilustración 7. Esquema metodológico



Fuente: Elaboración propia

### 3.1.1 Datos de entrada

Corresponde a los datos que son necesarios para poder realizar los indicadores que serán incorporados a los métodos de clasificación.

#### Cálculo de tiempos de desplazamiento

Esta información proviene de los cálculos realizados preliminarmente para la actualización de las condiciones de aislamiento, utilizando los datos censales (Censo abreviado 2017). El cálculo de los tiempos de desplazamiento (medido en horas) del municipio hacia una localidad rural, es el promedio de los tiempos de desplazamiento de cada una de las viviendas rurales levantadas en el pre censo del 2016.

#### Cálculo de distancias

Símil al cálculo de tiempos de desplazamiento, solo que se estiman kilómetros en vez de horas.

#### Base de datos de valor de combustible

Se debe construir una base de datos que contenga el valor del precio del diésel en cada una de las comunas. Para asignar el valor del precio por litro de diésel, se obtiene el valor promedio de todas las estaciones de servicio que se encuentran en la comuna. La información fue obtenida del sistema de información en línea de precios de combustibles en estaciones de servicio<sup>14</sup>.

14 <http://www.bencinaenlinea.cl/web2/>

### Índice de integración

Esta información proviene de los cálculos preliminares de las condiciones de aislamiento, basada en los tiempos de desplazamiento de la localidad donde se encuentra el municipio, hacia una serie de servicios propuestos en la metodología de Identificación de Localidades en Condición de Aislamiento (SUBDERE, 2012). La diferencia de estos resultados preliminares con los obtenidos en año 2012, se debe a que los tiempos de desplazamientos son medidos actualmente desde la localización del edificio consistorial donde funciona el municipio y no desde un punto representativo de la localidad (generalmente asignado a la plaza de armas o centro cívico)

### 3.1.2 Construcción de indicadores (variables del método de clasificación)

#### Indicador de concentración/dispersión

Es una síntesis de los niveles accesibilidad de las localidades hacia su sede comunal, entendiendo a esta como el lugar donde se ejerce el gobierno local y donde se localizan preferentemente los servicios básicos. Resulta evidente que, entre más lejanas de la cabecera comunal se encuentren las localidades, más difícil es atender las necesidades de su población.

El detalle de la construcción de este índice, a partir de la determinación de la matriz de accesibilidad general, se detallará en Capítulo 6.3.1.

### Indicador de costo por combustible

Este indicador es construido a partir del total de kilómetros que debe recorrer un vehículo estándar<sup>15</sup> a cada una de las localidades, en otras palabras, es la suma de los kilómetros desde el municipio a todas las localidades. El rendimiento del vehículo estándar es de 12 km/litros, con este supuesto, se pueden estimar los litros necesarios para recorrer, al menos una vez cada una de las localidades de la comuna.

Estimada la cantidad de diésel en litros por comuna, se puede determinar el costo en pesos (\$), utilizando el precio promedio de diésel por comuna.

### 3.1.3 Métodos de Clasificación

Las técnicas de clasificación no supervisada, llamadas también análisis de conglomerados, tienen como finalidad agrupar elementos en grupos homogéneos en función de "alguna similitud" entre ellos. En rigor, la conformación de un conglomerado se hace mediante una técnica de homogeneidad y reconocimiento de patrones.

Para un buen análisis de conglomerados es necesario tener en cuenta que debe ser:

- a) **Eficiente:** Utilizar tan pocos conglomerados como sea posible.
- b) **Efectivo:** Captura todos conglomerados estadística y teóricamente importantes.

La finalidad de la formación de conglomerados es las observaciones (comunales), se puedan clasificar y de esta forma encontrar grupos con "patrones de respuestas" similares.

Dentro de los métodos de clasificación exploratoria más usuales se cuentan: K medias y Jerárquico.

La clasificación de comunales permitirá identificar conjuntos de comunales, con características espaciales estructurales, que hacen difícil el despliegue territorial de un municipio con su sistema de localidades (colegios rurales, retiro de residuos sólidos domiciliarios, agua potable, manejo de emergencias, etc.).

15 Para este caso, se entiende como aquel que tiene un rendimiento de 12 kilómetros por litro de diésel.





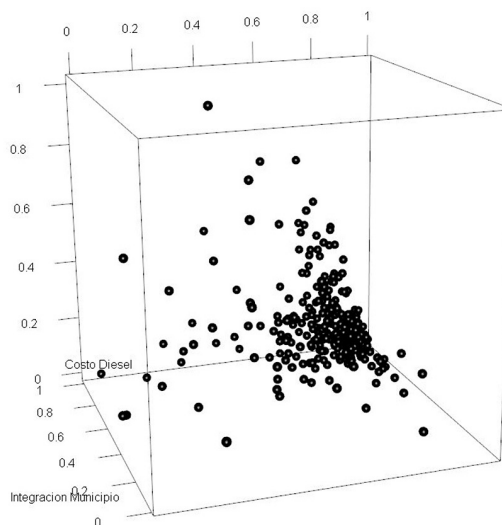
## IV. DESARROLLO

### 4.1 ANÁLISIS DE VARIABLES

En la Ilustración 8, se muestra la representación de las variables  $RHO_i$ <sup>16</sup>,  $INT_i$ <sup>17</sup>, Diésel<sup>18</sup>.

- Índice de Integración de la localidad donde se encuentra el municipio ( $INT_i$ ).
- Indicador de concentración/dispersión de localidades rurales en torno a la localidad donde se localiza el municipio ( $RHO_i$ )
- Costo en combustible, estandarizado entre 0 y 1. (Diésel).

Ilustración 8. Gráfico 3D Variables consideradas



Fuente: Elaboración propia

16 Indicador de concentración/dispersión de localidades rurales en torno a la localidad donde se localiza el municipio.

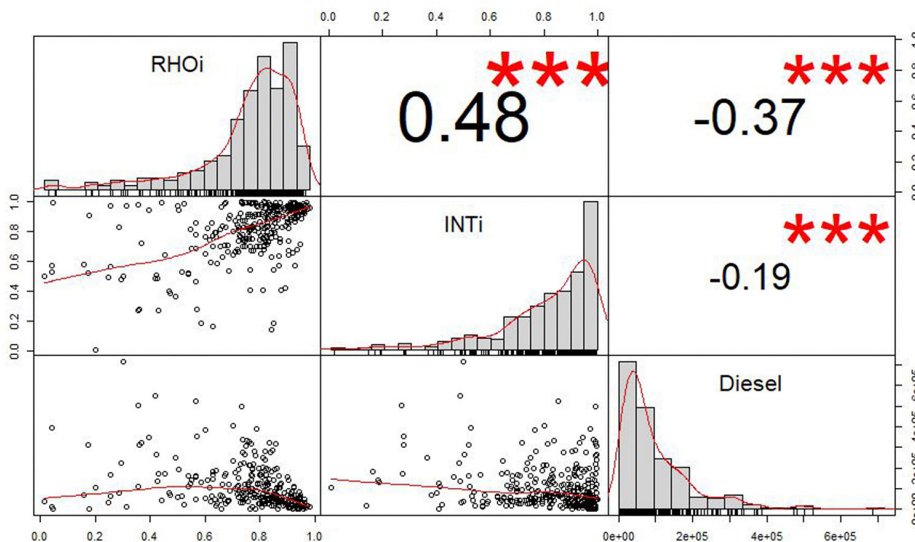
17 Índice de Integración de la localidad donde se encuentra el municipio.

18 Costo en combustible estandarizado entre 0 y 1.

Este tipo de ilustraciones permite observar si una variable incrementa o disminuye y cómo afecta a las demás. Los diagramas de dispersión muestran el grado de correlación (no ve causalidad) entre estas, sugiriendo si el tipo de relación es positiva, negativa o nula. En los diagramas se agrega una línea de tendencia que permite ver gráficamente el tipo de relación existente. Uno de los aspectos más relevantes de un gráfico de dispersión, es su capacidad para mostrar las relaciones no lineales entre las variables. Además, si los datos son representados por un modelo de mezcla de relaciones simples, estas relaciones son visualmente evidentes como patrones superpuestos.

El histograma permite obtener de manera general la distribución de cada una de las variables, respecto a las características cuantitativas. De esta manera, ofrece una visión de grupo, permitiendo observar una preferencia o tendencia. A partir de los histogramas se puede observar el grado de homogeneidad o el grado de variabilidad y, por ende, la dispersión de todos los valores que toman las partes. También es posible no evidenciar ninguna preferencia o tendencia, es decir, un comportamiento aleatorio.

**Ilustración 9.** Histogramas, gráficos de dispersión Correlaciones de las variables analizadas



Fuente: Elaboración propia

### El indicador de concentración/dispersión tiene:

- Una distribución asimétrica negativa (cola hacia la izquierda), muestra que los de valores tienden a concentrarse en torno a su municipio. Este gráfico da indicios de que la mayoría de los municipios tiene un nivel de servicios que los hace estar integrados, según los cálculos preliminares de la actualización de los datos censales para determinar las condiciones de aislamiento.
- Está correlacionado positivamente con el índice de integración del municipio. En el gráfico de dispersión se puede visualizar que ambas variables están concentradas en la esquina superior derecha, es decir, que la mayor parte de las comunas comparten valores cercanos a 1. Lo cual queda en evidencia en la forma similar que adopta el histograma de los índices de integración de las localidades donde se encuentra el municipio.
- Está correlacionado negativamente con el costo en diésel. En el gráfico de dispersión se puede visualizar que ambas variables están concentradas en la esquina inferior derecha. Esto quiere decir que la mayor parte de las comunas con valores próximos a 1 en la integración de la localidad donde se encuentra el municipio, obtiene valores cercanos a 0 en el costo total de combustible, lo cual es consistente ya que son menores los traslados para acudir a cada una de localidades,

es decir están más próximas a su municipio. Lo último queda en evidencia en la forma similar que adopta el histograma de los índices de integración de las localidades donde se encuentra el municipio.

### El índice de integración de la localidad donde se encuentra el municipio tiene:

- Una distribución asimétrica negativa (cola hacia la izquierda), muestra que la concentración de valores tiende a valores de concentración de asentamientos rurales en torno a su municipio. Este gráfico da indicios de que en general, la población rural de las comunas de Chile se encuentra próxima a su municipio. No obstante, se aprecia que existen datos anómalos hacia la cola izquierda, por lo se está en presencia de que existen comunas que tienen un indicador con valores próximos a 0, lo que significa que su sistema de asentamientos rurales se encuentra disperso en relación a la localización de su municipio.
- Está correlacionado negativamente con el índice de diésel, en el gráfico de dispersión se puede visualizar que ambas variables esta concentradas en la esquina inferior izquierda, es decir, que la mayor parte de las comunas con valores próximos a 1, obtiene valores cercanos a 0 en el costo total de combustible.

### El indicador de costo de diésel tiene:

- Una distribución asimétrica positiva (cola hacia la derecha), muestra que la concentración de valores tiende a valores de menor costo de traslados medido en diésel. Este gráfico da indicio de que no son muchos los municipios que deben incurrir en grandes costos de diésel para atender a su sistema de asentamientos humanos.

## 4.2 ELECCIÓN DEL MÉTODO DE CLASIFICACIÓN

Se desarrollaron los métodos a partir de los algoritmos de k - medias y de análisis jerárquico<sup>19</sup>.

Ambos métodos parecen consistentes, en cuanto a que la media de sus variables difiere de grupo a grupo. Para el método de k - medias, el algoritmo arroja que son 6 los grupos homogéneos, en cambio en el método jerárquico son 7. La Tabla 1, muestra los resultados de la clasificación para ambos métodos.

**Tabla 1.** Resumen de casos clasificados

Clases k medias	Clases Análisis Jerárquico							Total Clase K
	1	2	3	4	5	6	7	
1	0	0	0	15	0	3	0	18
2	9	0	0	1	0	5	0	15
3	0	0	0	0	1	9	0	10
4	0	111	18	0	0	0	0	129
5	0	0	30	0	17	0	0	47
6	0	2	31	12	0	0	52	97
Total clase J	9	113	79	28	18	17	52	316 <sup>20</sup>

Fuente: Elaboración propia

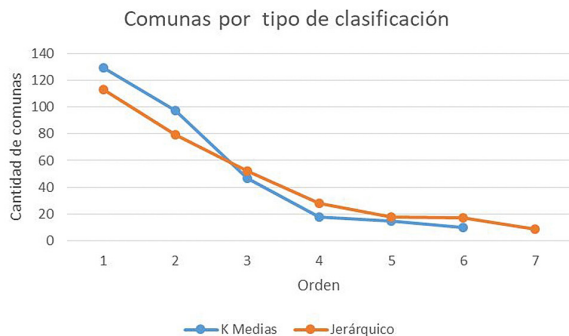
Al tratarse de técnicas de análisis exploratorio de datos, se debe tener en cuenta el sentido de la utilidad de la clasificación. Al ordenar de mayor a menor la cantidad de comunas Gráfico 1, se observa que el método jerárquico tiene una curva con menor pendiente, por lo que los cambios de cantidad de comunas no son tan abruptos.

19 Para ver el detalle del desarrollo del método véase el Capítulo 6.2.

20 Son 316 comunas que cuentan con levantamiento de viviendas del pre - censo 2016.



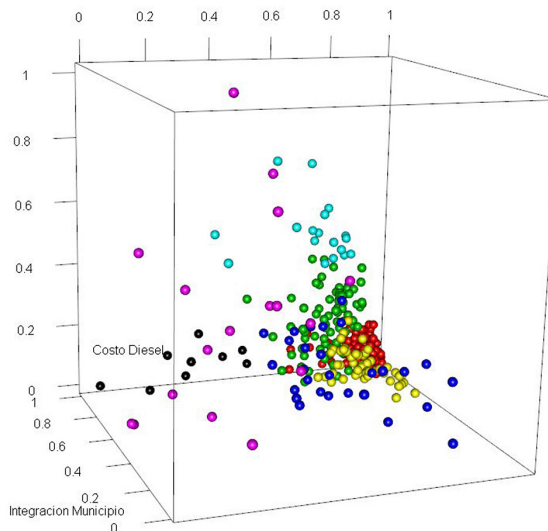
**Gráfico 1.** Decrecimiento por tipo de clasificación



Fuente: Elaboración propia

El método jerárquico con 7 clases, es el que se ocupará para la clasificación. En la Ilustración 10 se puede apreciar el resultado del algoritmo de análisis jerárquico, donde cada color representa una clase o grupo.

**Ilustración 10.** Gráfico 3D de variables y clasificación jerárquica



Fuente: Elaboración propia

### 4.3 COMUNAS SEGÚN CLASIFICACIÓN

Debido a que el método asigna una etiqueta (clase) aleatoriamente, se deben describir los grupos encontrados. Para determinar un “orden” se procede de la siguiente forma:

1. Se promedian, los promedios del índice de concentración/dispersión, con el índice de integración municipal.
2. Al promedio obtenido en 2 se resta el índice del costo del diésel.
3. El valor obtenido en 2 se ordena de mayor a menor.

En la Tabla 2, se expone la clasificación propuesta según las clases arrojadas por el método de agrupamiento jerárquico.

**Tabla 2.** Clasificación según clases del método de agrupamiento jerárquico

Clase	Promedios			Clasificación según complejidad
	Concentración dispersión	Integración sede municipal	Diésel	
2	0,8980	0,9345	0,0561	Muy Baja
7	0,8088	0,7114	0,0822	Baja
3	0,7626	0,8788	0,1935	Transición a Baja
1	0,3230	0,9560	0,0600	Media
4	0,6446	0,4961	0,1515	Transición a alta
5	0,7116	0,9032	0,4844	Alta
6	0,2821	0,4233	0,3672	Muy Alta

Fuente: Elaboración propia



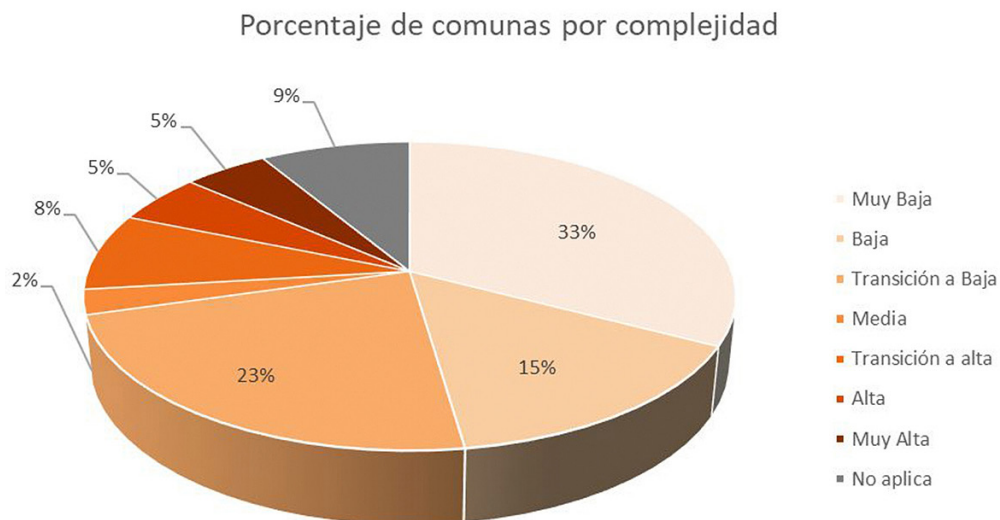
**Tabla 3.** Descripción de clases

Clasificación	Descripción
Muy Baja	Población rural concentrada en torno a su sede municipal, sede municipal con altos niveles de integración de servicios y costos en diésel bajos.
Baja	Población rural concentrada en torno a su sede municipal, sede municipal con niveles de integración de servicios medios y costo en diésel bajos.
Transición a Baja	Población rural semi concentrada en torno a su sede municipal, sede municipal con buenos niveles de integración de servicios y costos en diésel moderados.
Media	Población rural dispersa en torno a su sede municipal, sede municipal con buenos niveles de integración de servicios y costos en diésel bajos.
Transición a alta	Población rural semi dispersa, sede municipal con bajos niveles de integración y costos en diésel moderados.
Alta	Población rural semi dispersa, sede municipal con altos niveles de integración y costos en diésel altos.
Muy Alta	Población rural dispersa en torno a su sede municipal, sede municipal con bajos niveles de integración de servicios y costos en diésel altos.

**Fuente:** Elaboración propia

En la Tabla 3, se describen las características de las clasificaciones propuestas, en función de las variables ingresadas al método.

Gráfico 2. Distribución a nivel nacional

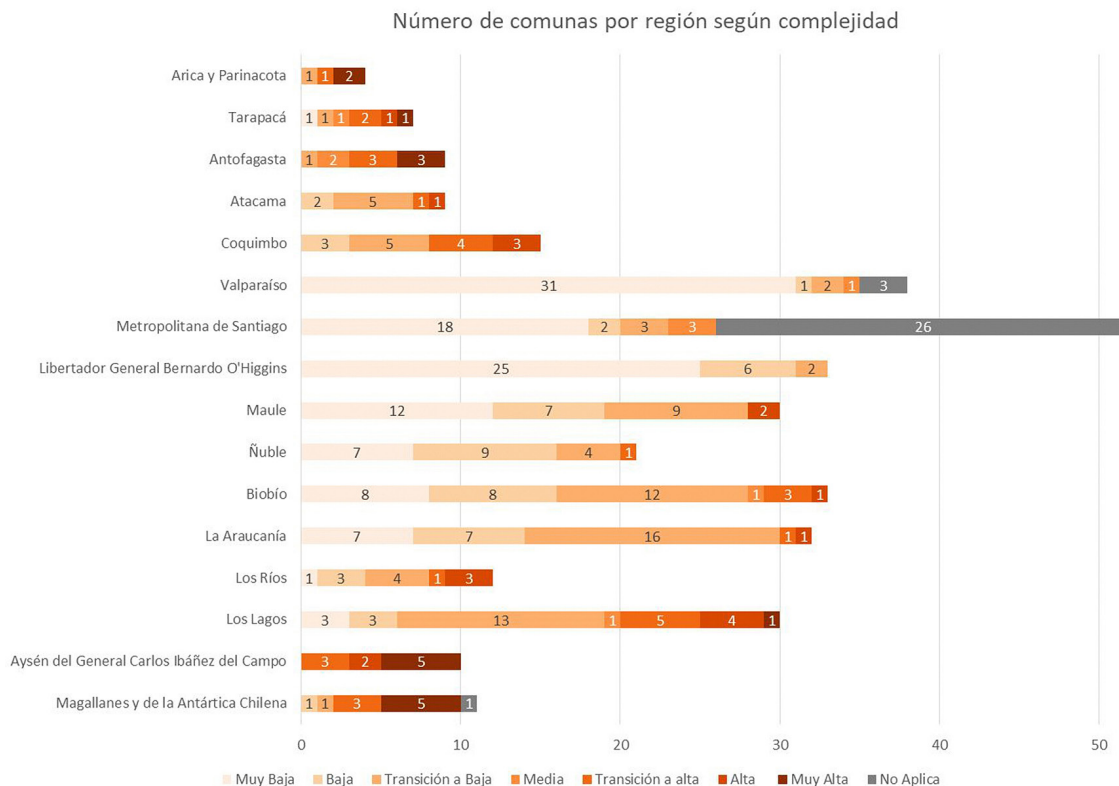


Fuente: Elaboración propia

En el Gráfico 2, se puede apreciar que el 48% de las comunas del país está clasificada entre los tipos baja y muy baja. Las comunas clasificadas como alta y muy alta equivalen al 14% del total de las comunas.

El Gráfico 3 muestra en términos absolutos, la cantidad de comunas por región según su clasificación de complejidad, siendo las regiones de Antofagasta, Los Lagos, Aysén y Magallanes, las que cuentan con un número mayor de comunas clasificadas de complejidad muy alta.

**Gráfico 3.** Distribución complejidad a nivel regional



Fuente: Elaboración propia

### 4.3.1 Resultados de Clasificación

A nivel regional se tiene que:

- La clasificación muy alta, está mayormente en las regiones australes y en menor medida en Antofagasta y Arica y Parinacota.
- Las regiones de Valparaíso, O'Higgins y Maule obtienen la mayor cantidad de comunas de Muy Baja complejidad, debido a que son en su mayoría comunas de menor superficie en Km<sup>2</sup>, con buena red de camino y una red vial densificada. En estas regiones se encuentran la mayoría de las ciudades intermedias en torno a la Ruta 5 y ciudades costeras.
- La Región de Los Lagos posee todas las posibles clasificaciones de comunas.
- La Región de Aysén es la que presenta las condiciones de transición de alta a muy alta, lo que la sitúa como la que presenta las mayores complejidades desde sus sistemas de asentamientos humanos.
- Las regiones de Valparaíso (1) y Metropolitana (26), tienen comunas que no tienen localidades rurales por lo que no aplica esta clasificación, lo que equivale al 9% de las comunas del país.
- En los mapas se puede apreciar que de la complejidad de los asentamientos humanos de la comuna, coincide como era de esperar con aquellas comunas que tienen mayor superficie, debido a que tienen que recorrer una distancia mucho mayor desde el municipio a ellas. Se encuentran preferentemente en los extremos norte y austral. Además, tienen los menos niveles de integración, según la actualización de variables del estudio de Identificación de Localidades en Condición de Aislamiento, realizado en el año 2018.
- Las comunas de la zona central de Chile, tienen mayor complejidad en los sectores cordilleranos, tanto de Los Andes como de la cordillera de la Costa (secano costero); en especial en la cordillera de Nahuelbuta.
- En cuanto a la Región de Arica y Parinacota, se puede apreciar que en general es compleja dada la dispersión de la población y la alta concentración de servicios e integración de la ciudad de Arica. La población rural se localiza preferentemente en los valles y quebradas.



## V. CONCLUSIONES

- La metodología propuesta permite clasificar a las comunas utilizando patrones de comportamiento de datos, lo cual permite reducir los juicios del especialista de fotointerpretación.
- Al no ser un índice compuesto o multifactorial, permite analizar el comportamiento de las variables por separado en cada grupo.
- Es una alternativa a los indicadores de densidad. La ventaja es que se incluyen los tiempos de desplazamiento entre las localidades rurales y la localidad donde se encuentra el municipio. Esto se hace tomando en cuenta la distribución y conectividad.
- Al no buscar determinar un ranking, se habla de grupos que tienen similitudes de comportamiento.
- La jerarquía del municipio puede precisarse más aún, ya que el índice de integración, proveniente del estudio de Identificación de Localidades en Condición de Aislamiento, tiene incorporado el acceso a capital comunal con una alta ponderación dentro de los criterios que definen dicho índice. A partir de los datos levantados de los talleres regionales, para la actualización de la metodología que determina las condiciones de aislamiento, se puede generar un mejor indicador que dé cuenta de la jerarquía en función de servicios de cómo los de comercio mayorista y minorista, financieros, actividades culturales, etc.
- El indicador de concentración / dispersión propuesto, permite sintetizar las relaciones funcionales del municipio con los asentamientos humanos, incorporando la población de cada uno de ellos.
- Este tipo de clasificación permite agrupar comunas con similar comportamiento. Esto permitirá que se puedan generar distintas estrategias de intervención en su territorio.

## VI. ANEXOS

### 6.1 LISTADO DE COMUNAS SEGÚN CLASIFICACIÓN

En las siguientes tablas se desglosa las comunas que perteneces a cada clasificación. Se incluye cartografía de una comuna ejemplificando cada una de las clasificaciones.

**Tabla 4.** Clasificación Muy Baja

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Tarapacá	Alto Hospicio	11	2491	88	0,9103	0,9643	0,0072
Valparaíso	Concón	24	2743	1062	0,9034	0,9585	0,0143
Valparaíso	Puchuncaví	46	2687	1920	0,9400	0,8995	0,0270
Valparaíso	Quintero	67	5039	2490	0,8679	0,9115	0,0785
Valparaíso	Los Andes	44	5691	1531	0,8146	0,9991	0,0686
Valparaíso	Calle Larga	60	4141	1563	0,9717	0,9923	0,0252
Valparaíso	Rinconada	28	2142	832	0,9673	0,9772	0,0094
Valparaíso	San Esteban	97	7483	3272	0,9225	0,9925	0,0699
Valparaíso	La Ligua	70	9381	7273	0,8496	0,9274	0,0956
Valparaíso	Cabildo	75	7215	2753	0,9010	0,8664	0,0941
Valparaíso	Papudo	11	942	430	0,9032	0,8201	0,0099
Valparaíso	Quillota	112	12186	4524	0,9260	0,9969	0,0683
Valparaíso	Calera	11	1886	692	0,9598	0,9860	0,0053
Valparaíso	Hijuelas	44	6253	2568	0,9351	0,9684	0,0305
Valparaíso	La Cruz	36	2646	980	0,9587	0,9849	0,0185

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Valparaíso	Nogales	36	3341	1373	0,9450	0,9741	0,0177
Valparaíso	San Antonio	45	4781	2432	0,8288	0,9938	0,0582
Valparaíso	Algarrobo	48	2918	3044	0,8949	0,8966	0,0408
Valparaíso	Cartagena	28	1892	824	0,9101	0,9746	0,0309
Valparaíso	El Quisco	29	1088	984	0,8701	0,9138	0,0243
Valparaíso	El Tabo	14	671	1060	0,9435	0,9435	0,0084
Valparaíso	Santo Domingo	38	4684	3443	0,6714	0,9845	0,0665
Valparaíso	San Felipe	71	7227	2562	0,9500	0,9745	0,0413
Valparaíso	Catemu	40	6185	2514	0,9519	0,8846	0,0233
Valparaíso	Llaillay	41	6636	2552	0,9159	0,8940	0,0238
Valparaíso	Panquehue	29	3467	1199	0,9349	0,9303	0,0127
Valparaíso	Putendo	67	10020	3873	0,9525	0,9620	0,0449
Valparaíso	Santa María	46	5453	2107	0,9718	0,9741	0,0204
Valparaíso	Quilpué	34	2112	1427	0,7369	0,9841	0,0676
Valparaíso	Limache	63	6952	2804	0,9043	0,9700	0,0458
Valparaíso	Olmué	92	5239	3231	0,8768	0,9529	0,0810
Valparaíso	Villa Alemana	13	1221	539	0,9419	0,9793	0,0056
Metropolitana de Santiago	Huechuraba	2	99	31	0,9374	0,9813	0,0000
Metropolitana de Santiago	Maipú	25	3433	965	0,9323	0,9605	0,0148
Metropolitana de Santiago	Pudahuel	32	4155	1327	0,8556	0,9706	0,0371
Metropolitana de Santiago	Quilicura	7	552	192	0,9192	0,9860	0,0027
Metropolitana de Santiago	Pirque	122	14851	4778	0,9308	0,9263	0,0705
Metropolitana de Santiago	Colina	161	28189	7749	0,8888	0,9606	0,1137



Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Metropolitana de Santiago	Lampa	160	21273	6972	0,9020	0,9429	0,1278
Metropolitana de Santiago	Tiltil	70	6242	2593	0,8266	0,8351	0,0871
Metropolitana de Santiago	San Bernardo	45	5065	1606	0,9269	0,9842	0,0269
Metropolitana de Santiago	Buín	142	13703	4231	0,9473	0,9789	0,0992
Metropolitana de Santiago	Calera de Tango	160	13847	4076	0,9435	0,9635	0,0738
Metropolitana de Santiago	Curacaví	79	12219	5105	0,8282	0,8914	0,0739
Metropolitana de Santiago	María Pinto	74	6923	2591	0,9301	0,8883	0,0674
Metropolitana de Santiago	Talagante	171	15028	4876	0,9123	0,9776	0,0881
Metropolitana de Santiago	El Monte	73	5878	2091	0,9017	0,9760	0,0447
Metropolitana de Santiago	Isla de Maipo	114	9278	3254	0,9131	0,9407	0,0711
Metropolitana de Santiago	Padre Hurtado	65	7522	2385	0,8767	0,9622	0,0537
Metropolitana de Santiago	Peñaflor	81	7137	2193	0,9478	0,9695	0,0272
Lib. Gral Bernardo O'Higgins	Rancagua	85	7591	2051	0,9209	0,9903	0,0736
Lib. Gral Bernardo O'Higgins	Codegua	71	5969	2038	0,9627	0,9379	0,0453
Lib. Gral Bernardo O'Higgins	Coínco	33	3935	1563	0,9561	0,9180	0,0142
Lib. Gral Bernardo O'Higgins	Coltauco	76	8386	3493	0,9491	0,8895	0,0536
Lib. Gral Bernardo O'Higgins	Doñihue	39	4463	1804	0,9628	0,9214	0,0198
Lib. Gral Bernardo O'Higgins	Graneros	52	3616	1222	0,9751	0,9521	0,0213

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Lib. Gral Bernardo O'Higgins	Machalí	19	1001	560	0,8065	0,9781	0,0356
Lib. Gral Bernardo O'Higgins	Malloa	57	7949	3160	0,9619	0,9509	0,0366
Lib. Gral Bernardo O'Higgins	Mostazal	55	4389	1876	0,9291	0,9207	0,0416
Lib. Gral Bernardo O'Higgins	Olivar	42	4320	1447	0,9831	0,9608	0,0169
Lib. Gral Bernardo O'Higgins	Peumo	28	2313	829	0,9503	0,9330	0,0206
Lib. Gral Bernardo O'Higgins	Pichidegua	66	10411	3956	0,9165	0,8804	0,0612
Lib. Gral Bernardo O'Higgins	Quinta de Tilcoco	47	5724	2137	0,9572	0,9287	0,0225
Lib. Gral Bernardo O'Higgins	Rengo	116	12879	4713	0,9471	0,9751	0,0800
Lib. Gral Bernardo O'Higgins	Requínoa	116	11685	4014	0,9624	0,9542	0,0702
Lib. Gral Bernardo O'Higgins	San Vicente	167	20557	8105	0,9442	0,9670	0,1198
Lib. Gral Bernardo O'Higgins	Marchihue	55	3987	1919	0,8820	0,8271	0,0580
Lib. Gral Bernardo O'Higgins	Chépica	80	6391	2482	0,9357	0,9314	0,0643
Lib. Gral Bernardo O'Higgins	Chimbarongo	155	14913	5494	0,9215	0,9282	0,1473
Lib. Gral Bernardo O'Higgins	Lolol	81	4382	1968	0,8110	0,8235	0,0795
Lib. Gral Bernardo O'Higgins	Nancagua	61	6493	2464	0,9572	0,9457	0,0330
Lib. Gral Bernardo O'Higgins	Palmita	77	9018	3321	0,8605	0,9463	0,0871
Lib. Gral Bernardo O'Higgins	Peralillo	58	4370	1886	0,9262	0,8900	0,0403
Lib. Gral Bernardo O'Higgins	Placilla	51	6363	2336	0,9574	0,9411	0,0263

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Lib. Gral Bernardo O'Higgins	Santa Cruz	119	15031	5931	0,9434	0,9596	0,0834
Maule	Talca	106	9441	3432	0,9101	0,9941	0,0847
Maule	Maule	115	10887	4207	0,9229	0,9620	0,0932
Maule	Pelarco	53	5835	2362	0,9390	0,9087	0,0426
Maule	Río Claro	114	9493	3912	0,9021	0,8484	0,1239
Maule	San Rafael	48	4223	1811	0,9426	0,9269	0,0271
Maule	Molina	119	8658	3766	0,9173	0,9610	0,1208
Maule	Rauco	64	4953	2131	0,8879	0,9432	0,0593
Maule	Romeral	84	8658	3475	0,9351	0,8996	0,0651
Maule	Sagrada Familia	109	10658	3938	0,9325	0,9222	0,0993
Maule	Colbún	135	11713	5300	0,9227	0,8967	0,1558
Maule	Villa Alegre	81	6545	2769	0,9341	0,9291	0,0578
Maule	Yerbas Buenas	116	11894	4516	0,9105	0,9636	0,1020
Ñuble	Bulnes	114	8002	3641	0,8572	0,8852	0,1003
Ñuble	Chillán Viejo	60	3498	1559	0,8324	0,9602	0,0605
Ñuble	Pinto	71	5369	2897	0,8268	0,8610	0,0930
Ñuble	San Ignacio	111	9277	3978	0,8601	0,8263	0,0964
Ñuble	Coihueco	154	17783	6888	0,8259	0,8644	0,1607
Ñuble	Ñiquén	89	9999	4319	0,8524	0,8546	0,0967
Ñuble	San Nicolás	107	6716	2943	0,8505	0,8769	0,0998
Biobío	Concepción	67	4517	1711	0,8570	0,9930	0,0806
Biobío	Chiguayante	4	75	32	0,6610	0,9647	0,0037
Biobío	Lota	4	108	108	0,8634	0,9719	0,0028
Biobío	Penco	19	467	231	0,8069	0,9637	0,0159
Biobío	San Pedro de la Paz	3	124	45	0,9368	0,9772	0,0018

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Biobío	Talcahuano	4	1429	133	0,7843	0,9881	0,0037
Biobío	Hualpén	10	788	248	0,9075	0,9633	0,0085
Biobío	Negrete	33	3682	1416	0,8985	0,8596	0,0215
La Araucanía	Gorbea	76	4254	2106	0,8447	0,8344	0,0828
La Araucanía	Perquenco	59	3341	1446	0,8643	0,8824	0,0466
La Araucanía	Pucón	83	10169	6829	0,8176	0,8955	0,1256
La Araucanía	Cholchol	100	7255	2782	0,8250	0,8678	0,0778
La Araucanía	Angol	81	4351	1984	0,7329	0,9982	0,1114
La Araucanía	Ercilla	62	4093	1373	0,8299	0,9170	0,0561
La Araucanía	Renaico	31	1898	713	0,8847	0,8983	0,0235
Los Ríos	Máfil	62	2852	1203	0,8926	0,8376	0,0569
Los Lagos	Castro	77	9763	4776	0,7904	0,9682	0,0868
Los Lagos	Chonchi	79	9144	4545	0,8537	0,8867	0,1169
Los Lagos	Curaco de Vélez	26	2723	1215	0,9262	0,8676	0,0113

Fuente: Elaboración propia

Tabla 5. Clasificación Baja

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Atacama	Caldera	42	1887	8310	0,7270	0,7740	0,1149
Atacama	Chañaral	47	1136	1901	0,7632	0,6987	0,1183
Coquimbo	Paiguano	49	4493	2369	0,9034	0,6441	0,0590
Coquimbo	Illapel	105	9796	4429	0,7618	0,7642	0,1805
Coquimbo	Salamanca	95	12544	4653	0,7460	0,6985	0,1387

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Valparaíso	Petorca	61	5723	2637	0,7490	0,7515	0,0774
Metropolitana de Santiago	Alhué	28	3660	1558	0,9254	0,5720	0,0230
Metropolitana de Santiago	San Pedro	101	9669	4168	0,8532	0,7967	0,0935
Lib. Gral Bernardo O'Higgins	Pichilemu	57	3491	2753	0,8428	0,6285	0,0797
Lib. Gral Bernardo O'Higgins	La Estrella	50	1908	1358	0,8134	0,7054	0,0396
Lib. Gral Bernardo O'Higgins	Litueche	49	2926	2051	0,8158	0,6999	0,0593
Lib. Gral Bernardo O'Higgins	Navidad	73	6635	5522	0,9050	0,6755	0,0600
Lib. Gral Bernardo O'Higgins	Paredones	77	4346	3024	0,8039	0,6521	0,0813
Lib. Gral Bernardo O'Higgins	Pumanque	50	3421	1670	0,9455	0,7134	0,0353
Maule	Curepto	98	6059	3072	0,7865	0,6995	0,1232
Maule	Empedrado	47	1134	534	0,8008	0,7470	0,0392
Maule	Chanco	88	5007	2095	0,8589	0,7643	0,0766
Maule	Pelluhue	45	3490	2290	0,8744	0,7524	0,0461
Maule	Hualañé	47	3914	1903	0,8732	0,7275	0,0589
Maule	Licantén	24	1688	775	0,9356	0,6780	0,0216
Maule	Vichuquén	47	3132	2901	0,8699	0,6353	0,0486
Ñuble	Quillón	77	7062	5256	0,8011	0,7706	0,0734
Ñuble	Yungay	70	4484	2623	0,8420	0,7548	0,0730
Ñuble	Quirihue	88	2121	1147	0,7764	0,6750	0,0814
Ñuble	Coelemu	68	4560	2164	0,8308	0,7273	0,0563
Ñuble	Ninhue	51	3714	1684	0,8314	0,7392	0,0395
Ñuble	Portezuelo	61	3058	1343	0,7309	0,7651	0,0472

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Ñuble	Ranquil	75	4168	2151	0,9024	0,7333	0,0604
Ñuble	Treguaco	61	3627	1788	0,7383	0,6980	0,0481
Ñuble	San Fabián	39	2614	1402	0,8142	0,6819	0,0558
Biobío	Santa Juana	93	4155	2630	0,7362	0,7787	0,1223
Biobío	Cañete	148	15043	5730	0,7435	0,7385	0,1712
Biobío	Contulmo	46	2970	1466	0,6827	0,7109	0,0575
Biobío	Los Álamos	33	2309	907	0,8200	0,7455	0,0416
Biobío	Antuco	30	1996	1147	0,9188	0,6584	0,0242
Biobío	Cabrero	91	7009	3222	0,8500	0,8056	0,0840
Biobío	Quilleco	47	4001	1786	0,8269	0,7617	0,0594
Biobío	Tucapel	52	2247	1299	0,8424	0,7939	0,0546
La Araucanía	Curarrehue	54	5213	2349	0,7370	0,7458	0,0809
La Araucanía	Melipeuco	57	3331	1629	0,8697	0,5720	0,0520
La Araucanía	Saavedra	110	9721	4445	0,7174	0,6769	0,1504
La Araucanía	Teodoro Schmidt	134	9642	4197	0,7994	0,7007	0,1756
La Araucanía	Toltén	92	5854	2545	0,7683	0,6751	0,1357
La Araucanía	Curacautín	104	4575	2719	0,8264	0,7175	0,1330
La Araucanía	Purén	72	4246	1699	0,8430	0,7597	0,0581
Los Ríos	Corral	37	1821	1124	0,6842	0,6625	0,0556
Los Ríos	Lanco	76	5145	2119	0,8123	0,7568	0,1045
Los Ríos	Futrono	98	6482	3231	0,7219	0,6631	0,1580
Los Lagos	Puqueldón	43	3921	1641	0,8492	0,5432	0,0254
Los Lagos	Queilén	48	3062	1493	0,6548	0,6988	0,0625
Los Lagos	Quellón	92	9378	4244	0,7620	0,7532	0,1322
Magallanes y de la Antártica Chilena	Natales	73	2297	1324	0,7716	0,7514	0,2236

Fuente: Elaboración propia

**Tabla 6.** Transición a baja

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Arica y Parinacota	Arica	85	16285	6242	0,7687	0,9892	0,1987
Tarapacá	Pozo Almonte	66	5616	5559	0,5467	0,8139	0,1723
Antofagasta	Mejillones	9	513	576	0,6388	0,7684	0,0263
Atacama	Tierra Amarilla	79	4084	1489	0,5043	0,9417	0,2658
Atacama	Vallenar	80	5898	3186	0,6903	0,9079	0,1262
Atacama	Alto del Carmen	105	5281	2919	0,7212	0,7336	0,2664
Atacama	Freirina	86	2452	1753	0,6619	0,8399	0,1703
Atacama	Huasco	30	1247	2460	0,5667	0,7812	0,0674
Coquimbo	La Serena	214	20414	8677	0,7432	0,9951	0,3701
Coquimbo	Coquimbo	106	13180	7436	0,6732	0,9931	0,2420
Coquimbo	Andacollo	35	1036	699	0,6747	0,8285	0,0387
Coquimbo	Vicuña	173	10778	4577	0,8204	0,7716	0,3893
Coquimbo	Punitaqui	111	5108	2383	0,7215	0,8951	0,2100
Valparaíso	Casablanca	142	8865	5350	0,8299	0,9265	0,1674
Valparaíso	Zapallar	48	2326	1800	0,7196	0,8086	0,0841
Metropolitana de Santiago	San José de Maipo	57	6981	2893	0,7833	0,9137	0,0832
Metropolitana de Santiago	Paine	176	26006	9908	0,9025	0,9643	0,1992
Metropolitana de Santiago	Melipilla	250	38903	14520	0,8179	0,9953	0,3374
Lib. Gral Bernardo O'Higgins	Las Cabras	146	15062	8297	0,7979	0,8645	0,2310
Lib. Gral Bernardo O'Higgins	San Fernando	116	10261	4462	0,8745	0,9909	0,1451
Maule	Constitución	138	8795	3627	0,8071	0,9150	0,2490
Maule	Pencahue	115	4807	2301	0,8626	0,9351	0,1385

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Maule	Cauquenes	187	7327	4150	0,7601	0,7982	0,2961
Maule	Curicó	212	16567	6474	0,7825	0,9907	0,3665
Maule	Teno	186	19389	7367	0,8989	0,9383	0,2181
Maule	Linares	154	15930	7363	0,8745	0,9947	0,1673
Maule	Longaví	176	21287	8543	0,8429	0,9503	0,2546
Maule	Parral	163	10843	5333	0,8208	0,9121	0,2365
Maule	Retiro	116	12911	5441	0,8621	0,9236	0,1462
Ñuble	Chillán	115	16092	6502	0,8387	0,9817	0,1268
Ñuble	El Carmen	82	7183	3022	0,6538	0,8174	0,0973
Ñuble	Pemuco	57	4280	1889	0,6364	0,7949	0,0680
Ñuble	San Carlos	245	19763	8207	0,8522	0,9561	0,2710
Biobío	Florida	119	6064	3416	0,7607	0,8149	0,1325
Biobío	Hualqui	83	3444	2020	0,7649	0,9186	0,0976
Biobío	Tomé	123	5406	3068	0,8189	0,9547	0,1594
Biobío	Arauco	113	8937	3718	0,7461	0,8751	0,1667
Biobío	Curanilahue	28	2114	1001	0,7967	0,7948	0,0395
Biobío	Laja	71	6167	2734	0,7395	0,8324	0,0871
Biobío	Mulchén	105	5478	2419	0,7095	0,8725	0,1419
Biobío	Nacimiento	61	3284	1937	0,6262	0,8561	0,0915
Biobío	Quilaco	39	2118	1175	0,7555	0,8235	0,0547
Biobío	San Rosendo	18	505	370	0,7705	0,8223	0,0149
Biobío	Santa Bárbara	67	5787	2490	0,7172	0,8421	0,1058
Biobío	Yumbel	157	8487	5346	0,8256	0,8019	0,1849
La Araucanía	Temuco	208	19250	7020	0,8011	0,9798	0,2308
La Araucanía	Carahue	155	11280	4646	0,6393	0,7939	0,2649



Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
La Araucanía	Cunco	181	8679	4576	0,7428	0,7400	0,3115
La Araucanía	Freire	208	16720	6587	0,7963	0,9153	0,2731
La Araucanía	Galvarino	85	7847	2800	0,7789	0,8071	0,0982
La Araucanía	Lautaro	237	10692	4186	0,8110	0,9262	0,2850
La Araucanía	Loncoche	169	7020	3448	0,7674	0,8290	0,2306
La Araucanía	Nueva Imperial	241	13629	5659	0,8302	0,8879	0,2721
La Araucanía	Padre Las Casas	267	30683	10125	0,8608	0,9698	0,2767
La Araucanía	Pitrufquén	108	8306	3999	0,7575	0,9023	0,1695
La Araucanía	Vilcún	171	12729	5052	0,8055	0,8348	0,2170
La Araucanía	Collipulli	101	6259	2589	0,6866	0,8877	0,1662
La Araucanía	Los Sauces	97	2847	1287	0,7788	0,8522	0,1042
La Araucanía	Lumaco	113	5515	2167	0,6293	0,7718	0,2020
La Araucanía	Traiguén	92	4493	1610	0,7576	0,9121	0,1179
La Araucanía	Victoria	191	9326	3869	0,7705	0,9963	0,2789
Los Ríos	Valdivia	125	11364	5171	0,7927	0,9653	0,1781
Los Ríos	Los Lagos	182	9836	4412	0,8094	0,7497	0,3487
Los Ríos	Mariquina	134	11319	4979	0,8248	0,8329	0,1800
Los Ríos	Paillaco	142	7736	3186	0,7930	0,8208	0,1655
Los Lagos	Puerto Montt	199	25759	11807	0,7759	0,9913	0,3077
Los Lagos	Calbuco	113	17989	7555	0,6077	0,8198	0,1979
Los Lagos	Fresia	140	4930	2467	0,8180	0,8227	0,2000
Los Lagos	Frutillar	170	5476	2650	0,8707	0,9188	0,2243
Los Lagos	Los Muermos	158	8962	4236	0,7723	0,8345	0,2671
Los Lagos	Llanquihue	115	3369	1409	0,8704	0,9640	0,1458
Los Lagos	Mauñín	80	7590	3889	0,6176	0,7425	0,1478

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Los Lagos	Ancud	152	10726	5862	0,7630	0,8084	0,3006
Los Lagos	Dalcahue	98	6562	2989	0,7734	0,8983	0,1427
Los Lagos	Puyehue	182	6669	3078	0,8209	0,7634	0,2778
Los Lagos	Río Negro	200	7101	3032	0,8594	0,9209	0,2770
Los Lagos	San Juan de la Costa	117	6342	3627	0,7207	0,8262	0,2115
Los Lagos	San Pablo	179	5421	2533	0,7290	0,9219	0,3076
Magallanes y de la Antártica Chilena	Punta Arenas	92	5660	4008	0,8324	0,9815	0,2015

Fuente: Elaboración propia

**Tabla 7.** Clasificación media

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Tarapacá	Iquique	39	2403	1488	0,3577	0,9790	0,1507
Antofagasta	Calama	35	7244	4262	0,5019	0,9720	0,0835
Antofagasta	Antofagasta	22	7769	699	0,2554	0,9754	0,0884
Valparaíso	Valparaíso	21	710	1225	0,5227	0,9868	0,0309
Metropolitana de Santiago	Puente Alto	1	12	5	0,1797	0,9052	0,0009
Metropolitana de Santiago	Lo Barnechea	12	2699	2209	0,3134	0,9696	0,0152
Metropolitana de Santiago	La Florida	1	117	44	0,0448	0,9934	0,0009
Bío Bío	Coronel	37	3188	1434	0,4422	0,9920	0,0612
Los Lagos	Quinchao	62	4879	2227	0,2889	0,8308	0,1086

Fuente: Elaboración propia

**Tabla 8.** Clasificación Transición a Alta

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Arica y Parinacota	General Lagos	100	683	696	0,6278	0,1636	0,2554
Tarapacá	Camiña	33	1242	1086	0,8942	0,4210	0,0319
Tarapacá	Colchane	61	1313	2005	0,8095	0,4108	0,1568
Antofagasta	Ollagüe	9	321	439	0,8402	0,1400	0,0231
Antofagasta	San Pedro de Atacama	45	5472	2165	0,5387	0,5226	0,1560
Antofagasta	Tocopilla	31	555	1525	0,6625	0,6277	0,0918
Atacama	Diego de Almagro	27	670	698	0,6027	0,6528	0,0742
Coquimbo	La Higuera	73	2948	2103	0,6275	0,7016	0,2305
Coquimbo	Canela	139	7133	4269	0,7102	0,6621	0,2483
Coquimbo	Los Vilos	83	4288	2727	0,4545	0,6414	0,2423
Coquimbo	Combarbalá	166	7324	4344	0,7034	0,6453	0,3250
Ñuble	Cobquecura	96	3559	2396	0,7639	0,5398	0,1164
Biobío	Lebu	59	2045	1047	0,5674	0,6839	0,2258
Biobío	Tirúa	49	6679	2443	0,6878	0,5367	0,0602
Biobío	Alto Biobío	35	5823	2094	0,5141	0,5303	0,0909
La Araucanía	Lonquimay	85	6464	3209	0,5390	0,5181	0,2246
Los Ríos	Lago Ranco	118	7627	4130	0,6702	0,7086	0,2256
Los Lagos	Cochamó	65	4019	2238	0,5669	0,4637	0,2188
Los Lagos	Quemchi	98	5840	2734	0,5514	0,6837	0,1976
Los Lagos	Futaleufú	27	765	507	0,6456	0,5880	0,0387
Los Lagos	Hualaihué	62	5303	2753	0,4269	0,4867	0,1923
Los Lagos	Palena	21	636	433	0,6210	0,5917	0,0527
Aysén del Gral. Carlos Ibáñez del Campo	Guaitecas	4	138	78	0,7632	0,4095	0,0024

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Aysén del Gral. Carlos Ibáñez del Campo	Cochrane	32	649	461	0,4710	0,4017	0,1116
Aysén del Gral. Carlos Ibáñez del Campo	O'Higgins	29	621	360	0,8783	0,4082	0,1748
Magallanes y de la Antártica Chilena	Laguna Blanca	58	274	285	0,5848	0,1854	0,1834
Magallanes y de la Antártica Chilena	Río Verde	56	617	231	0,4735	0,3826	0,0965
Magallanes y de la Antártica Chilena	Primavera	64	1158	472	0,8514	0,1826	0,1941

Fuente: Elaboración propia

**Tabla 9.** Clasificación Alta

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Tarapacá	Huara	72	1621	2328	0,3681	0,7202	0,4489
Atacama	Copiapó	142	2975	1924	0,7382	0,9854	0,4694
Coquimbo	Ovalle	341	23733	11116	0,7325	0,9911	0,6922
Coquimbo	Monte Patria	258	15413	7555	0,7030	0,8003	0,5643
Coquimbo	Río Hurtado	165	4277	2763	0,7672	0,8131	0,4432
Maule	San Clemente	264	21603	10669	0,8335	0,9350	0,4481
Maule	San Javier	216	12978	5852	0,8385	0,9582	0,3858
Biobío	Los Ángeles	373	51244	19916	0,7619	0,9934	0,4582
La Araucanía	Villarrica	342	18998	12041	0,8440	0,9170	0,3988
Los Ríos	Panguipulli	248	19266	9996	0,5997	0,7675	0,5362
Los Ríos	La Unión	219	11374	5154	0,7374	0,8723	0,3850

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Los Ríos	Río Bueno	295	14229	6583	0,5662	0,8549	0,7314
Los Lagos	Puerto Varas	222	12368	5707	0,7867	0,9745	0,5440
Los Lagos	Osorno	338	13634	5368	0,8557	0,9928	0,4422
Los Lagos	Puerto Octay	213	6946	3341	0,7509	0,8377	0,4626
Los Lagos	Purranque	229	5966	2745	0,7782	0,9033	0,3918
Aysén del Gral. Carlos Ibáñez del Campo	Coihaique	150	7850	5008	0,7502	0,9974	0,4311
Aysén del Gral. Carlos Ibáñez del Campo	Aisén	136	3651	1704	0,3974	0,9441	0,4863

Fuente: Elaboración propia

**Tabla 10.** Clasificación Muy Alta

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Arica y Parinacota	Camarones	62	1255	948	0,3596	0,2759	0,4357
Arica y Parinacota	Putre	174	613	1396	0,4186	0,4879	0,7663
Tarapacá	Pica	21	5384	659	0,1574	0,5788	0,0985
Antofagasta	Sierra Gorda	14	10158	569	0,2482	0,4905	0,0488
Antofagasta	Taltal	52	2194	688	0,2569	0,5651	0,2309
Antofagasta	María Elena	9	1532	166	0,0414	0,5722	0,0199
Los Lagos	Chaitén	82	3428	1481	0,1761	0,5205	0,4323
Aysén del Gral. Carlos Ibáñez del Campo	Lago Verde	36	850	604	0,3556	0,2709	0,7040

Región	Comuna	Número de Localidades	Población Rural	Número de Viviendas	Índice de Concentración Dispersión	Índice de Integración sede municipal	Índice de Costo de Diésel
Aysén del Gral. Carlos Ibáñez del Campo	Cisnes	89	2403	1215	0,0443	0,5263	0,5489
Aysén del Gral. Carlos Ibáñez del Campo	Tortel	36	523	320	0,2023	0,0090	0,1455
Aysén del Gral. Carlos Ibáñez del Campo	Chile Chico	46	1703	1166	0,4221	0,5044	0,3688
Aysén del Gral. Carlos Ibáñez del Campo	Río Ibáñez	49	2635	1797	0,4927	0,3632	0,3496
Magallanes y de la Antártica Chilena	San Gregorio	102	759	371	0,5718	0,2790	0,4910
Magallanes y de la Antártica Chilena	Cabo de Hornos	16	189	97	0,0160	0,4992	0,0511
Magallanes y de la Antártica Chilena	Porvenir	160	739	747	0,3044	0,4997	1,0000
Magallanes y de la Antártica Chilena	Timaukel	39	405	214	0,4319	0,2717	0,2408
Magallanes y de la Antártica Chilena	Torres del Paine	51	1203	225	0,2961	0,4823	0,3101

Fuente: Elaboración propia

Las siguientes comunas son consideradas en la clasificación muy alta por sus condiciones espaciales:

Territorios insulares de Rapa Nui, Juan Fernández y la Antártica Chilena.

## 6.2 MÉTODOS DE CLASIFICACIÓN UTILIZADOS

En este capítulo se detallan la fundamentación, desarrollo y resultados de cada uno de los métodos utilizados, así como la fundamentación de la elección de uno de estos.

Los modelos de agrupamiento más conocidos son el algoritmo de K-Medias y el Agrupamiento Jerárquico. Ambos modelos de clasificación son no supervisados, es decir, no se conocen cuantas son las clases en que se pueden agrupar los datos. Este tipo de modelo trabaja en base a encontrar semejanzas entre los datos, agrupándolos según las características de las variables

### 6.2.1 Agrupamiento por K medias<sup>21</sup>

El método de K-Medias basa su funcionamiento en agrupar los datos de entrada en un total de k conjuntos definidos por un centroide, cuya distancia con los puntos que pertenecen a cada uno de los datos es la menor posible. En términos generales, el algoritmo puede resumirse como:

1. Definir un total de k centroides al azar.
2. Calcular las distancias de cada uno de los puntos de entrada a los k centroides, y asignar cada punto al centroide cuya distancia sea menor.
3. Actualizar la posición de los k centroides, calculan-

do la posición promedio de todos los puntos que pertenecen a cada clase.

4. Repetir los pasos 2 y 3 hasta que los centroides no cambien de posición y, por lo tanto, las asignaciones de puntos entre clases no cambie.

Sin embargo, la cantidad óptima de centroides k a utilizar no necesariamente se conoce de antemano, por lo que es necesario aplicar una técnica conocida como el Método del Codo a fin de determinar dicho valor. Básicamente, este método busca seleccionar la cantidad ideal de grupos a partir de la optimización de la WCSS (Within Clusters Summed Squares).

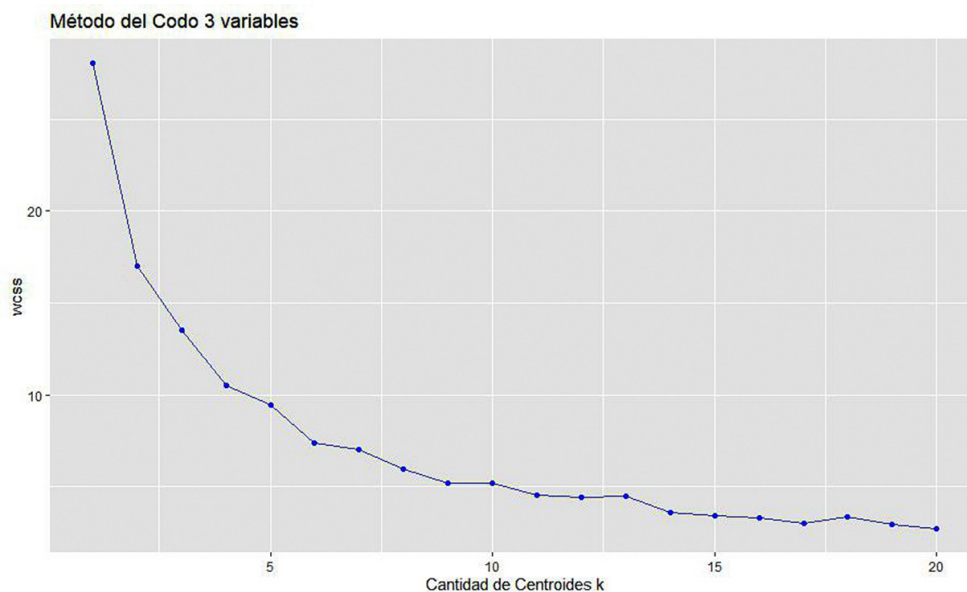
Por otro lado, ya que los centroides iniciales se generan al azar, pueden obtenerse resultados distintos en cada corrida del algoritmo, e incluso debido a las ubicaciones iniciales de los centroides, al final pueden obtenerse soluciones que son mínimos locales en vez del global real del conjunto de datos. Para solventar este problema, se propone el algoritmo de k-means++ a fin de escoger los centroides iniciales que garantizaran la convergencia adecuada del modelo.

Para implementar el modelo de K-Medios, se debe determinar la cantidad óptima de centroides a utilizar a partir del Método del Codo. Para ello, se aplicará la función kmeans al conjunto de datos, variando en cada caso el valor de k, y acumulando los valores de WCSS<sup>22</sup>, tal como se ve en Gráfico 4 WCSS.

21 Introducción a los Modelos de Agrupamiento (Clustering) en R, Ronald Delgado 23 de Junio de 2018. <https://rpubs.com/rdelgado/399475>

22 Suma cuadrática de cada clúster.

Gráfico 4. WCSS



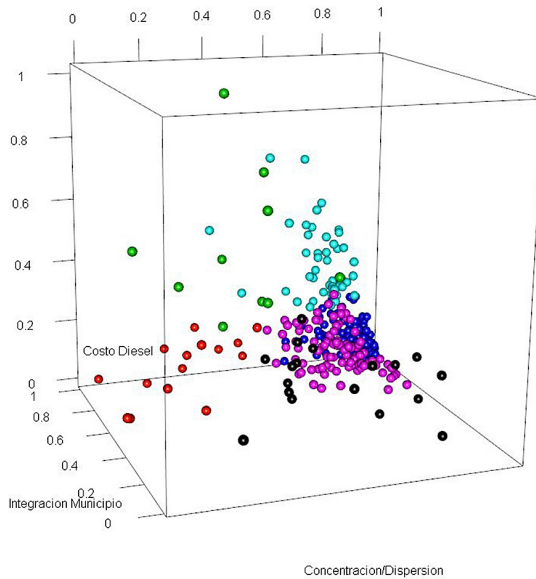
Fuente: Elaboración propia

Se puede observar que, a medida que aumenta el número de clases  $k$ , el valor de los WCSS disminuye. A partir de  $k \geq 6$ , se observa que ya no se producen variaciones importantes del valor WCSS al aumentar

los valores de  $k$ . El resultado de la clasificación para  $k=6$  clases, se aprecia en la Ilustración 11. Resultado clasificación  $K= 6$ .



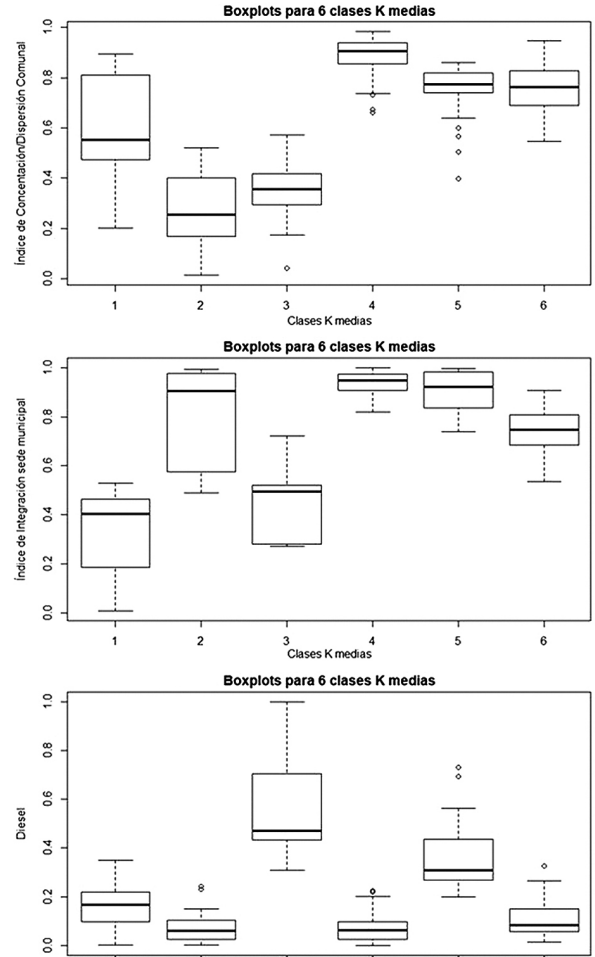
**Ilustración 11.** Resultado clasificación K= 6



Fuente: Elaboración propia

A continuación, por medio de gráficos box-plot Ilustración 12 y estadísticas descriptivas, se muestra el comportamiento de cada de las variables, por cada uno de los grupos obtenidos, en la ejecución de este método. Ilustración 12 y resumen de estadísticas descriptivas en Tabla 11.

**Ilustración 12.** Boxplots k= 6



Fuente: Elaboración propia

**Tabla 11.** Estadísticas descriptivas de clasificación k medias k=6

Clases k Medias		1	2	3	4	5	6
Número de Comunas		18	15	10	129	47	97
Índice de Concentración Dispersión comunal	Promedio	<b>0,6059</b>	<b>0,2721</b>	<b>0,3317</b>	<b>0,8899</b>	<b>0,7596</b>	<b>0,7580</b>
	Mín.	0,2023	0,0160	0,0443	0,6610	0,3974	0,5467
	Máx.	0,8942	0,5227	0,5718	0,9831	0,8608	0,9455
	Desvest	0,1929	0,1648	0,1433	0,0636	0,0917	0,0919
Índice de integración de sede municipal	Promedio	<b>0,3484</b>	<b>0,7968</b>	<b>0,4567</b>	<b>0,9374</b>	<b>0,9050</b>	<b>0,7394</b>
	Mín.	0,0090	0,4905	0,2709	0,8201	0,7400	0,5367
	Máx.	0,5303	0,9934	0,7202	0,9991	0,9974	0,9079
	Desvest	0,1530	0,2087	0,1425	0,0442	0,0809	0,0819
Índice de Costo de Diésel	Promedio	<b>0,1583</b>	<b>0,0821</b>	<b>0,5506</b>	<b>0,0698</b>	<b>0,3553</b>	<b>0,1067</b>
	Mín.	0,0024	0,0009	0,3101	0,0000	0,1987	0,0149
	Máx.	0,3496	0,2423	1,0000	0,2243	0,7314	0,3250
	Desvest	0,0886	0,0757	0,2119	0,0520	0,1219	0,0657

Fuente: Elaboración Propia

Al analizar, se puede advertir que los grupos clasificados difieren entre sí, es decir, ninguna clase tiene valores similares para las 3 variables consideradas en el análisis. Esto confirma que la elección de 6 grupos parece ser adecuada. Para el grupo k=5, se observan algunos datos estadísticamente anómalos en el índice de concentración/dispersión y de costo de diésel, esto quiere decir, que en esas variables existen valores que se escapan de lo “esperable”.

## 6.2.2 Agrupamiento Jerárquico

El Agrupamiento Jerárquico es un método de agrupamiento que basa su funcionamiento en encontrar jerarquías en los datos de entrada a partir de generar grupos basados en la cercanía o semejanza de tales datos. En el caso aglomerativo, se empieza calculando los puntos de los datos de entrada que estén más cercanos y se crea un grupo entre ellos. Luego se calculan los siguientes pares más cercanos y de manera ascendente se van generando grupos de clases que, de manera visual, podrán observarse a partir de la construcción de un Dendrograma.

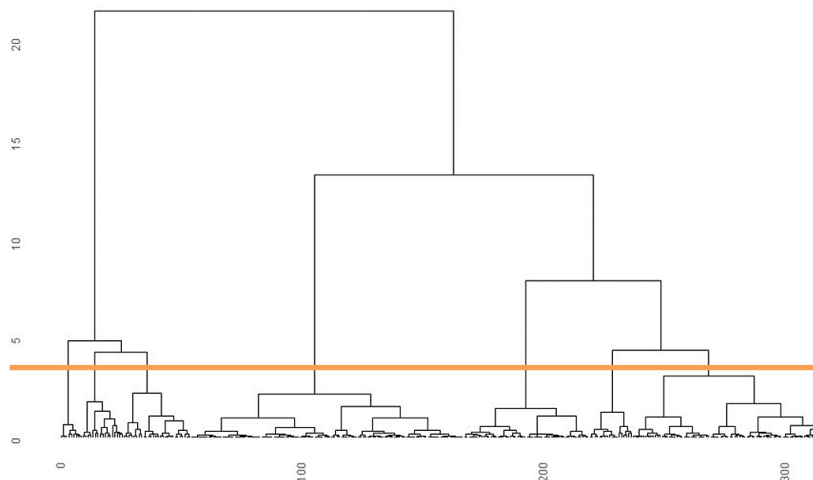
Así, como el primer paso para determinar un número de clases es el método de k medias, para el agrupamiento jerárquico hay que construir un dendrograma (Ilustración 13), las clases estarán definidas por la cantidad de líneas verticales.

En el eje horizontal del dendrograma, se tiene cada uno de los datos que componen el conjunto de entrada, mientras que en el eje vertical se representa la distancia

euclídea que existe entre cada grupo, a medida que éstos se van jerarquizando.

Cada línea vertical del diagrama representa un agrupamiento que coincide con los puntos arrojados por ésta, y como se ve en el dendrograma, estos van formándose progresivamente hasta tener un solo gran grupo determinado por la línea horizontal superior.

**Ilustración 13.** Dendrograma

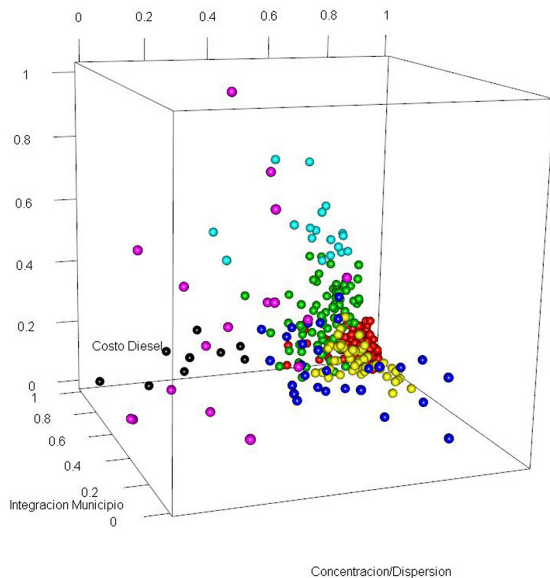


**Fuente:** Elaboración propia

Así, al ir descendiendo en la jerarquía, de un solo grupo pasa a 2, luego a 3, luego a 6, y así sucesivamente. Una manera de determinar la cantidad de clases adecuadas es cortando el dendrograma a aquella altura del

diagrama que mejor representa los datos de entrada. Para el caso de análisis, el corte propuesto en el dendrograma da como resultado 7 clasificaciones o clases, como puede apreciarse en la Ilustración 14.

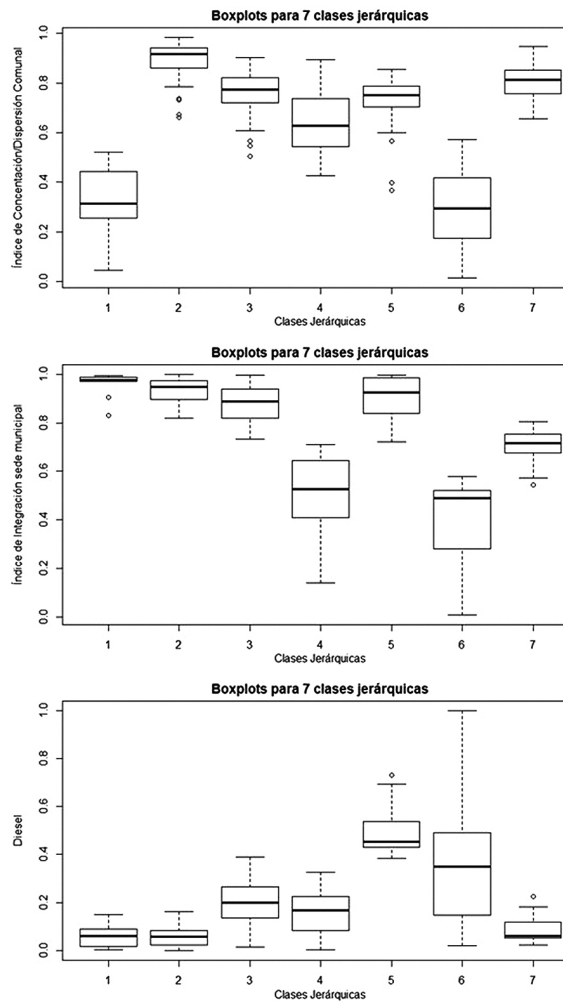
**Ilustración 14.** Resultado clasificación J= 7



Fuente: Elaboración propia

A continuación, por medio de gráficos box-plot (Ilustración 15) y estadísticas descriptivas, se muestra el comportamiento de cada de las variables, por cada uno de los 7 grupos obtenidos, en la ejecución de este método. Ilustración 12 y resumen de estadísticas descriptivas en la Tabla 12.

**Ilustración 15.** Boxplots j= 7



Fuente: Elaboración propia

**Tabla 12.** Estadísticas descriptivas de clasificación jerárquica J=7

Clases Clasificación Jerárquica		1	2	3	4	5	6	7
Número de Comunas		9	113	79	28	18	17	52
Índice de Concentración Dispersión comunal	Promedio	<b>0,3230</b>	<b>0,8980</b>	<b>0,7626</b>	<b>0,6446</b>	<b>0,7116</b>	<b>0,2821</b>	<b>0,8088</b>
	Mín.	0,0448	0,6610	0,5043	0,4269	0,3681	0,0160	0,6548
	Máx.	0,5227	0,9831	0,9025	0,8942	0,8557	0,5718	0,9455
	Desvest	0,1545	0,0611	0,0837	0,1324	0,1415	0,1623	0,0686
Índice de integración de sede municipal	Promedio	<b>0,9560</b>	<b>0,9345</b>	<b>0,8788</b>	<b>0,4961</b>	<b>0,9032</b>	<b>0,4233</b>	<b>0,7114</b>
	Mín.	0,8308	0,8201	0,7336	0,1400	0,7202	0,0090	0,5432
	Máx.	0,9934	0,9991	0,9963	0,7086	0,9974	0,5788	0,8056
	Desvest	0,0539	0,0477	0,0767	0,1697	0,0870	0,1543	0,0575
Índice de Costo de Diésel	Promedio	<b>0,0600</b>	<b>0,0561</b>	<b>0,1935</b>	<b>0,1515</b>	<b>0,4844</b>	<b>0,3672</b>	<b>0,0822</b>
	Mín.	0,0009	0,0000	0,0149	0,0024	0,3850	0,0199	0,0216
	Máx.	0,1507	0,1607	0,3893	0,3250	0,7314	1,0000	0,2236
	Desvest	0,0522	0,0382	0,0870	0,0851	0,0978	0,2752	0,0466

Fuente: Elaboración propia

Al analizar, se puede advertir que los grupos clasificados difieren entre sí, es decir, ninguna clase tiene valores similares para las 3 variables consideradas en el análisis. Esto confirma que la elección de 7 grupos parece ser adecuada. Para el grupo J=6, que la distancia intercuartílica es mayor a las otras clases.

### 6.2.3 Elección del método de agrupamiento

A partir de lo desarrollado en los puntos 5.1.1 y 5.1.2, ambos métodos parecen consistentes, en cuanto a que la media de sus variables difiere de grupo a grupo.

En la Tabla 13, se muestra el conteo resumen de cuantas comunas fueron clasificadas para cada método por clases. Se debe tener claro que las etiquetas de las clases no coinciden con las características predomi-

nantes de la clasificación, es decir, que la clase  $k=1$  del método de agrupamiento  $k$  medias no es equivalente a  $j=1$ , en efecto, para  $k=1$  se clasifican 18 comunas y para  $j=1$  solo 9.

**Tabla 13.** Resumen de casos clasificados

Clases $k$ medias	Clases Análisis Jerárquico							Total Clase $K$
	1	2	3	4	5	6	7	
1	0	0	0	15	0	3	0	18
2	9	0	0	10	0	5	0	15
3	0	0	0	0	1	9	0	10
4	0	111	18	0	0	0	0	129
5	0	0	30	0	17	0	0	47
6	0	2	31	12	0	0	52	97
<b>Total clase <math>J</math></b>	<b>9</b>	<b>113</b>	<b>79</b>	<b>28</b>	<b>18</b>	<b>17</b>	<b>52</b>	<b>316<sup>20</sup></b>

Fuente: Elaboración propia

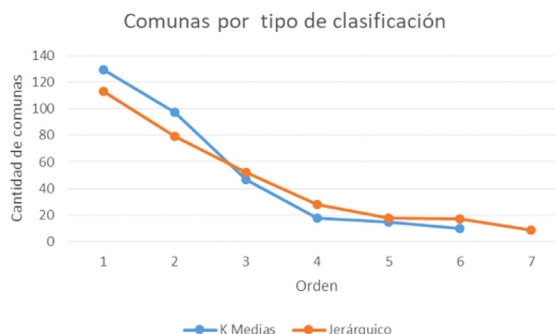
Tampoco es correcto asociar si un par de clases tienen el mismo número de comunas es el mismo grupo, solo que el método lo etiquetó de manera distinta, para  $k=1$  y  $j=5$  se tienen 18 comunas. Sin embargo, las comunas clasificadas para  $k=1$ , se distribuyen entre los grupos  $j=4$  y  $j=6$ .

Las clasificaciones que son más similares entre ambos métodos serían,  $K=4$  con  $J=2$ .

Al tratarse de técnicas de análisis exploratorio de datos, se debe tener en cuenta el sentido de la utilidad de la clasificación. Al ordenar de mayor a menor la cantidad de comunas Gráfico 1, se observa que el método jerárquico, tiene una curva con menor pendiente, por lo que los cambios de cantidad de comunas no son tan abruptos. Además, las clases con mayor cantidad de comunas,  $j=2$  con 113 es menor a  $k=4$  que tiene 129 comunas.

EL método jerárquico con 7 clases, es el que se ocupará para la clasificación.

**Gráfico 5.** Decrecimiento por tipo de clasificación



Fuente: Elaboración propia

## 6.3 CONCEPTOS TEÓRICOS

### 6.3.1 Conceptos de econometría espacial

La dependencia o autocorrelación espacial aparece como una consecuencia de la existencia funcional entre lo que ocurre en un punto determinado del espacio y lo que ocurre en otro lugar (Cliff and Ord, 1973; Paelink y Klaaseen, 1979; Anselin 1988a). Es decir, el valor que toma una variable en una zona no se explica por condiciones internas, sino que también por el valor de la misma variable en sus zonas vecinas.

Las matrices de ponderación espacial, tienen la finalidad de recoger las interdependencias de las diferentes relaciones de un punto en el espacio con lo que ocurre en otro.

La dependencia espacial, a diferencia de la temporal (unidireccional), es multidireccional (una zona puede no solo estar afectada por la zona contigua a ella, sino que además, por otras muchas que la rodean, que al igual que ellas pueden influir sobre aquellas). Este es el llamado problema frontera, el cual aparece como consecuencia de que la dependencia espacial no se limita a las zonas que incluyen una muestra, sino que se extiende a unidades espaciales para las cuales no se dispone de la información no existiendo, por lo tanto, una única solución. Esto imposibilita la ocupación de un operador de retardo, en el caso de una dependencia temporal.

La solución de multidireccionalidad en el contexto espacial pasa por definir una matriz que se denomina de pesos espaciales, retardos o de contactos  $W$ .

$$W = \begin{bmatrix} 0 & w_{12} & \dots & w_{1N} \\ \vdots & \ddots & & \vdots \\ w_{N1} & w_{N2} & \dots & 0 \end{bmatrix}$$

Siendo  $W$  una matriz no aleatoria cuyos elementos  $W_{ij}$  reflejan la intensidad de la interdependencia entre un par de zonas  $i$  y  $j$ .

Sin embargo, para definir los pesos no existe una definición ampliamente aceptada, solo que los pesos sean positivos y finitos (Anselin 1980). Lo más usual es que se ocupe el concepto de contigüidad física, donde el elemento  $W_{ij}$  vale 1, en el caso si son físicamente adyacentes y 0 en caso contrario.

Esta definición de contigüidad naturalmente asume la existencia de un mapa, a partir del cual se pueda discernir la existencia o no de fronteras. Si las distintas unidades espaciales son irregulares, esta matriz se puede obtener de manera directa. Sin embargo, no diferencia aquellas unidades que tienen una frontera en común muy amplia de las que apenas tienen un borde en común.





Cuando las unidades espaciales son puntos y no áreas, como pueden ser las ciudades en una jerarquía urbana, el significado de la contigüidad puede transformarse en la noción de distancia. En este caso estableceríamos un umbral de distancia a partir del cual consideraríamos a dos unidades espaciales como contiguas, asignándolas por tanto un 1.

El concepto de contigüidad binaria, fue extendido por Cliff y Ord (1973, 1981) para incluir una medida general de la interacción potencial entre dos unidades espaciales. La especificación correcta de los elementos de esta matriz es uno de los temas metodológicos más difíciles y discutidos en econometría espacial.

La sugerencia inicial de Cliff and Ord consiste en la utilización de una combinación de medidas de distancia (inversa de la distancia, una función exponencial negativa de la distancia) y la longitud relativa de la frontera que tienen en común dos unidades espaciales en relación a la longitud de frontera que tiene una unidad espacial con el resto de unidades de su alrededor. Los elementos resultantes serán por tanto asimétricos, a menos que ambas unidades espaciales tengan la misma longitud de frontera. Así Cliff and Ord (1973 y 1981) sugieren la siguiente definición:

$$W_{ij} = d_{ij}^{-a} \beta_{ij}^b$$

Donde:

$d_{ij}$ : es la distancia que separa las unidades  $i$  y  $j$ .

$\beta_{ij}$ : es la longitud relativa de la frontera en común, entre  $i$  y  $j$ , con relación al perímetro de  $i$ .

$a$  y  $b$ : parámetros a estimar.

Para evitar la complejidad de estimar los parámetros  $a$  y  $b$  Anselin (1980) sugiere definir los parámetros a priori.

En la misma línea, Dacey (1968) sugirió que los elementos de esta matriz de ponderación espacial, también tuvieran en cuenta el área relativa de las unidades espaciales.

$$W_{ij} = \gamma_i \beta_{ij} \alpha_j$$

Donde:

$\beta_{ij}$ : es la longitud relativa de la frontera en común, entre  $i$  y  $j$ , con relación al perímetro de  $i$ .

$\alpha_j$ : es un factor de contigüidad binario.

$\gamma_i$ : es el área de la región  $i$  en relación al área total del sistema.

Anselin (1980) propone la utilización de una matriz inversa de las distancias al cuadrado, de manera que el valor de la independencia entre dos zonas  $i$  y  $j$  disminuye a medida que la distancia entre estas aumenta.

Ambas medidas están muy relacionadas con las características físicas sobre un mapa de las unidades espaciales. Al igual que con las medidas de contigüidad binaria, estas medidas son poco útiles cuando las unidades espaciales son puntos en lugar de áreas. También son menos significativas cuando la interacción espacial del fenómeno bajo consideración viene determinada por factores puramente económicos, que poco tienen que ver con la configuración espacial de las fronteras en un mapa físico.

Como consecuencia, algunos autores han sugerido el uso de matrices que tengan una relación directa con el fenómeno que se pretende explicar. Bodson y Peeters (1975) introdujeron una matriz de accesibilidad general (con elementos calibrados entre cero y uno) que combina en una función logística la influencia de diversos canales de comunicación entre regiones, como pueden ser las carreteras, las líneas ferroviarias, y otros nexos de comunicación.

$$W_{ij} = \sum_{n=1}^N K_n \left\{ \frac{a}{1 + b * e^{-c_j d_{ij}}} \right\}$$

Donde:

$K_n$ : es la importancia relativa del medio de comunicación  $n$  (carreteras, líneas ferroviarias, conexiones marítimas y aéreas, ciclo vías, etc.)

$N$ : número de nexos de comunicación.

$d_{ij}$ : es la distancia que separa las unidades  $i$  y  $j$ .

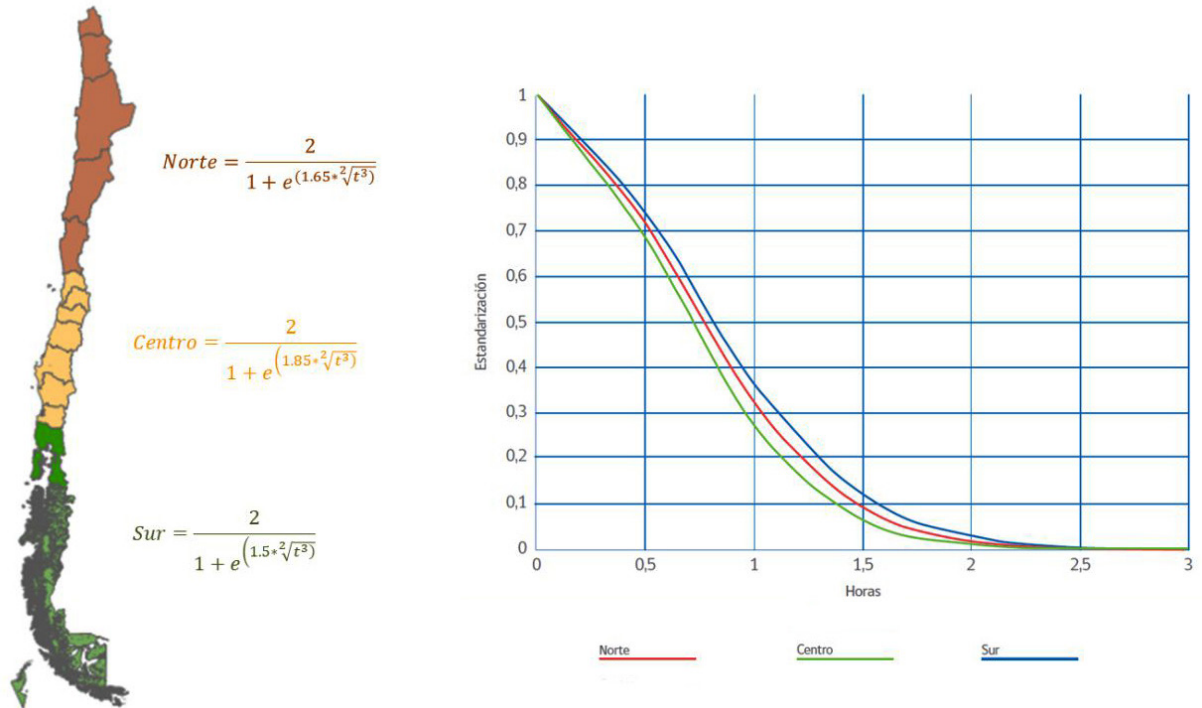
$a$ ,  $b$  y  $c_j$ : parámetros a estimar.

Uno de los problemas importantes para determinar la matriz  $w_{ij}$  de accesibilidad general es la determinación de los parámetros, pues la validez de las estimaciones de  $w_{ij}$ , está condicionada a la adecuada estimación de dichos parámetros.

Un procedimiento usual con el trabajo de matrices es el de estandarización de la matriz  $W$ . Procedimiento el cual divide cada elemento  $w_{ij}$  por la suma total de la fila a la que pertenece, obteniéndose como resultado que la suma de los elementos  $w_{ij}$  sea igual a uno. El objetivo de la estandarización es determinar la influencia total que recibe una zona de sus vecinas, independientemente del número de vecinos de cada una de las zonas. No obstante, la estandarización de  $W$  no siempre es adecuada, dado que se trata de distancias, por lo cual la matriz estandarizada carecería de significado Anselin (1988a).

En la Ilustración 16 se muestra la representación gráfica de la matriz de accesibilidad por macrozona geográfica del país. Un aspecto a resaltar es que, a partir de las 2:30 horas de traslado, los valores de la matriz tienden a 0 no importando su zona geográfica.

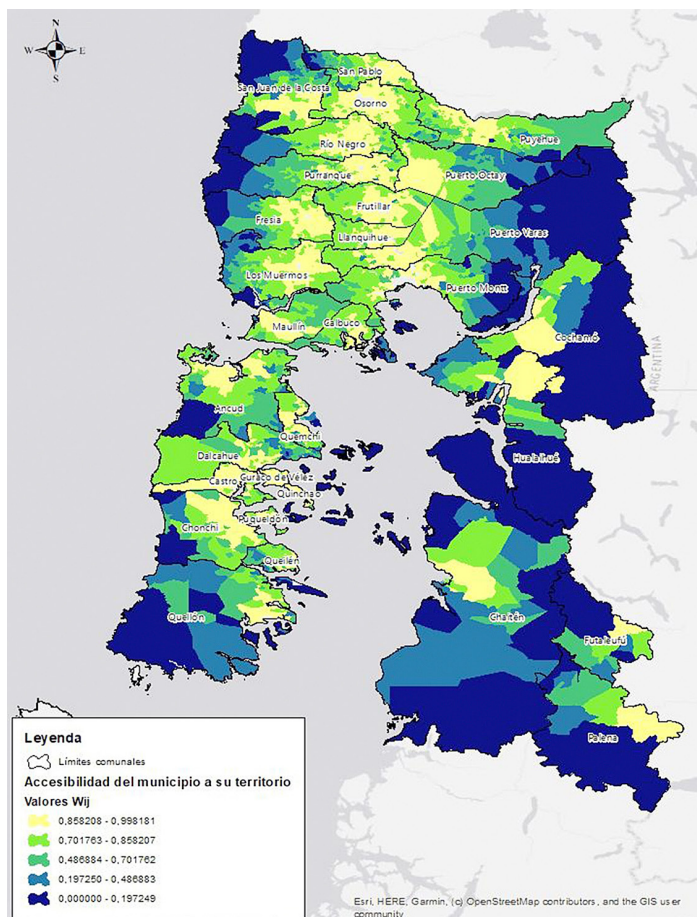
**Ilustración 16.** Matriz de Accesibilidad por macrozona



Fuente: Elaboración propia

En la Ilustración 17 se muestra los resultados de la matriz de accesibilidad del territorio comunal a su municipio.

**Ilustración 17.** Matriz Accesibilidad (Wij) Región de Los Lagos



Fuente: Elaboración propia

### 6.3.2 Indicador de Concentración/ Dispersión

Para determinar el grado de "dispersión" y "concentración", es usual ver en diversos estudios relacionados espacios rurales y geografía humana el índice de Demangeon cuya formulación matemática es:

$$I = \frac{E * N}{T}$$

Donde:

$I$ : es el valor del índice.

$E$ : es la población total de los lugares o puntos habitados que no es la localidad donde se encuentra el municipio.

$N$ : es el número de lugares que existen dentro de la comuna.

$T$ : Es la población total de la comuna.

Cuanto mayor es el número de lugares habitados y mayor sea el número de personas que residen en ellos, más elevado es el valor del índice. Si todos los habitantes del municipio residieran en un solo lugar, el índice sería igual a cero.

Este índice es muy usual en estudios de espacios rurales y geografía humana y local. Sin embargo, no rescata las particularidades de los territorios subcomunales.

El indicador que se propone es  $\rho_i$ , es el cociente entre la suma total de población ponderada de cada localidad rural  $j$ , con el total de la población en la comuna  $i$ .

$$\rho_i = \frac{\sum_{j=1}^N W_{ij} * \rho_j}{\sum_{j=1}^N \rho_j}$$

Donde:

$\rho_i$ : indicador concentración - dispersión de la comuna  $i$ .

$\rho_i \rightarrow 1$ ; máxima concentración.

$\rho_i \rightarrow 0$ ; máxima dispersión.

$w_{ij}$ : matriz de ponderación espacial de la sede municipal de la comuna  $i$ , a las localidades  $j$  pertenecientes a la comuna  $i$ .

$\rho_j$ : población de la localidad rural  $j$ .

$N$ : número de localidades de la comuna  $i$ .

El indicador  $\rho_i$ , muestra el grado de asociación, medido como concentración dispersión, de la sede municipal de una comuna respecto a sus localidades.

## 6.4. INFORMACIÓN UTILIZADA

Las bases de datos construidas se basan en los resultados preliminares de la actualización de localidades rurales según el Censo de 2017. Los puntos medidos tanto para distancia en km y horas, fueron determinados desde el municipio a todas las viviendas rurales catastradas en el Precenso 2016.

Para el cálculo por comuna se suman todas las distancias, en kilómetros, desde las localidades a la sede comunal. Para determinar cuántos litros de diésel se ocupan para ir a cada una de las localidades, se asume un gasto equivalente a 12 km/lt. El costo por litro de combustible se encuentra disponible en la página [www.bencinaenlinea.cl](http://www.bencinaenlinea.cl), desde donde se extrae el valor promedio del litro de diésel. En caso de no existir servicentro, se asigna como valor referencial valor promedio por litro de la comuna más cercana.

### 6.4.1 Base de datos de tiempos de desplazamiento y distancias

Los tiempos de desplazamientos fueron calculados mediante el software SIG TransCAD, utilizando el algoritmo de Dijkstra que se basa en teoría de grafos. La información base asume los siguientes supuestos.

- La cartografía de localidades y de la red de caminos es una representación fiel de la realidad.
- En senderos o huellas solo se transita a pie a una velocidad que ha determinado el modelo de la Subsecretaría de Transportes. Lo que implica que no se considera el tránsito de vehículos de doble tracción o tracción animal. Esto puede influir en que algunos resultados tengan tiempos sobrestimados.
- Las condiciones para el viaje son óptimas. Ello implica que el viaje se realiza sin congestión vehicular, sin condiciones atmosféricas adversas, sin cortes de camino y sin problemas para conseguir un medio de transporte que viaje a la velocidad máxima determinada por el modelo de velocidades de desplazamiento.
- El estudio se basa en cálculos de tiempos mínimos de desplazamiento desde las diferentes localidades hacia los centros de interés que son analizados. El tiempo se obtiene de un modelo econométrico que otorga velocidades a los caminos, según pendiente, material de carpeta, zona geográfica y curvatura.
- De la misma manera, se calcula la distancia, en kilómetros, de las localidades hacia la sede comunal (municipios).



## BIBLIOGRAFÍA

HERRERA M, PAZ J, Y CID, J (2012) Introducción a la Econometría Espacial. Instituto de Estudio Laborales y del Desarrollo Económico.

BARONIO A, VIVANCO A, RABANAL C (2012) Introducción a la Econometría Espacial, Dependencia y Heterogeneidad.

PEREZ J J (2006) Econometría Espacial Y Ciencia Regional. Red de revistas científicas de América Latina, el Caribe, España y Portugal.

PUMAN D, SAINT - JULIEN T (2010) Análisis Espacial - Las Interacciones. Analyse Spatiale. Les interactions, Armand Colin, 2ed, 2010. Serie GEOLibros N°21 Instituto de Geografía UC, 2014.

CARVAJAL L, POCH M (2012) Identificación de Localidades en Condición de Aislamiento, SUBDERE (2012)

CÓRDOVA O, POCH M (2019) Actualización de Variables para la Identificación de Localidades en Condición de Aislamiento, SUBDERE (2019).







Subsecretaría  
de Desarrollo  
Regional y  
Administrativo

Gobierno de Chile