

METODOLOGÍA PARA LA IDENTIFICACIÓN DE LOCALIDADES EN CONDICIÓN DE AISLAMIENTO

SUBDERE

División de Políticas y Estudios
Departamento de Estudios y Evaluaciones
Unidad de Análisis Territorial

Luis Carvajal, Geógrafo
Matías Poch, Ingeniero Civil Geógrafo
Rodrigo Osorio, Ingeniero en Informática

Santiago, 2013

ÍNDICE

CAPÍTULO I DEFINICIONES Y SUPUESTOS.	4
1.1 DEFINICIONES CONCEPTUALES Y PRÁCTICAS	4
1.2 OBJETIVOS METODOLÓGICOS	6
1.3 SUPUESTOS METODOLÓGICOS	6
CAPÍTULO II METODOLOGÍA PARA LA CREACIÓN DE INDICE DE AÍSLAMIENTO.	7
2.1 PROPUESTA TEÓRICA	9
2.2 SOFTWARES A UTILIZAR	10
2.3 INFORMACIÓN CARTOGRÁFICA	11
2.3.1 <i>Construcción y análisis de Cobertura Red de Aislamiento</i>	13
2.3.2 <i>Construcción y análisis de Cobertura Localidad</i>	27
2.3.3 <i>Implementación de información base común en SIG</i>	28
2.4 ESTANDARIZACIÓN Y CLASIFICACIÓN DE INDICADORES	39
2.5 COMPONENTE GRADO DE INTEGRACIÓN	47
2.5.1 <i>Ámbito Educación</i>	49
2.5.2 <i>Ámbito Salud</i>	49
2.5.3 <i>Ámbito Acceso a Centro Proveedor de Bienes y Servicios</i>	50
2.6 COMPONENTE CONDICIONES GEOGRÁFICAS ESTRUCTURALES	51
2.6.1 <i>Condiciones Físicas, Ambientales y de Centralidad</i>	52
2.6.2 <i>Acceso a Centros Político Administrativo</i>	57
2.7 IMPLEMENTACIÓN VISOR WEB DE RESULTADOS	57
2.7.1 <i>Construcción de visor de resultados y complemento con cartografía web e imágenes Googlemaps</i>	57
2.7.2 <i>Implementación</i>	58
2.8 MECANISMO DE VERIFICACIÓN	62
2.8.1 <i>El problema de nombres y tipo de localidades</i>	63

PRESENTACIÓN.

El presente estudio es un insumo metodológico y de diagnóstico - con una mirada nacional - para la implementación de políticas públicas regionales y sectoriales. Considera un universo total de 36.053 localidades de todo el territorio continental, desde pequeños asentamientos urbanos (villorrios, aldeas, caseríos) hasta zonas urbanas altamente pobladas, excluyendo los territorios que por ley han sido considerados como "especiales": Isla de Pascua y Juan Fernández, además de la Antártica Chilena.

A diferencia de otros estudios que se han realizado, se incorporó un enfoque "macrozonal", agrupando regiones para su análisis. Conceptualizando, desde ahora, que las variables que condicionan el aislamiento se comportan de manera distinta según zona geográfica.

Los resultados, obtenidos a partir de fuentes oficiales, deben ser considerados como un insumo base para la realización de las políticas regionales y sectoriales de apoyo a los territorios aislados, los que - atendidas las dinámicas territoriales - deben complementarse con datos y antecedentes levantados localmente, en función de los requerimientos y prioridades de dichas políticas.

CAPÍTULO I DEFINICIONES Y SUPUESTOS.

1.1 Definiciones conceptuales y prácticas

En el Decreto Supremo Nº608 publicado el 27 de noviembre de 2010, que establece la Política Nacional de Desarrollo de Localidades Aisladas, se señala que “se entiende por territorio aislado aquel con bajo nivel de accesibilidad, con escasa población y alta dispersión de ésta, baja presencia y cobertura de servicios básicos y públicos, y que, a consecuencia de estos factores, se encuentra en una situación de desventaja y desigualdad social respecto del desarrollo del país”.

A partir de lo anterior, se hizo necesario articular una definición de “Localidades Aisladas”, con el fin de permitir la identificación de estos territorios, que facilite la implementación de esta Política.

Localidad

En este estudio se entenderá como Localidad a un lugar en el espacio, representado por un punto en un mapa, que es representativo de un conjunto de viviendas habitadas por personas. Este punto es una abstracción de la realidad y en el caso de las entidades rurales, el punto sintetiza la dispersión de viviendas, por lo que su ubicación puede resultar arbitraria. Las Localidades de este estudio se clasifican en las siguientes categorías:

- i) **Sedes Comunales:** Es un asentamiento humano en donde se encuentra el municipio de una comuna, con nombre propio reconocido por la mayoría de sus habitantes. Todas las comunas que pertenecen a una Metrópolis, Grandes Áreas Urbanas y Ciudades se encuentran en esta categoría, a través de un punto representativo en el espacio, del área más densamente construida. Las sedes comunales para este estudio son 343, se excluyen las sedes comunales de las comunas de Isla de Pascua, Juan Fernández y Antártica. Todas las sedes comunales pertenecen a una de las categorías Entidad Poblada (urbana o rural) definidas por el Instituto Nacional de Estadísticas, INE.
- ii) **Zonas Urbanas:** Es un asentamiento humano de carácter urbano que no es la sede comunal de una comuna. Algunos casos de este tipo son: Tongoy (Coquimbo), Monteáguila (Cabrero), Capitán Pastene (Lumaco). Las Zonas Urbanas de este estudio son 150. El punto representativo en el mapa de la Zona Urbana se encuentra en el punto medio del polígono de la zona urbana, que se conoce como centroide.
- iii) **Entidad Rural:** Hace referencia a un asentamiento humano de carácter rural, cuyo número total de habitantes no supera las 3.000 personas. La información referida a las localidades se ha obtenido a partir de un catastro realizado por la Dirección de Planeamiento, DIRPLAN, del Ministerio de Obras Públicas, MOP, en conjunto con el INE, y que se encuentra disponible, con información del Censo de Población del año 2002, por medio de cartografía digital, a la cual ha tenido

acceso la SUBDERE. El punto representativo en el mapa, corresponde a un lugar dentro de un polígono (área), donde generalmente se encuentra la mayor concentración de viviendas.

- iv) **Localidad Aislada:** Corresponde a un punto en el espacio, habitado por menos de 3.000 habitantes, que cuenta con bajos niveles de integración (acceso a bienes y servicios del estado y de privados), con dificultades de acceso, y que por consecuencia de lo anterior, se encuentra en una situación de desventaja y desigualdad social respecto del desarrollo del país. Una localidad aislada se encuentra definida en este estudio por la relación existente entre los componentes de Aislamiento Estructural (variables morfológicas, clima y División Político Administrativa) y Grado de Integración (corresponde a la capacidad que tiene el sistema regional para atenuar estas condiciones desventajosas y lograr niveles de integración que permitan que los territorios sobrepasen, aminoren o mitiguen las condiciones de aislamiento, y puedan acceder a las dinámicas y servicios sociales, económicos, políticos, y cívicos, entre otros de los que gozan la mayoría de los habitantes del país).

El universo de datos a analizar en este estudio corresponde a las **Localidades** (Sedes Comunales, Zonas Urbanas y Entidad Rural) del país, con excepción de las localidades de las comunas que tienen estatutos y/o leyes especiales: Isla de Pascua, Juan Fernández y Antártica.

Finalmente, un **Territorio Aislado** es concebido como la agrupación de localidades aisladas, que se encuentran espacialmente en similares condiciones que permitan su agrupación, facilitando de esta manera la aplicación de políticas públicas. El "Territorio Aislado", corresponde, por lo tanto, a una delimitación instrumental para la aplicación de políticas, que puede variar en función de las necesidades de las diferentes instituciones que aplican proyectos, programas y políticas en el territorio.

Una vez establecidas estas consideraciones, el universo del estudio quedó conformado por **36.053** localidades.

Macrozona:

Una Macrozona corresponde a una agrupación funcional de regiones del país. En este estudio se ocupan tres macrozonas¹, para realizar los cálculos y estandarizaciones de las variables, de modo de obtener los indicadores. La utilización de macrozonas obedece a la intención de reconocer de manera explícita en los cálculos para determinar las localidades en condición de aislamiento, la heterogeneidad del territorio nacional, que presenta diferencias importantes, como por ejemplo: entre el desierto de la zona Norte, la tundra de las regiones australes y el clima mediterráneo de las comunas de la zona centro-sur, lo que se expresa en la cantidad de población, la manera de ocupación del territorio, las distancias y las extensiones de las comunas.

¹ Subdere (L.Carvajal-M.Poch), *Estudio Identificación de Territorios Aislados 2011*, p. 10, 2011

1.2 Objetivos Metodológicos.

1. Adaptar la metodología de identificación de territorios aislados, utilizada en el estudio de Territorios Aislados 2011², a una escala de mayor detalle que permita identificar localidades con condiciones de aislamiento.
2. Automatizar el proceso de cálculo de las localidades con mayores condiciones de aislamiento, con el fin de permitir que los gobiernos regionales puedan reasignar ponderaciones, por medio de la implementación de una Base de Datos en PostGres.

1.3 Supuestos metodológicos.

Con el fin de poder operacionalizar el estudio, que cuenta con un universo de 36.053 casos, es necesario tener presente que se trabaja con una serie de supuestos, que posibilitan el análisis de la información.

1. El *primer supuesto* está relacionado con la velocidad 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 (capitales regionales, sedes comunales, establecimientos de salud, establecimientos educacionales, entre otros). El tiempo se obtiene de un modelo matemático que otorga velocidades a los caminos, según pendiente, material de carpeta y curvatura (entre otros), y supone que una persona se desplaza siempre a la máxima velocidad que permite el modelo por cada trecho de camino.
2. El *segundo supuesto* es que 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.
3. El *tercer supuesto* es que en senderos o huellas solo se transita a pie a una velocidad que ha determinado el modelo de la Subsecretaría de transportes³. 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.
4. El *cuarto supuesto*, es que la cartografía de localidades y de la red de caminos es una representación fiel de la realidad.

² Subdere (L.Carvajal-M.Poch), *Estudio Identificación de Territorios Aislados 2011*, 91 pág., 2011

³ Actualización de la metodología de identificación de zonas aisladas para el otorgamiento del subsidio de transportes, pag. 31, Subsecretaría de Transportes, 2010.

CAPÍTULO II METODOLOGÍA PARA LA CREACIÓN DE INDICE DE AÍSLAMIENTO.

La metodología que se presenta en este documento, permite identificar las localidades que presentan condiciones de aislamiento. El universo de observaciones corresponde a 36.053 localidades del territorio nacional, con excepción del que pertenece a las comunas de Isla de Pascua, Juan Fernández y Antártica Chilena.

Es de carácter cuantitativo y la escala de este estudio es de 1:50.000, que corresponde a la escala a la que han sido levantado la mayoría de los datos geográficos que se utilizarán.

Con el fin de aprehender elementos que reconozcan la diversidad del territorio nacional, se ha incorporado (al igual que en el "Estudio Identificación de Territorios Aislados 2011"⁴) una agrupación de regiones, que permite la comparación en zonas parecidas y la aplicación de normalización de datos a grupos homogéneos. Esto implica a la vez, que los resultados entre diferentes Macrozonas no son comparables.

Las macrozonas utilizadas en este estudio son las siguientes e *Ilustración 1*

⁴ SUBDERE (Carvajal, L. – Poch, M.), op. cit.

Tabla 1

Macrozona	Regiones incluidas
Norte	Arica y Parinacota
	Tarapacá
	Antofagasta
	Atacama
	Coquimbo
Centro	Valparaíso
	Metropolitana
	O'Higgins
	Maule
	Biobío
	Araucanía
	Los Ríos
Sur	Los Lagos
	Aisén
	Magallanes

Fuente: Elaboración propia.

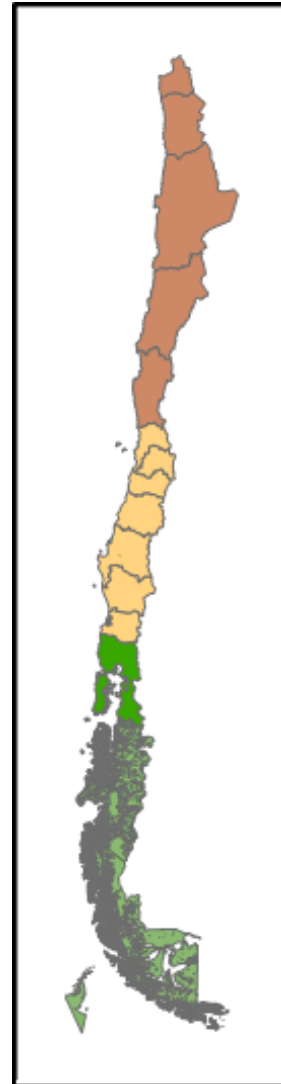


Ilustración 1

Fuente: Elaboración propia.

2.1 Propuesta Teórica

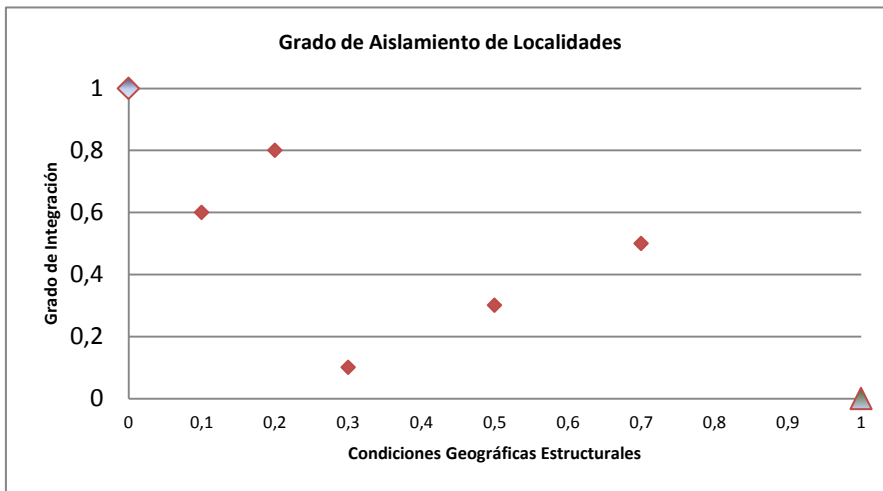
Dado que se reconoce, la situación dicotómica entre integración y aislamiento, se ha utilizado el "*Grado de Integración*" como un componente esencial en la identificación del aislamiento. Esto significa que a medida que existe un mayor grado de integración, se cuenta con un menor nivel de aislamiento. Sin embargo, no es suficiente para determinar el grado de aislamiento, por lo que fue necesario incorporar otro componente, como es la "*Condición Geográfica Estructural*" de una localidad, para lo cual se cuenta con datos que nos permiten un acercamiento teórico y práctico hacia este concepto. A medida que el condicionamiento geográfico estructural es mayor, implica mayores dificultades de acceso, por lo que actúa en favor del aislamiento.

Por lo tanto, se concibe el aislamiento como la posición relativa de un territorio entre los componentes de "*Grado de Integración*" y la condicionante "*Condiciones Geográficas Estructurales*".

De la relación entre ambos componentes, se obtiene una coordenada representada en un plano cartesiano, que indica el grado de aislamiento de cada una de las localidades.

En la *Gráfico 1* "Representación gráfica del aislamiento", la posición representada por el rombo de mayor tamaño (parte superior del eje "y", implica que las localidades que se ubican en esa posición ostentan el mínimo nivel de aislamiento, mientras que la localidad representada por un triángulo (parte derecha de eje "x") ostenta el mayor grado de aislamiento. El resto de las posiciones indica diferentes grados que deben ser determinados por las instituciones que aplicarán las políticas públicas.

Gráfico 1



Fuente: Elaboración propia.

Con el fin de poder generar un ranking de aislamiento, se genera un índice que relaciona los componentes señalados anteriormente. La fórmula de Grado de Aislamiento, es la definida en el Estudio de Territorios Aislados 2011:

$$(2 * \text{Grado de Integración}) - \text{Condiciones Geográficas Estructurales} = \text{Grado de Aislamiento}$$

Al multiplicar el componente de integración por 2, se reconoce la importancia que tiene para el aislamiento el nivel de integración. Se reafirma, de este modo, que las condiciones geográficas son una condicionante para la integración, pero no constituyen el elemento esencial para determinar el grado de aislamiento. Esto además permite metodológicamente generar un rango de resultados del índice que va desde 2 (que representa el máximo valor de *Grado de Integración* (2) y el mínimo valor de *Condiciones Geográficas Estructurales* (0)), hasta el -1 (que representa el máximo valor de *Condiciones Geográficas Estructurales* (1) y el mínimo valor de *Grado de Integración* (0)). Generar este rango permite lograr una mejor discriminación y análisis de índice de los tradicionales 0 a 1 ó 0 a 100.

Lo anterior implica metodológicamente, que si el resultado final (*Grado de Aislamiento*), es menor a 0, estamos en presencia de una localidad cuyo *Grado de Integración* es insuficiente para hacer frente a las *Condiciones Geográficas Estructurales*, por lo tanto, teóricamente esta localidad se encuentra en condición de aislamiento.

2.2 Softwares a utilizar.

En cuanto a software, se utilizan los programas:

1. **TransCad 5.0** (se procesan los cálculos de tiempo de traslado, y manejo de Bases de Datos)

2. **ArcMap 9.3** (se prepara la información que se utilizará para la revisión de los resultados)
3. **gvSIG 1.11** (se utiliza para el cálculo de geometrías, para determinar los tiempos de cada arco)
4. **PostGres** (Se utiliza para manejar los resultados de los cálculos y obtener el índice de aislamiento)
5. **Servidor de Mapa OpenLayer** (se utiliza para realizar parte de la revisión de resultados)
6. **Autodesk Infrastructure Map Server 2012** (se utiliza para la revisión de resultados)
7. **Excel 2010** (se utiliza para la generación de tablas dinámicas, para el análisis de resultados)

Los resultados que arroja el procedimiento de la ruta mínima son de aproximadamente 340 millones de registros.

2.3 Información Cartográfica.

La información base es de origen cartográfico digital y puede ser integrada en un SIG. Para realizar cualquier procesamiento de información es necesario que estas coberturas estén en un mismo sistema de referencia y proyección.

En la *Tabla 2 Información SIG* se precisan las características de las coberturas⁵ a utilizar en estudio.

⁵ Cobertura es el nombre técnico de los mapas digitales.

Tabla 2 Información SIG

Cobertura ⁶	Tipo	Formato Nativo	Fuente	Año	Datum	Huso
Red de Aislamiento	Línea	ESRI Shp	SUBDERE	2011	WGS 84	19 Sur
Red de Interconexión	Línea	ESRI Shp	DIRPLAN-MOP	2005	WGS 84	19 Sur
Red Vial	Línea	ESRI Shp	SUBTRANS	2011	WGS 84	19 Sur
Localidades	Punto	ESRI Shp	SUBDERE	2011	WGS 84	19 Sur
Sedes Comunales	Punto	ESRI Shp	SUBDERE	2010	WGS 84	19 Sur
Zonas Urbanas	Punto	ESRI Shp	SUBDERE	2011	WGS 84	19 Sur
Entidades Rurales	Punto	ESRI Shp	DIRPLAN-MOP	2005	WGS 84	19 Sur
Establecimientos Salud	Punto	ESRI Shp	MINSAL	2010	WGS 84	19 Sur
Establecimientos Educación	Punto	ESRI Shp	MINEDUC	2010	WGS 84	19 Sur
Habitabilidad	Polígono	ESRI Shp	SUBDERE-PUCV	2003	WGS 84	19 Sur

Fuente: Elaboración propia.

La **Red de Aislamiento** de la SUBDERE, está basada en la Red de Interconexión de la Dirección de Planeamiento del MOP, DIRPLAN, y en la Red Vial de la Subsecretaría de Transporte. Las **Localidades** están basadas en las coberturas de Sedes Comunales, Zonas Urbanas y Entidades Rurales.

La implementación de estas coberturas geográficas se realiza a través del *software TransCAD 5.0*⁷, con el objeto de almacenar, mostrar, y analizar datos de transporte. Se diferencia de los demás software informáticos de transporte, porque combina en una sola plataforma integrada las propiedades de un SIG y las capacidades de modelación del transporte. *TransCAD* puede usarse para todos los modos de transporte y a cualquier escala geográfica o nivel de detalle. En este caso, para los requerimientos del estudio, resultó la herramienta adecuada para procesar nuestros datos.

El método utilizado para obtener los indicadores de acceso, consistió en un procedimiento de cálculo de tiempo de desplazamiento desde **todas las localidades de Chile, hacia:**

⁶ Es un conjunto de información georreferenciada (puntos, líneas o polígonos) que representan una serie homogénea de datos geográficos que cubren un espacio determinado. *Layers* o capas de información.

⁷ www.caliper.com

- a) Todos los establecimientos educacionales
- b) Todos los establecimientos de salud
- c) Todas las principales ciudades
- d) Todas las sedes comunales⁸
- e) Todas las localidades con sucursales bancarias
- f) Las respectivas capitales provinciales
- g) Las respectivas capitales regionales

Una vez incorporadas las variables a la bases de datos de la Red de Aislamiento, se procede a filtrar los datos con el procedimiento del camino más corto, "algoritmo de rutas mínimas desarrollado por el Holandés Wybe Dijkstra, que es clave para determinar rutas mínimas, utilizando las variables de distancia y tiempo según sea el caso"⁹, para obtener, para cada localidad el menor tiempo de traslado hacia cada una de las variables medidas.

El hardware utilizado consistió en procesadores de cuatro núcleos, 16 Gb RAM, 1 GB de Ram de Video, 2 discos duros de 1 TB entre otras.

2.3.1 Construcción y análisis de Cobertura Red de Aislamiento.

Para construir la Red de Aislamiento se optó por unir la Red Vial de la Subsecretaría de Transporte, con la Red de Interconexión de DIRPLAN-MOP. La ventaja de la Red Vial de Transporte es presenta un nivel de detalle mucho mayor que la Red de Interconexión, en caminos, senderos y huellas del territorio nacional. En cambio, la Red de Interconexión de DIRPLAN tiene la ventaja de que posee, las principales vías de los países limítrofes además de las rutas de conexión marítima¹⁰.

Además de unir geoméricamente ambas coberturas, resulta imprescindible llevar a cabo una revisión topológica¹¹, que permita realizar operaciones de cálculo de redes. Los errores topológicos más comunes que se enfrentan en este tipo de tareas, se exponen en la *Ilustración 2*.

⁸ Para dimensionar este procedimiento, por ejemplo en el caso del cálculo a las sedes comunales, se multiplica el universo de localidades por la cantidad de municipalidades medidas en el estudio: 36.053 (localidades) x 342 (comunas) = 12.330.126 registros con tiempo de traslado. Esto implica que el resultado de tan solo esta operación arroja un archivo con más de 12 millones de filas que debe ser procesado.

⁹ SUBDERE (Carvajal, L.-Poch M.), op.cit., p19

¹⁰ Esto resulta de vital importancia en la zona austral del país.

¹¹ Relaciones espaciales entre elementos gráficos.

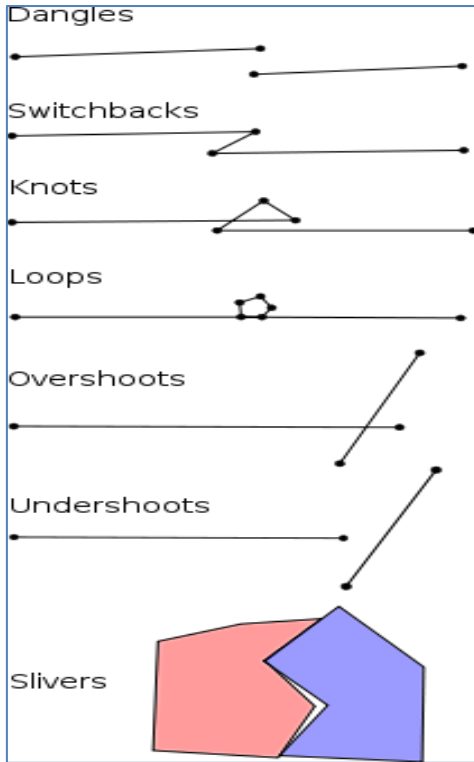


Ilustración 2
Fuente: Errores topológicos y de digitalización en SIG. Tony Rotondas 2008.

Los errores más usuales en este tipo de casos son los arcos colgantes (Dangles), arcos intersectados (Overshoots), y nodos no conectados (Undershoots). Estos errores resultan muy perjudiciales para la realización de cálculos de redes, entregando valores equivocados, debido a que no es posible utilizar el camino que topológicamente se encuentra cortado. Por ello es necesario detectar este tipo de errores, y arreglarlos con el fin de que la red permita realizar los cálculos de acceso.

Una vez que se genera la Red de Aislamiento correcta, desde el punto de vista topológico, incorpora el tiempo de desplazamiento por cada segmento a su base de datos.

Características Red Vial Subsecretaría de Transporte¹².

Esta Red es una actualización para el año 2011 y se obtuvo de la unión entre la red vial MOP (Dirección de Vialidad) y los caminos del IGM. Para ello, se procedió de la siguiente forma:

¹² Subsecretaría de Transporte y Telecomunicaciones, *Actualización de la Metodología de identificación de Zonas Aisladas para el Otorgamiento de Subsidios al Transporte*, 2010.

1. Recopilación de Información Red Vial.

Se recopilaron dos bases de información de red vial proveniente de diferentes fuentes y a distintas escalas:

- Red Vial del MOP (Dirección de Vialidad) escala 1:1.000, año 2008
- Caminos (huellas y senderos) IGM escala 1:50.000, año 2008.

2. Diagnóstico de las Bases de Información Red Vial.

Para efectuar el diagnóstico de la información se realizaron dos tipos de revisiones, la primera hace referencia a la información tabular y la segunda de carácter topológico, es decir, una revisión a la integridad de los datos.

3. Revisión Analítica.

Se realizó una revisión analítica a las bases de información con el objetivo de identificar su estado, sistema de referencia, información tabular y observaciones generales.

Adicionalmente, se procedió a revisar la información tabular de los datos, es decir, nombre de campos, tipo y número de registros por cada base de información *Tabla 3 Campos base de datos de coberturas.*

Tabla 3 Campos base de datos de coberturas.

Bases de Información	Campos			Registros
	Nombre	Tipo	Largo	
Red Vial MOP	ROL	Texto	20	7.369
	CODIGO	Texto	20	
	NOMBRE	Texto	80	
	CARPETA	Texto	10	
	LARGO	Double		
	OBSERV	Texto	70	
	ROL_PROP	Texto	20	
Red Vial IGM	Tipo	Texto	30	483.498

Fuente: Actualización de la Metodología de identificación de Zonas Aisladas para el otorgamiento de Subsidios al Transporte.

La cobertura Red Vial generada por el estudio "Actualización de la Metodología de identificación para Zonas Aisladas para el otorgamiento de Subsidios al Transporte" se detalla en *Tabla 4 Campos de base de datos cobertura Red Vial.*

Tabla 4 Campos de base de datos cobertura Red Vial

CAMPO	TIPO DE CAMPO	SIGNIFICADO
ID	Double	Identificador
LENGTH	Double	Largo en kilómetros
ROL	Texto	Número de rol
CODIGO	Texto	Código identificador de ruta
NOMBRE	Texto	Nombre camino
CARPETA	Texto	Tipo de carpeta
ROL_PROP	Texto	Número de rol propuesto
CODIGO_1	Texto	Código ruta
KM_INICIO	Double	Kilómetro inicio camino
KM_TERMINO	Double	Kilómetro término camino
MAT_CARP	Texto	Tipo de material de carpeta
ACH_CALZA	Double	Ancho calzada (metros)
EST_CALZA	Texto	Estado calzada
NUM_PISTA	Short Integer	Número de pistas
COMUNA	Texto	Nombre comuna
PROVINCIA	Texto	Nombre provincia
REGION	Texto	Nombre región
T_CALZA	Short Integer	Tipo calzada
T_CARP	Short Integer	Tipo carpeta (variable coper)
LARGO_1	Double	Largo (variable coper)
ANCHO	Double	Ancho calzada (metros; variable coper)
ALTITUD	Double	Altitud calzada (msnmm; variable coper)
PSUBIDA	Double	Pendiente de subida (m/km; variable coper)
PBAJADA	Double	Pendiente de bajada (m/km; variable coper)
TSUBIDA	Double	Tramo en subida (metros, variable coper)
CURVA	Double	Curvatura (°/km; variable coper)
IRI	Double	Coeficiente de rugosidad (variable coper)
FFL	Double	Factor de fricción lateral (variable coper)
FTNM	Double	Factor por transporte no motorizado (variable coper)
PTNA	Double	Porcentaje de tramo por no adelantar (variable coper)
VBR	Double	Velocidad de buses rurales
VBI	Double	Velocidad de buses interurbanos

CAMPO	TIPO DE CAMPO	SIGNIFICADO
IDRV	Long Integer	Identificador red vial
Tpo_Cam	Double	Tiempo caminata (horas)
Tiempo	Double	Tiempo (horas)
TARIFA	Long Integer	Tarifa (pesos)
PENDIENTE	Double	Pendiente media (%)

Fuente: Actualización de la Metodología de identificación de Zonas Aisladas para el otorgamiento de Subsidios al Transporte.

El estudio "Actualización de la Metodología de identificación para Zonas Aisladas para el otorgamiento de Subsidios al Transporte" calcula dos tipos de tiempos:

1. Viaje en red de transporte público: utilizando el modelo de costos de operación (COPER), en su versión chilena. Este modelo fue desarrollado por el Ministerio de Obras Públicas (Dirección de Vialidad) y actualizado por SECTRA en 1999. El modelo de costos de operación está basado en el módulo VOC del HDM-III (Highway Design and Maintenance Standards Model), el cual es un modelo desarrollado por el Banco Mundial para evaluar opciones de inversión en carreteras o rutas rurales en países en vías de desarrollo. El modelo HDM III - COPER, se utiliza principalmente en caminos que presentan flujo libre, es decir, para aquellos en que la interacción entre vehículos no es una variable económicamente importante. Entre los cálculos que realiza está el estimar la velocidad media de los vehículos para un arco de características dadas.
2. Tiempo de Acceso a la red de transporte público: se generó un modelo conceptual que supone que la velocidad de desplazamiento en caminata varía principalmente según el grado de pendiente que tenga la ruta. Así, el tiempo de acceso a la red de transporte público, o tiempo de caminata, se calcula sumando los cuocientes entre la distancia recorrida por cada grado de pendiente y la velocidad de desplazamiento respectiva. Además, este tiempo de desplazamiento podría estar afectado por las condiciones climáticas del lugar donde se realice, por lo que se agregan dos variables al análisis: una que representa un ponderador debido a la temperatura, y otra que representa un ponderador debido a la precipitación. Estos ponderadores afectarían el tiempo de desplazamiento total, dependiendo de los valores de las variables respectivas. El modelo que mejor se ajusta es uno lineal que considera solo la pendiente.

El tiempo de desplazamiento por red vial en este estudio no corresponde al de transporte público, por lo tanto los tiempos de este campo no son considerados. Además solo existe registro de este dato en aquellos segmentos donde existe una ruta de transporte público. El tiempo de caminata sí será considerado como dato relevante para este estudio, ya que se encuentra calculado para el tipo de carpeta Huella y Sendero.

De todos los campos generados en la "Actualización de la Metodología de identificación para Zonas Aisladas para el otorgamiento de Subsidios al Transporte", servirán para este estudio: ID, LENGTH, CARPETA, REGIÓN, Tpo_Cam y PENDIENTE.

Estimación tiempo de desplazamiento por segmento en red vial para Red de Aislamiento.

El tiempo de desplazamiento para segmento de arco vial, está determinado por las velocidades del modelo de transporte privado para el estudio "Accesibilidad territorial – Fronteras interiores", de DIRPLAN.

$$Vel = 40.629 * DHor + 40.434 * DAsf + 25.957 * DRip + 21.862 * DTie - 10.448 * DZon - 0.38 * Pendiente + 0.456 * Factor$$

Donde:

Vel : Velocidad.

DHor : Variable Dummy de carpeta hormigón.

DAsf : Variable Dummy de carpeta asfalto.

DRip : Variable Dummy de carpeta ripio.

DTie : Variable Dummy de carpeta tierra.

DZon : Variable Dummy de zona geográfica (regiones desde Arica y Parinacota hasta Maule es 0, de Biobío a Magallanes es 1).

Pendiente: pendiente media del arco calculada a partir de la diferencia de cota de sus nodos, expresada en porcentaje.

Factor : Cuociente entre la distancia real del arco, y la distancia recta que une el nodo inicial y final. Este cuociente se expresa en porcentaje, es así que 100% representan un camino recto.

De la formula anterior se concluyó en el estudio "Accesibilidad territorial – Fronteras interiores"

Tabla 5 Cálculo de valores de velocidades por arco.

Tabla 5 Cálculo de valores de velocidades por arco

Parámetros	Resultados Km/Hr.
Velocidad mínima para arcos con carpeta de Hormigón y asfalto	40
Velocidad mínima para arcos con carpeta de tierra o ripio	25
disminución de velocidad para un aumento del 1% de la pendiente del arco	0,38
Aumento de velocidad para un aumento del 1% del factor de sinuosidad	0,456

Fuente: Estudio accesibilidad Territorial – fronteras Interiores, DIRPLAN-MOP 2003. 2004.

Para calcular las velocidades para red vial terrestre y, por ende, el tiempo, es necesario:

1. Estandarizar campo carpeta a las categorías: hormigón, asfalto, ripio y tierra.
2. Calcular el cociente de sinuosidad para cada segmento del arco.

La estandarización propuesta del material de la carpeta de la Red Vial de la Subsecretaría de Transporte para utilizar fórmula de velocidades de transporte privado se muestra en la *Tabla 6 Estandarización de material de carpeta vial*.

Tabla 6 Estandarización de material de carpeta vial.

Material Carpeta	Nº arcos	Dhor	Dasf	Drip	Dtie
Adoquín	77	0	77	0	0
Asfalto	24604	0	24604	0	0
Faenas	864	0	864	0	0
Hormigón	11011	11011	0	0	0
Imprimación reforzada	316	316	0	0	0
Imprimación	1295	0	1295	0	0
Ripio	74186	0	0	74186	0
Tierra	26719	0	0	0	26719
Tratamiento superficial	11335	0	11335	0	0
TSD	2414	0	2414	0	0
TSS	267	0	267	0	0

Fuente: Elaboración propia.

Para calcular el factor de sinuosidad se prueban 3 formas de determinación, utilizando 2 software *ArcMap9.3* y *gvSIG 1.11*.

ArcMap 9.3 permite extraer la longitud de cada segmento (arco) de la cobertura de línea. Además permite extraer las coordenadas X, Y o Norte, Este, de cada uno de los nodos que componen segmento (inicio y final). Con las coordenadas es posible sacar la distancia plana o euclidiana entre dos puntos (inicio y final), y por lo tanto, conseguir el cuociente del factor de sinuosidad.

El cuociente de sinuosidad es calculado a partir del gestor de extensiones *SEXTANTE* de *gvSIG 1.11* "extraer geometrías de líneas". A pesar de ser solo un comando, genera información como: longitud del segmento, distancia lineal entre nodos de inicio y final, orientación y dirección (en grados sexagesimales) promedio en la cual trazada el arco y sinuosidad (que no es el factor, si no que el inverso de él).

Para ver que método de cálculo *Factor* se ocupará para determinar las velocidades promedios, en red vial, para la cobertura Red de Aislamiento (SUBDERE), es necesario tomar una muestra de segmentos de la cobertura Red de Interconexión (DIRPLAN) y comparar los tiempos estimados, con los tiempos estimados por:

1. Cuociente calculado en *ArcMap 9.3*, ocupando la fórmula de distancia lineal entre dos puntos y la longitud del segmento
2. Cuociente calculado por extensión *SEXTANTE* en *gvSIG*, ocupando longitud y distancia recta derivadas de la aplicación extracción de geometrías de líneas.
3. Cuociente calculado por el inverso de la sinuosidad entregada por la extensión extracción de geometrías de líneas.

Se deben eliminar todos los segmentos que presenten el error tipo *loop*, véase *Ilustración 2*, ya que se indeterminará el cuociente por estar dividido por 0. Esto se produce debido a que el nodo de inicio y final se encuentran en la misma coordenada.

Una vez estimados los tiempos por segmento se plantean medidas de ajuste, con las cuales se comparan los valores obtenidos por DIRPLAN con los estimados por los tres métodos de cálculos señalados anteriormente.

De la aplicación de cada uno de los métodos se evalúan sus medidas de ajuste, lo que permite determinar que método es el más idóneo para calcular los tiempos en la red vial que se utilizará en el estudio.

Las medidas de ajuste se basan en los datos obtenidos de la comparación de los observados (en este caso, los estimados por DIRPLAN) y los modelos simulados para determinar la conducta del modelo (estimados por las tres variaciones de cálculo del factor de sinuosidad). Los dos criterios sobre los cuales las medidas se juzgan aquí son: el grado para lo cual la sensibilidad atraviesa un rango de magnitudes de error, y la habilidad para comparar el modelo de ejecución por sobre resultados diferentes.

El primero de estos criterios indica al analista, en un sentido cualitativo, cuán bien un modelo se desempeña y a qué grado este desempeño difiere entre los modelos separados. El segundo criterio permite que los modelos sean comparados en diferentes situaciones o parámetros

Existen seis medidas de calidad de ajuste que han sido evaluadas como herramienta en el análisis sobre los dos criterios mencionados. Tal evaluación se realiza empleando la comparación entre lo estimado y lo observado (estimación DIRPLAN), que adopta la siguiente notación:

$[T_{ij}^*]$ = Tiempo estimado entre i, j .

$[T_{ij}]$ = Tiempo observado entre i, j .

T = Tiempo total observados.

\bar{T} = Tiempo medio observado por zonas.

n = Es el número de zonas.

Se revisaron las siguientes medidas de ajuste:

1. **Coefficiente de Determinación (R^2):** Esta es uno de las medidas de evaluación más usadas pero ha sido considerada insensible en relación a los parámetros diseñados por *Wilson et al (1969) and Black (1973)*. Esta cantidad indica que proporción de la variación total en la respuesta T_{ij} se explica con el modelo ajustado, es decir:

$$R^2 = \frac{\text{Var. explicada}}{\text{Var. total } T_{ij}}$$

Se calcula:

$$R^2 = 1 - \frac{\sum_{ij} (T_{ij} - T_{ij}^*)^2}{\sum_{ij} (T_{ij} - \bar{T})^2}$$

2. **Chi Cuadrado (χ^2):** Es una distribución de probabilidades, que contiene un parámetro y una cantidad de grados de libertad. Además, no puede ser negativa y todos sus valores están tabulados. Las pruebas de bondad de ajuste usan datos de muestra como base para aceptar o rechazar los supuestos que se hace acerca de la distribución. Las suposiciones se enuncian como la hipótesis nula, y se realiza calculando las frecuencias que se esperan si la hipótesis nula fuera verdadera; luego la diferencia entre las frecuencias observadas en la muestra f_0 y las frecuencias esperadas f_e , se usan para calcular el χ^2 de la muestra

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$$

Finalmente, el χ^2 de la muestra se compara con el χ^2_{av} , de las tablas para decidir si la hipótesis cero debe rechazarse o aceptarse. Se destaca por ser una estadística muy sensible, pero existen algunas dificultades con su uso. Valores intercambiables de menos de seis requieren datos de matrices agregadas y no existe un método

consistente para hacer esto. La eliminación de celdas pequeñas puede resultar en modelos de desempeño con evaluaciones erróneas, ya que tiempos más largos pueden ser mejor representados en los datos eliminados.

$$\chi^2 = \sum_{ij} \frac{(T_{ij} - T_{ij}^*)^2}{T_{ij}^*}$$

3. **Phi (Φ):** El estadígrafo phi es muy similar a la ganancia de información o a la información relativa, mientras el valor de phi es más grande, el modelo de ajuste es menos exacto.

La ganancia de Información se define como:

$$I(t^*, t) = \sum_{ij} t_{ij} \ln \left(\frac{t_{ij}}{t_{ij}^*} \right)$$

Donde:

$$t_{ij} = \frac{T_{ij}}{T} \wedge t_{ij}^* = \frac{T_{ij}^*}{T}$$

Entonces al reemplazar en la ganancia de información. Se obtiene Φ y al dejar la expresión sin cambios se obtiene ΔL

$$\phi = \sum_{ij} T_{ij} \left| \ln \frac{T_{ij}}{T_{ij}^*} \right|$$

4. **La Diferencia en el Logaritmo de Posibilidad:** *Wilson (1976)* enumera las ventajas de usar este tipo de medida, la cual se calcula de la misma manera que el estadígrafo Φ , usado para probar los estándares de ajuste en muchos modelos desagregados tales como aquellos informados por *Domencich y Macfadden(1975)*.

$$\Delta L = \sum_{ij} T_{ij} \left[\ln \frac{T_{ij}}{T} - \ln \frac{T_{ij}^*}{T} \right]$$

5. **Desviación estándar de residuos:** *Openshaw y Connolly (1977)* hicieron uso extensivo de esta medida en sus evaluaciones de funciones alternativas de detención. Principalmente mide la dispersión de datos.

$$A = \sqrt{\sum_{ij} \frac{(T_{ij} - T_{ij}^*)^2}{n^2 - 1}}$$

6. **Error Absoluto Medio:** Es la medida más simple y más probada, donde compara como diferencia absoluta el valor real con el estimado, según la cantidad de zonas (n) que represente y se define como:

$$MABSERR = \frac{\sum_{ij} (|T_{ij} - T_{ij}^*|)}{n^2}$$

El estudio del comportamiento comparativo de las seis medidas de ajuste, se realiza en base de tres criterios fundamentales, que son:

- Sensibilidad a los cambios de error.
- Similitud en la magnitud de las ciudades con diferentes tamaños.
- Consistencia con otras estadísticas.

Según lo anterior, se sugieren dos medidas como indicadores útiles estos son el **phi normalizado** y la **estadística de error absoluto medio normalizado**. Estas son igualmente sensibles a los cambios en el error en diferentes niveles, y tienen la propiedad importante de ser comparables entre diferentes matrices.

7. La versión sugerida de la estadística phi se define por:

$$\phi^* = \sum_{ij} \frac{T_{ij}}{T} \left| \ln \frac{T_{ij}}{T_{ij}^*} \right|$$

Un valor cero indica que las matrices observadas y simuladas son idénticas.

Una ventaja adicional de la estadística de phi es que el error es valorado de acuerdo al número observado o la proporción de lo observados. Los errores de la muestra son más grandes en los datos más pequeños. La estadística phi automáticamente descuenta estos errores.

8. La versión sugerida de la estadística del **valor absoluto medio** se define como:

$$MABSERR^* = \frac{1}{n^2} \sum_{ij} \left| \frac{T_{ij} - T_{ij}^*}{\bar{T}} \right|$$

Este es el radio del error absoluto medio en relación a la magnitud de intercambio medio.

Dado que tanto la matriz observada como la estimada tienen a su interior numerosas celdas con ceros, imposibilita la aplicación de χ^2, Φ (estándar y sugerida) y el error medio absoluto estándar, diferencia logaritmo de la posibilidad, dejando como medida de ajuste R^2 , la desviación estándar de residuos y el valor del error medio absoluto y valor absoluto medio (sugerida).

Una vez calibrados los distintos modelos, se calculan los diferentes estadígrafos que dan cuenta del grado de ajuste.

Es importante señalar que el modelo propuesto por DIRPLAN fue calibrado a partir de una red construida a partir de cartografía a escala 1:250.000.

Estimación tiempo de desplazamiento por segmentos en red marítima y red vial países limítrofes para Red de Aislamiento.

Para la estimación del modo de transporte marítimo¹³, se mantienen los tiempos catastrados e implementados en la Red de Interconexión DIRPLAN¹⁴, así como también los tiempos de la red vial de los países limítrofes (ver pag. 13).

A las carpetas que fueron identificadas como "CALLE", se les asigna una velocidad de 40 Km/Hr.

Las Sedes Comunales, Zonas Urbanas y Entidades Rurales no se encuentran unidas topológicamente a la Red de Aislamiento, por lo que es necesario crear un "puente" que conecte estos puntos de interés (con todos su atributos en la base de datos) con la Red de Aislamiento. Este puente se llama CONECTOR, que para este estudio une (por medio de una línea recta) las **Localidades** con el nodo¹⁵ más cercano de la Red de Aislamiento. Este conector solo tiene ID¹⁶ y longitud de línea (creados automáticamente por el sistema), pero no posee el campo tiempo en su base de datos. Para incorporar esta información se asume que la velocidad en que se recorre este conector es 5 Km/HR.

Los otros elementos a conectar a la red como son los establecimientos de salud y educación, se conectan al nodo más cercano que a su vez sea CONECTOR. Esto significa que a pesar de no estar geoméricamente en las mismas coordenadas, se puede asignar tiempo para ese arco igual a cero. Esto significa que al ejecutar el algoritmo de rutas mínimas, se está asignando matemáticamente un establecimiento, ya sea de salud o educación, al nodo que fue conectado una localidad a la red. De lo anterior se desprende que hay dos "tipos de conectores en la Red de Aislamiento: los conectores de localidades, que tienen una velocidad de 5km/h y los conectores de establecimientos de educación y salud, con una velocidad de 0km/h.

Una vez que se obtiene una red topológicamente correcta, se construye a partir de esta, una red lógica (grafo) (Network extensión *.net en TransCAD*) que permite realizar los cálculos algorítmicos.

La creación de la topología de redes, consiste en la construcción de las relaciones espaciales de los elementos espaciales considerados para el cálculo. Para este estudio entre las Localidades y la Red de Aislamiento.

¹³ Ver capítulo 2.3.1

¹⁴ Si no se cuenta con la información, esta se completa asumiendo que una velocidad de 12 nudos \approx 40 Km/Hr.

¹⁵ En SIG, las líneas están definidas por vectores compuestos por un nodo de inicio, un arco y un nodo final.

¹⁶ ID es la sigla que se utiliza en los SIG para hacer referencia a los identificadores que asigna el software

Es importante señalar que la Red de Aislamiento constituye una abstracción de la realidad. Es un dibujo que muestra la disposición y característica de la red de caminos de Chile. La Red es el resultado de la unión de diferentes capas de información digitalizadas a su vez por diferentes instituciones (IGM, MOP-Vialidad). Este dato resulta relevante al considerar que su construcción es a partir de una escala 1:50.000, en especial los datos relacionados con senderos y huellas. En la medida que nos acercamos a un área densamente poblada, la densidad de la red¹⁷ aumenta considerablemente, incidiendo directamente en los "errores de digitalización" señalados anteriormente, lo que repercute finalmente en los cálculos algorítmicos sobre la red lógica. El otro problema se relaciona con el año en que ha sido capturado el dato, debido a que no hay una actualización permanente de esta información (huellas y senderos), por lo que en zonas con una alta densidad de la red, es frecuente encontrarnos con la falta de caminos en general (apreciables en Google Earth). Además de ello, es necesario tener en consideración que la información proveniente de la Dirección de Vialidad se hace cargo de los caminos "enrolados" por el Ministerio de Obras Públicas, lo que implica que hay una serie de caminos privados que pueden ser de tierra, ripio, o asfalto, que no han sido considerados en esta red.

Los problemas señalados en el párrafo anterior, se puede afirmar que a escala 1:50.000 o inferior, los resultados se comportan bastante bien y constituyen un acercamiento aceptable hacia la realidad del aislamiento en Chile. Sin embargo, a escala local (1:20.000, 1: 10.000, 1: 5.000 o superior), las dificultades mencionadas inciden en resultados que pueden estar sobreestimados, al calcular rutas más largas de lo real, debido a la falta de uniones entre caminos, o la ausencia de algunos de ellos debido a la desactualización de la información de senderos y huellas¹⁸, y por el nivel de detalle que no puede ser abarcado con información a escala 1:50.000.

La Red de Aislamiento se configura a nivel nacional de la siguiente manera *Ilustración 3*.

¹⁷ Cantidad de líneas por unidad de superficie.

¹⁸ Este problema se aprecia principalmente en la zona aledaña de la ciudad de Santiago que ha experimentado una explosión demográfica importante, asociada a la creación de condominios y parcelas de agrado; además de explotaciones agrícolas que han mejorado sustancialmente sus caminos interiores y que no aparecen en los mapas de senderos y huellas.

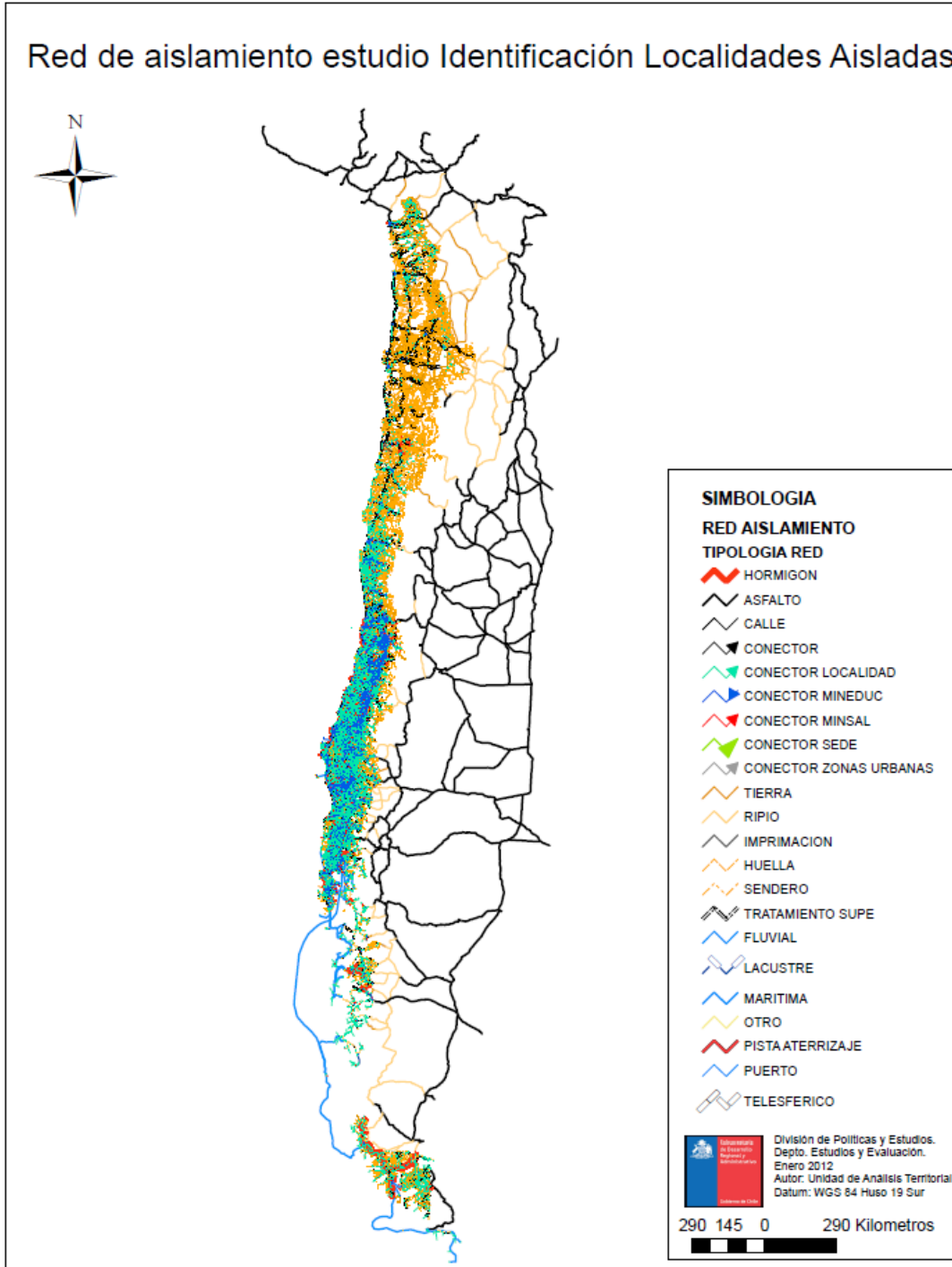


Ilustración 3
 Fuente: Elaboración propia.

2.3.2 Construcción y análisis de Cobertura Localidad.

Se define el concepto de **Localidad** como un lugar en el espacio, representado por un punto en un mapa, donde habitan personas. Clasificándolas en tres categorías: Sedes Comunales, Zonas Urbanas y Entidad Rural.

Para construir esta única cobertura, se realiza un proceso de unión espacial de cada una de las coberturas a un nodo representativo de la Red de Aislamiento (CONECTOR) (véase 2.3.1 *Construcción y análisis de Cobertura Red de Aislamiento*). Es importante señalar que existen casos donde la Sede Comunal es a la vez una Entidad Rural, pues tiene todas las características que define el INE, pero se opta por dejar para un mismo nodo los códigos de comuna y el código de entidad manzana (rural), para realizar cruces de información de bases de datos censales. El código de esta cobertura es alfanumérico.

Sedes comunales.

Es una cobertura de puntos georreferenciados, en la que su base de datos tiene el código único territorial (CUT)¹⁹, para la identificación de la comuna. Es el punto donde se encuentra localizada la cabecera comunal y que será tomado como representativo para efectos de cálculos. Además, al poseer el código único territorial permite integrar y gestionar bases de datos. La base cartográfica de esta cobertura fue corregida utilizando las imágenes satelitales de Google Earth. En general, el 10% de las sedes comunales fueron reubicadas utilizando la aplicación Google Earth en *TransCad 5,0*, disminuyendo considerablemente el error original producto de la utilización de una escala tan pequeña.

Zonas Urbanas.

Es una cobertura de puntos georreferenciados, que representa un asentamiento humano de carácter urbano que no es la sede comunal de una comuna. Para este estudio se crea un código para gestionar los cálculos, que corresponde a la concatenación del código CUT de la comuna a la que pertenece. De haber más de una Zona Urbana por comuna se agrega al final un dígito correlativo, según el orden alfabético, el nombre de uso y costumbre de sus habitantes. El punto corresponde al centroide²⁰ de una cobertura de polígonos de zonas urbanas construida por Mideplan.

¹⁹ CUT: Código Único Territorial, compuesto por un máximo de 5 dígitos. Los dos primeros dígitos indican el número de región, el tercer dígito hace referencia a la provincia, y los últimos dos dígitos se refieren a la comuna. Esta codificación fue generada en un trabajo conjunto entre Subdere y el Instituto Nacional de Estadísticas (INE), y fue oficializado vía Decreto Supremo del Ministerio del Interior, Diario Oficial del 8 de mayo de 2000.

²⁰ Es el punto que define el centro geométrico de un objeto.

Entidad Rural.

Es una cobertura de puntos georreferenciados, que representa un asentamiento humano de carácter rural²¹, cuyo número total de habitantes no supera las 3.000 personas. Esta información se ha obtenido a partir de un catastro realizado por la Dirección de Planeamiento del MOP, en conjunto con el INE, y que se encuentra disponible, con información del Censo de población del año 2002. Cada Entidad Rural posee un código de manzana entidad, el cual está compuesto por el código de comuna, código de distrito, zona censal y manzana entidad.

Esta cobertura fue construida a partir de una cartografía escala 1:50.000, con datos del año 2002 (Censo de Población y Vivienda).

Cobertura de localidades

La cobertura Localidad, quedó compuesta tabularmente por cuatro campos alfanuméricos (columnas) fundamentales:

1. CUT: corresponde al código de la comuna a la que pertenece una localidad.
2. COD_ZUR: corresponde al código asignado a cada Zona Urbana.
3. ENT: corresponde al código manzana entidad rural del censo 2002.
4. TERRITORIO: es el código que se le asigna a la localidad. Si es Sede Comunal se asigna el CUT (si la sede comunal es además una Entidad Rural aplica solo el CUT), en los casos de Zona Urbana (COD_ZURB) y Entidad Rural (ENT) se mantienen los códigos COD_ZURB y ENT.

Esta codificación permite enlazar relacionamente las localidades con bases de datos de la división político administrativa; de la cantidad de instituciones financieras²²; de los centros de salud; y de los centros educativos; que se encuentran presentes en el territorio. La habitabilidad se enlaza a las localidades de manera geométrica a través de una unión o enlace espacial. Esto implica que la localidad asume el valor de habitabilidad del polígono que se encuentra justo sobre el punto georeferenciado.

2.3.3 Implementación de información base común en SIG²³.

La implementación de la información base en común tiene como fin realizar el cálculo de tiempos de desplazamientos entre puntos de interés para procesar posteriormente los distintos indicadores, considerados para determinar el grado de aislamiento y/o integración.

²¹ El punto representa una dispersión de viviendas, que por ser rural, puede abarcar un área considerable.

²² Se construyó un listado por comuna de las instituciones financieras presentes en base de Superintendencia de Banco e Instituciones Financieras SIBF 2011, las cuales fueron relacionadas a la base de datos de Localidad por medio del campo común CUT.

²³ Sistema de Información Geográfico

Los cálculos referidos a distancia y tiempos son complejos, se basan en algoritmos y heurísticas de programación lineal entera (PPL), en cambio los índices requieren cálculos simples basados en resultados de tiempos (grado de integración).

El SIG *TransCad* posee herramientas de análisis de redes basada en teoría de grafos, para obtener de manera expedita los datos referidos a desplazamiento entre los distintos puntos de interés por atributo en torno a la red. La herramienta de cálculo de redes se basa en el algoritmo de rutas mínimas desarrollado por el Holandés Wybe Dijkstra, que es clave para determinar rutas mínimas utilizando las variables de distancia y tiempo, según sea el caso.

Debido a la complejidad del cálculo es necesario explicitar desde el punto de vista teórico el desarrollo del algoritmo de Dijkstra. Para ello se revisa la teoría de grafos, la formulación matemática y la aplicación algorítmica.

Teoría de Grafos²⁴.

Es una importante aplicación de las técnicas de programación lineal entera para la determinación de rutas de costo mínimo, en grafos. Para definir los términos del problema:

Sea $G(V,E,c)$ un grafo dirigido etiquetado, i.e.,

- V es un conjunto de vértices o nodos,
- $E \subseteq V \times V$ es un conjunto de arcos,
- $c: E \rightarrow \mathbf{R}$ es una función que etiqueta los arcos²⁵, i.e., para $e \in E$, $c(e)$ es la etiqueta de e .

Es usual que el grafo $G(V,E,c)$ sea finito, es decir, que el conjunto de vértices V lo sea. Como convención se supondrá que $V = \{1,2,\dots,n\}$. La etiqueta de un arco representa una especie de costo de la transición entre los vértices que se conectan.

Un camino de longitud m , desde el vértice u hasta el vértice u_{m-1} es una secuencia

$$r = \langle u_0, u_1, \dots, \dots, u_{m-1} \rangle \in V^+$$

Para $i = 0, \dots, m-1$, $(u_i, u_{i+1}) \in E$. En este caso, se define el costo del camino r como:

$$\text{costo } r := \sum_{i=0}^{m-1} c(u_i, u_{i+1})$$

El grafo $G(V,E,c)$ se puede representar con una matriz de distancias (directas) $D = (d_{ij}) \in \text{Mn}(\mathbf{R}^*)$, definida de modo que, para $1 \leq i, j \leq n$:

²⁴ Jaime Bohorquez y Rodrigo Cardozo, *Análisis de Algoritmos*, Universidad Los Andes, Colombia, 1992.

²⁵ $\mathbf{R}^* := \mathbf{R}^+ \cup \{0, \infty\}$

$d(i,i) := 0,$

$d(i,j) := c(i,j),$ si $(i,j) \in E,$

$d(i,j) := \infty,$ si $(i,j) \notin E$

La representación es tal, que se olvida el costo de las "lazos", i.e., si hay un arco con $c(i,i) > 0$, la representación lo "cambia" por otro de costo 0. La razón para esta decisión es el interés en determinar costos mínimos de caminos entre vértices; de esta forma, un camino que utilice un lazo de costo positivo siempre tiene costo mayor que otro que no lo usa. Por otro lado, los vértices originalmente no conectados, se consideran ligados por un arco de costo infinito.

A partir de multiplicaciones de matrices $M_n(\mathbf{R}^*)$ se puede determinar la matriz $D^* = (d_{ij}^*)$, donde d_{ij}^* es, el costo de un camino de costo mínimo del vértice i al vértice j . Entonces, D^* se llama la matriz de distancias mínimas.

Se plantean los siguientes problemas:

- Encontrar D^* .
- Dado un vértice u , encontrar $d^* u.$, i.e., $d^* uv$, para $v \in V$.
- Dados dos vértices u, v , encontrar $d^* uv$.

El problema de encontrar la matriz de distancias mínimas resuelve los otros dos problemas, y resolver el problema de costos mínimos de un vértice a todos los demás conlleva una solución para encontrar el costo mínimo entre dos vértices. Por el contrario, es fácil construir soluciones de los problemas más complejos a partir de soluciones para los simples.

Para calcular D^* , es decir, determinar, para $1 \leq i, j \leq n$, el valor de

$$d_{ij}^* = (\min r \mid r \text{ es un camino de } i \text{ a } j: \text{costo } r)$$

Una especificación del problema para una solución algorítmica:

Ctx: m : **array**[1..n,1..n] **of** \mathbf{R}^*

Pre: $m = D$

Pos: $m = D^*$

La búsqueda de una solución puede afrontarse con técnicas de programación.

Para $W \subseteq V, i, j \in V$:

$$R_{ij}^W := \{r \mid r = \langle i, w_1, w_2, \dots, w_m, j \rangle \text{ camino, para } k = 1, \dots, m: V \in W\}$$

$$C_{ij}^W := (\min: r \in R_{ij}^W : \text{costo } r)$$

De esta manera, para $1 \leq i, j \leq n$:

$$d_{ij}^* = C_{ij}^V$$

Entonces se puede establecer una recurrencia para calcular los $C_{ij}^W, W \subseteq V, i, j \in V$

$$C_{ij}^\emptyset = D_{ij}$$

$$C_{ij}^{W \cup \{v\}} = C_{ij}^W \min(C_{iv}^W + C_{vj}^W), \forall v \notin W$$

Si se denota $C^{\{1, \dots, k\}}_{ij}$ como C^k_{ij} , se tiene que:

$$C_{ij}^k = \begin{cases} D_{ij} & , \text{ si } k = 0 \\ C_{ij}^{k-1} \min(C_{ik}^{k-1} + C_{kj}^{k-1}), & \text{ si } 0 < k \leq n \end{cases}$$

Esta recurrencia es la base fundamental para establecer la corrección del llamado algoritmo de Floyd-Warshall:

proc FW (**var** m: **array**[1..n, 1..n] of R*)

{Pre: m = D}

{Pos: m = D*}

k := 0;

{Inv P: $0 \leq k \leq n \wedge m = C^k$ }

do $k \neq n \rightarrow i, j, k := 1, 1, k+1$;

{Inv P1: $1 \leq i \leq n+1 \wedge 1 \leq j \leq n \wedge m = C^{k-1}$ }

do $i \neq n+1$

$\rightarrow m[i, j] := m[i, j] \min (m[i, k] + m[k, j]);$

if $j \neq n \rightarrow j := j+1$

$j = n \rightarrow i, j := i+1, 1$

od

od

Para comprobar que los invariantes se conservan, debe notarse que:

$$C_{ik}^{1\dots k} = C_{ik}^{1\dots k-1}, C_{kj}^{1\dots k-1} = C_{kj}^{1\dots k}$$

o:

$$C_{ik}^k = C_{ik}^{k-1}, C_{kj}^{1\dots k-1} = C_{kj}^{1\dots k}$$

De manera que:

$$m[i, j] := m[i, j] \min(m[i, k] + m[k, j])$$

El algoritmo de Floyd-Warshall puede "decorarse" para resolver el problema relacionado de determinar una ruta de costo mínimo entre cada par de nodos. De este modo, se puede decidir, al construir un camino de costo mínimo, si se debe o no utilizar cada nodo intermedio considerado.

Esta idea se plasma al mantener una matriz *intmd*, de manera que, para $1 \leq i, j \leq n$:

$intmd_{ij} = v$, si existe una ruta que pasa por $v \in V$, con costo mínimo

= 0, si la ruta más corta de i a j es el arco $(i, j) \in E$

= -1, si no hay camino de i a j .

Más aun, se pueden definir los valores:

$intmd_{ij}^k = v$, si existe una ruta que pasa por $v \in V$, con costo mínimo, cuyos nodos intermedios están en $\{1, \dots, k\}$

= 0, si la ruta más corta de i a j es el arco $(i, j) \in E$

= -1, si no hay caminos de i a j , cuyos nodos intermedios están en $\{1, \dots, k\}$.

Por lo tanto:

$$intmd_{ij}^n = intmd_{ij}$$

Entonces se pueden modificar los algoritmos para que calculen los nodos intermedios de las rutas óptimas.

Algoritmo Dijkstra

Una variante del problema de determinar todas las distancias mínimas consiste en determinar, para un vértice fijo f (que actúa de *pivot*), llamado la fuente, las distancias a todos los demás vértices. En otras palabras, determinar d_{fv}^* , para $v \in V$, o bien, el vector D_f^* .

Una manera de solucionar el problema consiste en ejecutar el procedimiento de Floyd-Warshall y quedarse con la fila f de la matriz respuesta.

El algoritmo de Dijkstra resuelve el problema de la siguiente manera (la semántica del procedimiento `Escoja_min`):

```

proc Dijkstra (G(V,E,c): grafo, f: V; var df:array[1..n] of R* )
  {Pos: df = D*f}
  M:= {f};
  for v ∈ V → df[v]:= c[f,v] rof;
  {Inv P( ∀ v: v ∈ M : df[v]=D*fv ) ∧ (∀ v: v ∈ V\M : df[v]=CMfv) }
  do M≠V → w:= Escoja_min (df,V\M);
    {Q1: P ∧ df[w] ≤ df [V\M] ∧ w ∈ V\M }
    M:= MU{w};
    for v ∈ V\M ∧ (w,v) ∈ E
      → df[v]:= min{df[v],df[w]+c[w,v]}
    rof
  od

```

Como el w elegido en cada iteración es un nuevo miembro del conjunto M , supóngase que no fuera cierto que:

$$df[w] = D_{fw}^*$$

En este caso, debería existir un camino de costo óptimo de f a w , cuyos nodos intermedios no pueden estar todos en M (puesto que $df[w] = C_{fw}^M$). Sea u el primer nodo en este camino que no está en M .

Entonces:

$$\begin{aligned}
 D_{fw}^* &= D_{fu}^* + D_{uw}^* \\
 &= D_{fu}^M + D_{uw}^* , \text{ porque los antecesores de } u \text{ están en } M \\
 &= df[u] + D_{uw}^* , \text{ por el invariante} \\
 &< df[w] , \text{ porque se supone que } df[w] \text{ no es óptimo} \\
 &\leq df[u] , \text{ porque } w \text{ se escoge con } df[.] \text{ minimal.}
 \end{aligned}$$

Y se concluiría que $D^*uw < 0$, lo cual es contradictorio.

Por otra parte, para cada nodo en el nuevo $V \setminus M$, el último ciclo actualiza los valores para satisfacer el invariante.

El conjunto M se llama de esta manera porque "marca" los nodos para los que ya se conoce el óptimo. A $V \setminus M$ se le llama el conjunto de los "abiertos". Una forma de expresar el algoritmo con una estructura de datos que administre los abiertos es:

```

proc Dijkstra (G(V,E,c): grafo, f: V; var df:array[1..n] of R* )
    {Pos: df = D*f.}
    A:= V\{f};
    for v ∈ V    → df[v]:= c[f,v] rof;
    do A≠∅      → w:= Escoja_min(df,A);
                A:= A\{w};
                for v ∈ A ∧ (w,v) ∈ E
                    → df[v]:= df[v] min (df[w]+c[w,v])
                rof
    od

```

El algoritmo de Dijkstra se puede considerar formado por la secuenciación de una inicialización

```

INIC: A := V \ {f};
      for v ∈ V → df[v] := d[f,v] rof;
y de un ciclo:
      CICLO: do A ≠ ∅ → w := Escoja_min (df,A);
              A := A \ {w};
              for v ∈ A ∧ (w,v) ∈ E
                  → df[v] := min{df[v],df[w]+c[w,v]}
              rof
      od
  
```

El tamaño del problema es "n", el número de vértices del grafo. Por otra parte, sea e = |E| el número de arcos del grafo²⁶.

Como operaciones básicas:

- # : añadir un elemento de un conjunto
- \ : eliminar un elemento de un conjunto
- Escoja_min : escoger el elemento de un conjunto que alcanza el mínimo de una medida
- min : calcular el mínimo de dos números.

Así:

$$T_D(n) = T(\text{INIC}) + T(\text{CICLO})$$

$$T(\text{INIC}) = O(n) T(\#)$$

$$T(\text{CICLO}) = O(n) [T(\text{Escoja_min}) + T(\backslash)] + O(e) T(\text{min})$$

Es decir:

$$T_D(n) = O(n) [T(\#) + T(\text{Escoja_min}) + T(\backslash)] + O(e) T(\text{min})$$

La complejidad definitiva depende de las estructuras de datos que se utilicen. En otras palabras, ya que es concebible representar tanto el grafo $G(V,E,c)$ como el conjunto de

²⁶ Obsérvese que $0 \leq e \leq n^2$, de modo que $e = O(n^2)$

abiertos A con diversas estructuras de datos, el uso de recursos dependerá de la conveniencia de estas estructuras para realizar el algoritmo.

A continuación se muestran tres variantes de representación que arrojan complejidades diferentes para las ejecuciones del algoritmo de Dijkstra:

Variante 1:

- $G \approx$ matriz de distancias directas (adyacencias etiquetadas con costos): $O(n^2)$.
- $A \approx$ arreglo booleano: $O(n)$.

Como costo de las operaciones básicas, se supone:

- $T(\#) = O(1)$
- $T(\backslash) = O(1)$
- $T(\text{Escoja_min}) = O(n)$
- $T(\text{min}) = O(1)$

Entonces:

$$\begin{aligned} T_D(n) &= O(n) [O(1) + O(n) + O(1)] + O(e) O(1) \\ &= O(n^2 + e) \\ &= O(n^2). \end{aligned}$$

Variante 2:

- $G \approx$ lista de distancias directas: $O(n^2)$.
- $A \approx$ arreglo booleano: $O(n)$.

Como costo de las operaciones básicas, se supone:

- $T(\#) = O(1)$
- $T(\backslash) = O(1)$
- $T(\text{Escoja_min}) = O(n)$
- $T(\text{min}) = O(1)$

El análisis es idéntico al del caso anterior:

$$T_D(n) = O(n^2)$$

Variante 3:

- $G \approx$ lista de distancias directas: $O(n^2)$.
- $A \approx$ montón (inglés: heap) : $O(n)$.

Como costo de las operaciones básicas, se supone:

- $T(\#)$ = $O(\log n)$
- $T(\backslash)$ = $O(\log n)$
- $T(\text{Escoja_min})$ = $O(1)$
- $T(\text{min})$ = $O(1)$

$$T_D(n) = O(n) [O(\log n) + O(1) + O(\log n)] + O(e) O(1)$$

$$= O(n \log n + e)$$

Ahora, si el grafo $G(V,E,c)$ no tiene "demasiados arcos", i.e., si $e \leq n \log n$, se tendrá que:

$$T_D(n) = O(n \log n).$$

Este caso, de grafos con pocos arcos, es frecuente en la práctica. Es decir, se tiene una variante "práctica" del algoritmo de Dijkstra, mejor que $O(n^2)$.

Teniendo claro conceptualmente cómo se calcula rutas mínimas con grafos, es necesario construir un grafo en base a la información cartográfica. Como se mencionó con anterioridad el software a utilizar es *TransCAD*, por lo que se importan los *layers* de información desde el archivo nativo *ESRI Shape File .shp* a *.dbf* archivo nativo de *TransCAD*.

Creación de Topología²⁷ de Redes.

La creación de la topología de redes, consiste en la construcción de las relaciones de los elementos espaciales considerados para el cálculo. Para este estudio entre las localidades y la red de interconexión.

Las localidades no se encuentran unidas topológicamente a la Red de Aislamiento, por lo que es necesario crear un puente que conecte las localidades (con todos sus atributos en la base de datos) con la red de interconexión²⁸ (ver página 24). Los establecimientos de educación y salud se unen al nodo del CONECTOR más próximo, asumiendo una velocidad de 0 Km/HR.

²⁷ Es la definición de las reglas y relaciones espaciales entre líneas, puntos y polígonos.

²⁸ La red no incluye el trazado urbano, por lo que es necesario unir las sedes a la red por una línea ficticia, para permitir los procesamientos necesarios.

Una vez que se cuenta con una red gráfica topológicamente correcta, se construye a partir de esta una red lógica (grafo) (*Network* extensión *.net* en *TransCAD*). Esta red permite realizar los cálculos algorítmicos.

Cálculo de ruta mínima.

El sistema construido permite enlazar bases de datos referidas con la localización de diferentes puntos de interés, por lo tanto permite evaluar dentro de todas las rutas posibles la menor entre los nodos seleccionados. El resultado final debe ser un archivo que tenga formato de base de dato, dado el gran volumen de información se opta que las salidas en base de datos sean en formato de *TransCad* (*.bin*).

Para obtener las rutas mínimas se realiza una consulta que busca los trayectos desde "todas las localidades (36.053)" hacia "todas las localidades (36.053)". Esto implica que el sistema de información geográfico debe calcular "todas las posibilidades de trayectorias" entre todas las localidades (36.053), para encontrar las rutas mínimas. Por lo tanto el cálculo es de $36.053 * 36.053 = 1.299.818.809$ de registros como resultado. Esto implica un archivo con cerca de 1.300 millones de filas.

Con el fin de optimizar este cálculo, se ha optado por realizar cálculos escalonados:

- a) Desde todas las sedes comunales hacia todas las localidades ($343 * 36.053 = 12.366.179$ registros), lo que genera un archivo con más de 12 millones de registros.
- b) Desde todos los establecimientos educativos²⁹ analizados hacia todas las localidades ($1.827 * 36.053 = 65.868.831$), lo que genera un archivo con más de sesenta y cinco millones de registros.
- c) Desde todas los centros de salud³⁰ analizados hacia todas las localidades ($2.112 * 36.053 = 76.143.936$), lo que genera un archivo con más de setenta y seis millones de registros.

Estos tres procesamientos permiten realizar las consultas a la base de datos buscando por tipo de establecimientos, los tiempos mínimos. En el caso de las sedes comunales, es posible obtener la información relacionada con los tiempos mínimos hacia:

1. Sedes comunales
2. Capitales provinciales
3. Capitales regionales
4. Principales ciudades
5. Centros proveedores de servicios

Debido al gran volumen de datos generados, el procesamiento se hace inviable en herramientas tradicionales como hojas de cálculo (*Excel*), por lo que su procesamiento se realiza en *Transcad*. Los resultados de tiempos mínimos con 36.053 registros, son

²⁹ Un establecimiento puede tener más de un tipo de enseñanza, por lo que solo se mide una vez y después se realizan uniones de bases de datos para realizar filtros por tipo de enseñanza.

³⁰ Un establecimiento puede tener más de un tipo de atenciones, por lo que solo se mide una vez y después se realizan uniones de bases de datos para realizar filtros por tipo de atenciones.

procesados en un manejador de base de datos, que en este caso es *PostGres* que permite exportar los resultados a su herramienta espacial *PostGis*, que puede ser analizada en diferentes sistemas de información geográficos, como por ejemplo *gvSIG*.

2.4 Estandarización y clasificación de indicadores.

"Lo que no se define, no se puede medir. Lo que no se mide, no se puede mejorar. Lo que no se mejora, se degrada siempre."³¹

Todas las variables de este estudio están referidas a los tiempos de desplazamiento. La excepción es la habitabilidad, compuesta de cinco categorías que representan la agresividad del medio físico.

El proceso de estandarización corresponde a una transformación de un conjunto de datos en otro. Para efectos de este estudio, consiste en transformar el valor de los tiempos de desplazamiento medidos de las localidades a cada uno de los puntos de interés (Salud, Educación, División Política Administrativa, centros proveedores de bienes y servicios, y la ciudad principal) a un indicador cuyo rango sea entre 0 y 1.

Para estandarizar se ha optado por determinar funciones matemáticas en que los límites superiores e inferiores tomen valores entre 0 y 1, o bien, que converjan cuando la variable tienda al infinito.

Para determinar las funciones que resultan adecuadas para la estandarización, es necesario revisar algunos conceptos de econometría espacial. Esto permitirá tener una base teórica que sustente la definición de las funciones (transformaciones) para estandarizar las variables.

Debido a que en este estudio se han definido dos componentes: *Grado de Integración* y *Condiciones Geográficas Estructurales*, es necesario encontrar dos tipos de funciones, una para cada componente.

La medición del Grado de Integración se basa en la accesibilidad de los territorios, por lo tanto se sustenta a 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."

³¹ William Thomson. Lord Kelvin.

La filosofía de esta propuesta es estandarizar de las variables (ya sea para el componente de *Grado de Integración y Condiciones Geográficas Estructurales*) a partir de una matriz de accesibilidad o de ponderación espacial. Este tipo de matrices se fundamentan en base a los conceptos de “dependencia o autocorrelación 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. Es decir, el valor que toma una variable en una zona no se explica por condicionantes internas, sino que también por el valor de la misma variable en sus zonas vecinas³².

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 muchas otras que la rodean, que al igual que estas puede influir sobre aquellas.

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

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

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

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, asignándoles por tanto un 1.

Como consecuencia, se define una matriz de accesibilidad general (con elementos calibrados entre cero y uno) que combina en una función logística la influencia de modos de comunicación entre territorios, 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:

³² El concepto de vecindad debe ser entendido en función de las relaciones que se establecen entre dos puntos en el espacio y no vinculado exclusivamente a la proximidad cartográfica (Cartesiana o Euclidiana).

Kn : 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 (tiempo) que separa las unidades i y j .

a, b y c_j : parámetros a estimar.

Los parámetros a, b y c_j , en este caso, se estiman a partir del comportamiento de la variable a estandarizar, y su relación con el concepto que se desea buscar. Por ejemplo, un indicador de accesibilidad a hospital de alta complejidad no puede tender a 0 a partir de que el tiempo sea mayor a 15 horas, no resulta razonable siendo que según la experiencia empírica a partir de 4 horas se puede afirmar que una localidad está en condiciones de aislamiento con respecto a ese servicio. De esta forma, la estimación de los parámetros de este estudio dependerá mucho de la experiencia y sensibilidad que tiene el analista de datos y conocimiento del fenómeno que intenta medir. En otras palabras, al no existir una métrica ni unidades de medidas internacionales, para la medición del fenómeno de aislamiento, es necesario crear una métrica que dé respuesta al concepto que se intenta medir³³.

Por lo tanto, 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 los dichos parámetros.

Definición de matriz w_{ij} de ponderación espacial.

Debido a la naturaleza de los datos puntuales, se opta por utilizar distintas variantes de la matriz de accesibilidad general (con elementos calibrados entre 0 y 1).

Se proponen seis matrices de accesibilidad, una para cada macrozona y para las componentes de Grado de Integración y Condiciones Geográficas Estructurales. Sin embargo, se opta por modificar el orden de los parámetros a calibrar, por lo tanto pierde la característica de función logística, pero conserva la propiedad de que cada uno de sus elementos se encuentra dentro del rango cero y uno.

$$w_{ij} = \frac{a}{b + e^{-c \cdot d_{ij}}}$$

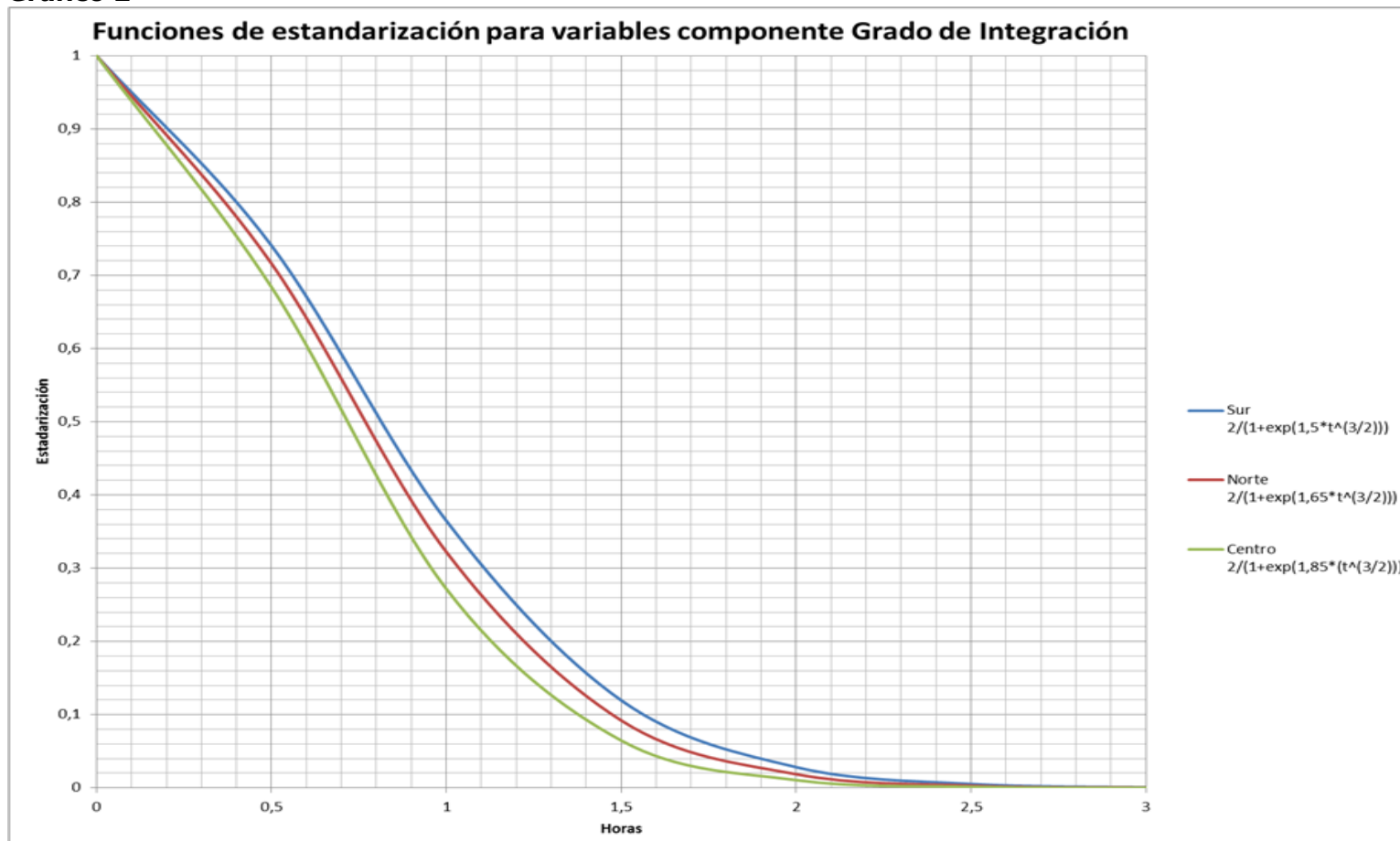
Las matrices de ponderación espacial se diferencian entre sí, debido a los diferentes parámetros a, b y c (definidos a priori).

³³ Si no se puede medir, no se puede probar. Lord Kelvin (creador de escala natural para plantear las ecuaciones termodinámicas).



En *Gráfico 2 "Funciones de estandarización para variables de Componente de Integración"* se observan los valores las matrices propuestas, para el componente Integración y para cada macrozona.

Gráfico 2



Fuente: Elaboración propia.

$$Norte = \frac{2}{1 + e^{(1.65 \cdot \sqrt{t^3})}}$$

$$Centro = \frac{2}{1 + e^{(1.85 \cdot \sqrt{t^3})}}$$

$$Sur = \frac{2}{1 + e^{(1.5 \cdot \sqrt{t^3})}}$$

En la *Tabla 7 Valores de estandarización variables componente Grado de Integración*. Se observan los valores de estandarización, a partir de valores de tiempo para cada macrozona.

Tabla 7 Valores de estandarización variables componente Grado de Integración.

HORAS	NORTE	CENTRO	SUR
0	1	1	1
0,5	0,716318227	0,684144	0,74088
1	0,322217899	0,271746	0,364851
1,5	0,092068371	0,064674	0,119534
2	0,018627524	0,010623	0,028332
2,5	0,00293655	0,001333	0,005307
3	0,000377971	0,000134	0,000824
3,5	4,06339E-05	1,1E-05	0,000109
4	3,7012E-06	7,47E-07	1,23E-05

Fuente: Elaboración propia.

Para el *Componente Geográfica Estructural* se proponen funciones que distan de las propuestas para el componente *Grado de Integración*. Donde el parámetro *a*, aparece multiplicando una función exponencial y se incorpora una sustracción. A pesar de estas modificaciones, conserva la propiedad de que cada uno de sus elementos se encuentra dentro del rango entre 0 y 1.

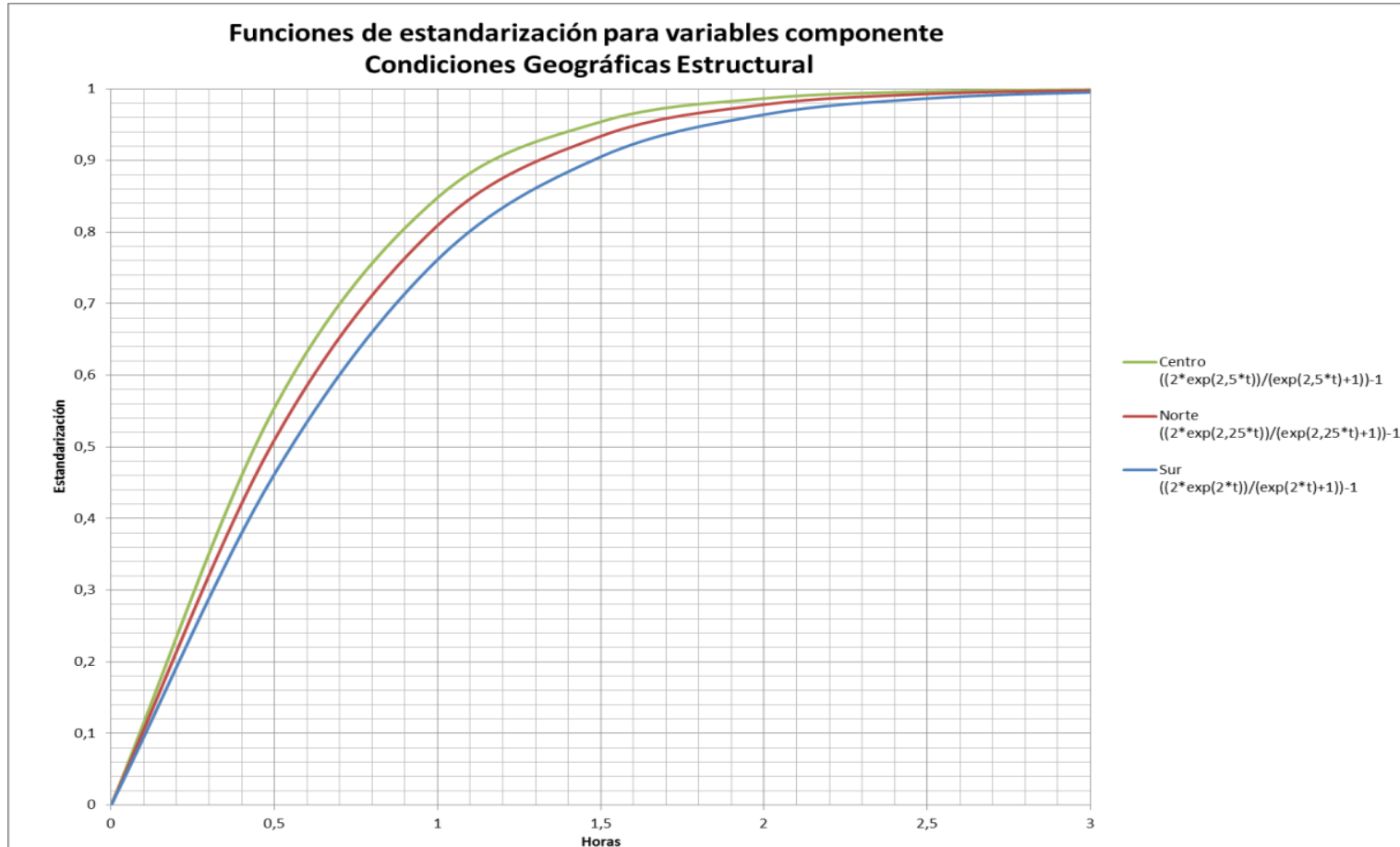
En el



Gráfico 3 "Funciones de estandarización para variables de Componente Geográficas Estructural" se observan los valores de las matrices propuestas, para el componente Geográfico Estructural y para cada macrozona.



Gráfico 3



Fuente: Elaboración propia.

$$Norte = \frac{2e^{2.25*t}}{e^{2.25*t} + 1} - 1$$

$$Centro = \frac{2e^{2.5*t}}{e^{2.5*t} + 1} - 1$$

$$Sur = \frac{2e^{2*t}}{e^{2*t} + 1} - 1$$

En la *Tabla 8 Valores de estandarización variables componente Geográfica Estructural.* se observan los valores de estandarización, a partir de valores de tiempo para cada macrozona.

Tabla 8 Valores de estandarización variables componente Geográfica Estructural.

HORAS	NORTE	CENTRO	SUR
0	0	0	0
0,5	0,509829974	0,554599722	0,46211716
1	0,80930107	0,84828364	0,76159416
1,5	0,933828043	0,95404526	0,90514825
2	0,978026115	0,986614298	0,96402758
2,5	0,992812795	0,996146531	0,9866143
3	0,997660979	0,998894443	0,99505475
3,5	0,999240031	0,999683128	0,9981779
4	0,999753211	0,999909204	0,9993293

Fuente: Elaboración propia.

2.5 Componente Grado de integración.

Al medir el grado de integración, medimos el conjunto de variables que indican la situación de las localidades en relación a los centros políticos, urbanos, y de bienes y servicios ofrecidos por el mercado en el país. Mide la situación de cada una de las localidades en comparación a las otras localidades. La finalidad es encontrar una medida de integración.

Uno de los problemas asociados a la integración de los territorios y, por ende, asociado al aislamiento, es la dificultad para acceder a los servicios básicos que los habitantes necesitan para vivir y desarrollarse. Por ello, el Estado debe garantizar que todos los habitantes de la nación tengan un acceso equitativo a los servicios que provee. Estos servicios se localizan principalmente en los centros urbanos, en las sedes comunales, en las principales ciudades o en los centros proveedores de bienes y servicios.

Para su medición se han seleccionado un grupo de variables agrupadas en tres ámbitos: Educación, Salud y Centros Proveedores de Bienes y Servicios (*ver Tabla 9 Ponderaciones, ámbitos, indicadores y variable componente Grado de Integración*).

El nivel de acceso que tenga un territorio a un conjunto de servicios mínimos para desarrollar cualquier actividad, es un factor determinante en la definición de su integración. La selección de servicios, depende de la disponibilidad de datos oficiales y disponibles para la totalidad del país. Estas condicionantes disminuyen considerablemente la cantidad de variables que pueden ser utilizadas.

El componente Grado de Integración está definido como un índice que es producto de una ponderación entre los tres ámbitos.

$$\text{Grado Integración} = 0.3 * \text{Educación} + 0.4 * \text{salud} + 0.3 * \text{Centro proveedores de bienes y servicios}$$

Tabla 9 Ponderaciones, ámbitos, indicadores y variable componente Grado de Integración.

Ámbito	Ponderador ámbito	Sub-ámbito	Ponderación Sub-ámbito	Indicador	Ponderación desagregada por Indicador	Variable
Educación	30			Acceso a Establecimiento de enseñanza Básica	14	Horas al establecimiento más cercano.
				Acceso a Establecimiento de Enseñanza Media	14	Horas al establecimiento más cercano.
				Acceso a Establecimiento de enseñanza Parvularia	2	Horas al establecimiento más cercano.
Salud	40	Acceso Establecimientos de atención primaria	11	Acceso Establecimientos de atención primaria	11	Horas al establecimiento más cercano según su Servicio de Salud.
		Acceso de Urgencias	29	Acceso a cualquier establecimiento de urgencias (Hospital de cualquier complejidad, Posta y SAPU)	11	Horas al establecimiento más cercano según su Servicio de Salud.
				Acceso Hospital de Baja Complejidad	3	Horas al establecimiento más cercano según su Servicio de Salud.
				Acceso de Hospital Mediana Complejidad	6	Horas al establecimiento más cercano según su Servicio de Salud.
Acceso de Hospital Alta Complejidad	9	Horas al establecimiento más cercano según su Servicio de Salud.				
Centro Proveedores de Bienes y Servicios	30			Acceso a Centro proveedor de servicios con al menos 1 banco	6	Horas a lugar (sede comunal) a la institución bancaria más cercana.
				Acceso a Centro proveedor de bienes y servicios con 2 o 3 sucursales bancarias	9	Horas a lugar (sede comunal) con 2 o 3 instituciones bancarias más cercana.
				Acceso a Centro proveedor de bienes y servicios con más de 4 sucursales bancarias	15	Horas a lugar (sede comunal) cuatro o más instituciones bancarias más cercana.

Fuente: Elaboración propia.

El nivel de acceso es medido por el tiempo de desplazamiento de las localidades (georreferenciadas) a cada uno de los tipos de servicios (georreferenciados) véase 2.3.3 *Implementación de información base común en SIG.*. En muchos casos los servicios tienen jurisdicción territorial, es decir, tiene asignado un lugar definido donde ir. Esto significa modificar la programación del algoritmo de Dijkstra, con el objeto de medir el tiempo del desplazamiento entre los nodos (localidades) y los nodos de destino (servicio con jurisdicción, georreferenciado). La solución de este problema no es trivial, debido a que se debe construir una matriz en base de datos que permita realizar las restricciones de jurisdicción, tal que acote el conjunto de solución a la jurisdicción territorial de los servicios. Este cálculo respetando la jurisdicción, por las dificultades descritas, su procesamiento y obtención de resultados demoran más que la solución por ruta mínima convencional.

2.5.1 Ámbito Educación.

El ámbito de Educación se divide en tres indicadores: acceso a establecimientos de enseñanza Básica; establecimientos de enseñanza Media y, establecimientos de enseñanza Parvularia. La ponderación asignada a cada una de estas variables se basa en el reconocimiento constitucional de los 12 años de educación obligatoria, lo que implica que el acceso a establecimientos de enseñanza primaria tiene el mismo peso relativo que el acceso a establecimientos de enseñanza secundaria. La baja ponderación de la educación Parvularia, se debe a que cuenta con un nivel de cobertura mucho menor y que no es de carácter obligatoria.

De manera similar a lo ocurrido en la construcción de la cobertura de Localidad, existen establecimientos que imparten más de un tipo de enseñanza. En este caso se debe generar en la base de datos de nodos (véase 2.3.3 *Implementación de información base común en SIG.*) que permitan realizar cada uno de los cálculos separados por tipo de enseñanza impartida.

2.5.2 Ámbito Salud.

Al igual que en el Ámbito Educación, existen establecimientos de salud que imparten más de un tipo de servicios, se debe generar en la base de datos de nodos (véase 2.3.3 *Implementación de información base común en SIG.*), que permitan realizar cada uno de los cálculos separados para los distintos tipos de prestaciones de salud.

El ámbito salud se divide en dos sub-ámbitos: acceso a atención primaria y acceso a establecimientos de salud de urgencias (este sub ámbito se divide en cuatro indicadores).

Para el caso de acceso atención primaria se mide el tiempo de desplazamiento de la localidad a los establecimientos de salud, según la siguiente tipología de establecimiento: convenio de atención cerrada (CAC), clínica en convenio (CCONV),

centro de diagnóstico terapéutico (CDT), centro de referencia de salud (CRS), centro comunitario de salud familiar (CECOF), centro de salud familiar (CESFAM), consultorio de salud mental (COSAM), centro de salud urbano (CSU), consultorio de salud rural (CSR) y posta salud rural (PSR). El valor que toma la variable es el menor tiempo, según servicio de salud al que pertenece la localidad. Por lo tanto, no se trata de un indicador ponderado de los 10 tipos de establecimientos mencionados, si no que el que se encuentra a menor tiempo de la localidad, respetando la jurisdicción del servicio de salud³⁴ al que pertenece la localidad. Este indicador tiene una alta ponderación dentro del ámbito salud (se plantea como un sub-ámbito) debido a que es el que la población recurren con mayor frecuencia.

Para el acceso a urgencias, el indicador acceso a cualquier establecimiento de urgencias opera de la misma manera que la atención primaria. El valor que toma la variable es el menor tiempo entre cualquier tipo de hospital, posta y SAPU. Para el caso de los indicadores acceso a hospitales de alta, media y baja complejidad, se evalúan por separado, pero reconociendo la lógica del diseño jerárquico-territorial del sistema de salud nacional. Se plantea que a mayor tiempo de desplazamiento al sistema en general, menor será la integración de la población. Sin perjuicio de lo anterior, se está más integrado, cuando un habitante está menos distante de un hospital de mayor complejidad, ya que esté podrá cubrir más prestaciones que las otorgadas por hospitales de menor complejidad. De manera inversa, un habitante que tiene acceso cercano a centros de menor complejidad estará más aislado que aquel que tiene acceso cercano a centro de mayor complejidad.

En ambos casos se respeta la jurisdicción de los servicios de salud.

2.5.3 Ámbito Acceso a Centro Proveedor de Bienes y Servicios.

Por último, el ámbito de acceso a Centro Proveedor de bienes y servicios, se divide en tres indicadores: acceso a lugares con al menos una sucursal bancaria; acceso a lugares con 2 ó 3 sucursales bancarias; y acceso a lugares con más de 4 sucursales bancarias. La presencia de instituciones financieras compitiendo por la demanda de servicios financieros, es un indicio del dinamismo económico y tamaño de mercado de bienes al cual puede acceder una localidad. Se debe reconocer la lógica de que a mayor número de instituciones bancarias existe un mayor mercado, se plantea que a mayor tiempo de desplazamiento al sistema en general, menor será la integración de la población. Sin perjuicio de lo anterior, se está más integrado cuando un habitante está menos distante de un mayor número de instituciones bancarias. Dado esto, para efecto de cálculo de horas de viaje, se asignan ponderaciones distintas a los centros, según su número de instituciones bancarias.

Se asume que un habitante que se encuentra a menos tiempo de un número mayor de instituciones bancarias, que de un número menor de instituciones bancarias, recurrirá

³⁴ El territorio y las sedes de los servicios de salud, es establecido por un decreto supremo del Minsal, por orden del presidente.

al primero y no será “perjudicado” por el mayor tiempo de desplazamiento del segundo. Para resolver el caso anterior, en un caso como éste el mayor desplazamiento al centro de menor complejidad se homologa al de menor tiempo, y solo se aplican las ponderaciones correspondientes.

La información de instituciones financieras fue georreferenciada a nivel de sede comunal, lo que implica una escala aproximada de 1:100.000, que se considera suficiente para este nivel de análisis. Para el caso de los establecimientos de educación y establecimientos de salud, es información georreferenciada oficial emanada de los respectivos ministerios (Mineduc y Minsal).

2.6 Componente Condiciones Geográficas Estructurales.

El Aislamiento se considera como una situación desventajosa para una comuna, los factores que lo determinan son principalmente físicos y demográficos, con leves variaciones en el tiempo y que son los elementos estructurantes de una situación de aislamiento.

El supuesto metodológico que sustenta esta componente, es que el territorio tiene “características” o “condiciones” que lo hacen aislado. Esta característica es bastante determinista, y su modificación en el tiempo es escasa o poco probable.

Lo anterior considera que la acción del Estado tiene pocas probabilidades de modificar una situación de aislamiento de una comuna.

La División de Políticas y Estudios (SUBDERE) consideró para el estudio Diagnóstico de Territorios Aislados 2011, pertinente medir los elementos físicos y demográficos, disponibles en bases de datos digitales, que inciden en el aislamiento de un territorio, que fueron agrupados en dos ámbitos: Condiciones Físicas Ambientales y de Centralidad; y Acceso a Centro Político Administrativo (ver *Tabla 10. Condiciones Geográficas Estructurales*).

En la *Tabla 10. Condiciones Geográficas Estructurales*, mide el conjunto de indicadores que condicionan a un determinado territorio en relación al resto del territorio.

Tabla 10. Condiciones Geográficas Estructurales

Ámbito	Ponderador ámbito	Indicador	Ponderación desagregada por indicador	Variable
Condiciones Físicas Ambientales y de Centralidad	50	Nivel de habitabilidad	23	Índice de Habitabilidad
		Acceso a Ciudad Principal	27	Horas a la ciudad principal más cercana.
Acceso a Centro Político Administrativo	50	Acceso a Capital Regional	15	Horas a Capital Regional
		Acceso a Capital Provincial	5	Horas a Capital Provincial
		Acceso a Sede Comunal	30	Horas a Sede Comunal

Fuente: Elaboración propia.

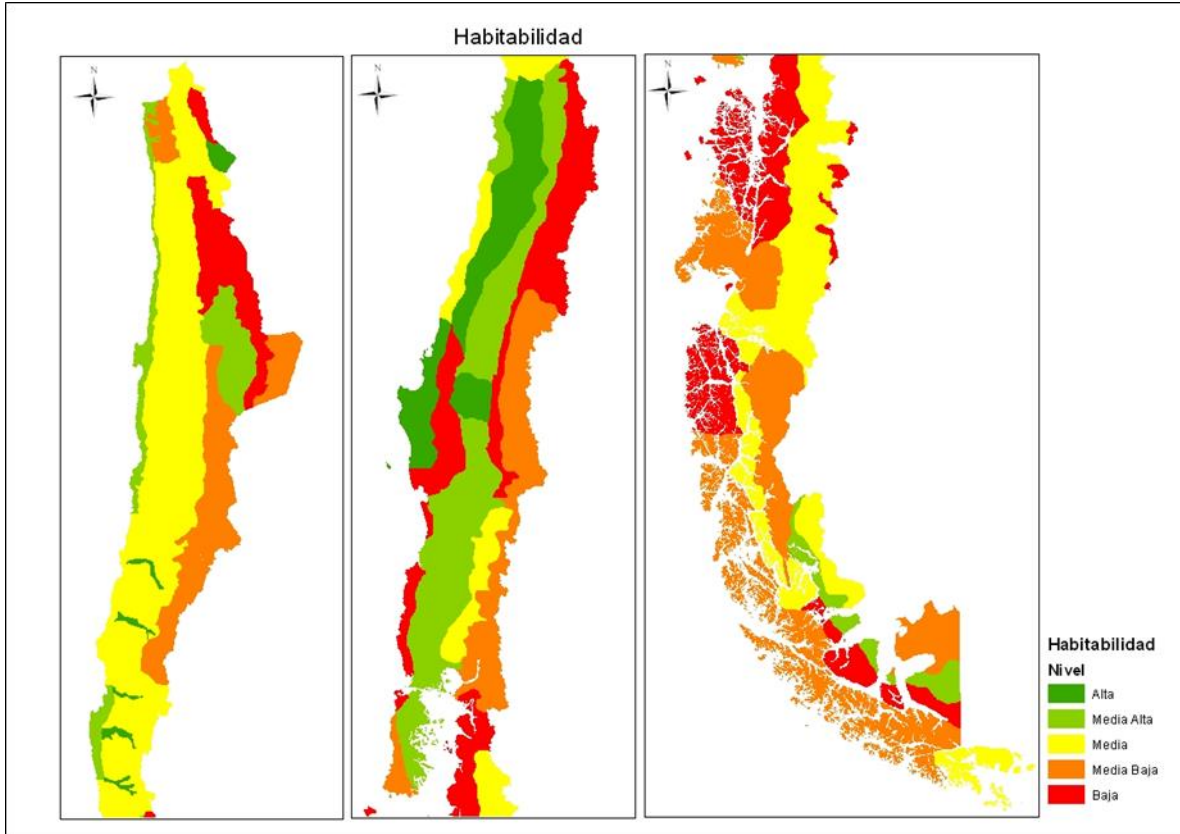
2.6.1 Condiciones Físicas, Ambientales y de Centralidad.

El ámbito Condiciones Físicas Ambientales y de Centralidad está compuesto por dos ámbitos: Nivel de Habitabilidad y Accesos de Ciudades Principales.

El Concepto de Habitabilidad se entiende como “el conjunto de condiciones favorables y adversas que le dan al espacio potencialidades y restricciones para ser ocupado”. El Nivel de Habitabilidad fue determinado espacialmente por la Escuela de Geografía de la Universidad Católica de Valparaíso con motivo del Estudio “Metodología para la identificación de Territorios Especiales”³⁵. Uno de los productos de este estudio fue la determinación áreas según *Ilustración 4 Niveles de Habitabilidad*.

³⁵ SUBDERE-PUCV 2002.

Ilustración 4 Niveles de Habitabilidad.



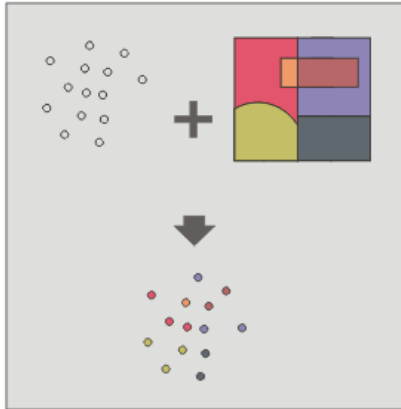
Fuente: Elaboración propia en base información digital en base al estudio "Metodología para la identificación de Territorios Especiales" SUBDERE-PUCV 2002.

El dato de habitabilidad está disponible para todo Chile en una cobertura geográfica en formato ESRI shape file .shp. Para capturar el dato, el procedimiento es que a cada localidad se agrega la característica climática a través de una unión o enlace espacial de capas de información (*spatial join*), en el Software gvSIG³⁶ 11.1, logrando así atribuir las localidades con la característica de habitabilidad del lugar.

El enlace espacial consiste en un geoproceto, que permite transferir los atributos de una capa a otra en base a una característica común. A diferencia del *join* de las bases de datos relacionales, en este caso la característica común no es un que un campo de las dos tablas tome el mismo valor (la clave del *join*), sino que los elementos relacionados de las dos capas cumplan unos criterios espaciales *Ilustración 5 Esquema Enlace Espacial*.

³⁶ Proyecto de desarrollo de SIG en software libre impulsado por la Conselleria de Infraestructuras y Transporte de la Generalitat Valenciana.

Ilustración 5 Esquema Enlace Espacial.



Fuente: Manual gvSIG 1.11.

El geoproceso **Enlace Espacial** implementado por la extensión de geoprocesamiento de gvSIG, permite seguir dos tipos de criterios espaciales para establecer el enlace espacial:

1. **Vecino más próximo (relación 1->1)**. Asigna a un elemento de la capa origen los atributos del elemento más próximo de la capa enlazada. En el caso de que el elemento más próximo interseccione (o esté contenido para el caso de polígonos) al elemento original, habiendo por tanto varias intersecciones, el algoritmo tomará el primer elemento analizado de las posibles intersecciones.
2. **Contenido en (relación 1->M)**. Relaciona un elemento de la capa origen con varios elementos de la capa destino (en concreto, con aquellos que son intersectados). En este caso la capa origen no heredará los atributos de la capa relacionada, sino que la operativa será muy parecida a la del geoproceso *Disolver*. Para los M elementos relacionados con un elemento de la capa origen, se dará al usuario la posibilidad de escoger una o varias funciones resumen (media, mínimo, máximo, sumatorio) que se aplicarán sobre los atributos numéricos de la capa enlazada.

El acceso a la Ciudad Principal se fundamenta a que en ella existe una serie de servicios de primera necesidad y ocio.

El acceso a los servicios de primera necesidad y de ocio, es un elemento esencial para determinar los grados de aislamiento. El supuesto metodológico es que a mayor dificultad de acceso a los centros urbanos, que proveen una serie de servicios complementarios a los provistos por el Estado, mayor es el grado de aislamiento del territorio.

El acceso se mide calculando el menor tiempo de desplazamiento desde las localidades a la ciudad principal más cercana *Tabla 11 50 Ciudades con más población en Chile*.

Tabla 11 50 Ciudades con más población en Chile.

Ciudad	Región	Población (Censo 2002)
Santiago Metropolitano³⁷	Metropolitana	5.631.839
Concepción Metropolitano³⁸	Biobío	848.023
Valparaíso Metropolitano³⁹	Valparaíso	824.006
La Serena - Coquimbo	Coquimbo	296.253
Antofagasta	Antofagasta	285.255
Temuco - Padre Las Casas	Araucanía	268.437
Rancagua⁴⁰	O'Higgins	236.363
Iquique - Alto Hospicio	Tarapacá	214.586
Talca	Maule	208.907
Arica	Arica y Parinacota	175.441
Puerto Montt - Puerto Varas	Los Lagos	175.140
Chillán - Chillán Viejo	Biobío	165.528
Los Ángeles - Nacimiento	Biobío	138.856
Calama	Antofagasta	136.600
Copiapó - Tierra Amarilla	Atacama	134.561
Osorno	Los Lagos	132.245
Quillota - La Calera⁴¹	Valparaíso	128.874
Valdivia	Los Ríos	127.750
Punta Arenas	Magallanes	116.005
San Antonio - Cartagena - Sto. Domingo⁴²	Valparaíso	106.255

³⁷ Incluye la provincia de Santiago, Puente Alto, San Bernardo, Peñaflores, Buin Lampa, Colina, Padre Hurtado, Bajos de San Agustín, Batuco, Pirque, El Principal, Lo Herrera, Alto Jahuel, Champa, Pintué-La Guachera, Estación Colina y Viluco.

³⁸ Incluye Concepción, Chiguayante, Hualpén, Talcahuano, San Pedro de la Paz, Coronel, Lota, Penco y Tomé.

³⁹ Incluye Valparaíso, Viña del Mar, Concón, Quilpué, Villa Alemana y Placilla de Peñuelas.

⁴⁰ Incluye Machalí, Gultro y Los Lirios.

⁴¹ Incluye La Cruz e Hijuelas.

Ciudad	Región	Población (Censo 2002)
Curicó	Maule	104.124
Ovalle	Coquimbo	66.405
Linares	Maule	65.130
Los Andes - San Esteban	Valparaíso	62.669
Melipilla	Metropolitana	53.522
San Felipe	Valparaíso	50.865
Talagante	Metropolitana	49.957
San Fernando	O'Higgins	49.519
Limache - Olmué	Valparaíso	45.327
Coihaique	Aisén	44.850
Angol	Araucanía	43.801
Vallenar	Atacama	43.750
La Unión - Río Bueno	Los Ríos	40.669
Constitución	Maule	33.914
Rengo	O'Higgins	30.891
Cauquenes	Maule	30.771
Curanilahue	Biobío	30.126
San Carlos	Biobío	29.359
Castro	Los Lagos	29.148
Villarrica	Araucanía	27.408
Ancud	Los Lagos	27.292
Molina	Maule	27.203
Parral	Maule	26.397
Quintero - Las Ventanas	Valparaíso	24.676
Victoria	Araucanía	23.977
Tocopilla	Antofagasta	23.352
El Monte	Metropolitana	22.284
Illapel	Coquimbo	21.826
Mulchén	Biobío	21.819
Graneros	O'Higgins	21.615

Fuente: Observatorio Urbano MINVU.

⁴² Incluye Las Cruces y Las Brisas

Para el cálculo de los tiempos se utilizará el Sistema de información Geográfico *TransCad* (véase 2.3.3 *Implementación de información base común en SIG*). Los resultados se ordenarán en rangos, se estandarizarán y transformarán en un indicador para ser utilizado en el cálculo.

2.6.2 Acceso a Centros Político Administrativo.

El ámbito Acceso a Centros Político Administrativo se compone de tres indicadores: Accesos a Capital Regional; Acceso a Capital Provincial y Acceso a Sede Comunal. El acceso a la sede comunal tiene una ponderación mayor, porque al medir el aislamiento a escala local, el nivel de gobierno más cercano a la población adquiere un nivel de importancia mucho mayor, ya que es la primera instancia del Estado que entrega soluciones a la población. Adicionalmente, se considera que logra discriminar de mejor manera la situación de aislamiento de las diferentes localidades.

2.7 Implementación visor web de resultados.

La implementación de un visor de mapas web permitirá publicar los resultados de este estudio en una plataforma informática, a la cual se puede acceder desde la intranet institucional o internet, posibilitando el acceso universal a la información y evitando la instalación de un sistema de información geográfico en cada equipo que lo requiera.

La posibilidad de mostrar los resultados del estudio sobre mapas comerciales (como Google Maps, Bing o Yahoo), permite también revisar la cartografía ocupada para realizar los cálculos, pudiendo identificar errores de tipificación de carpetas (encontrar una carretera donde la red de interconexión identifica una huella, por ejemplo), segmentos de caminos, senderos o huellas sin digitalizar, o errores de georeferenciación en localidades, centros de salud y establecimientos educacionales.

2.7.1 Construcción de visor de resultados y complemento con cartografía web e imágenes Goglemaps.

Para la implementación del visor web con los resultados del estudio, se han considerado la simplicidad de uso y la visualización de mapas de Google, como los principales requerimientos funcionales. En base a estos requerimientos, como al uso de software libre, se eligen las siguientes tecnologías para el desarrollo del visor:

GeoServer:

Este software permite la publicación y edición de mapas desde distintas fuentes, siendo la referencia en los estándares definidos por *Open Geospatial Consortium*

(OGC⁴³) para la creación de servidores web de atributos o *WFS*, que permiten la edición y análisis de los mapas servidos a través de un servidor web de mapas o *WMS*.

Mediante *GeoServer* se implementa un *WMS*, quien responde las solicitudes de datos con imágenes, de manera que puedan mostrarse en un navegador web. La solicitud se realiza mediante una dirección de internet o URL, donde se especifican los parámetros para los datos requeridos, como el sistema de coordenadas, la extensión espacial, el formato de la imagen de salida, la capa de dónde se deben extraer los datos, entre otros. Las fuentes de datos podrán estar tanto en archivos cartográficos (shapefiles) como en bases de datos geográficas.

Postgres – Postgis:

Postgres es un sistema de gestión de base de datos relacionales libre y que posee una extensión para el manejo de datos geográficos, *Postgis*. Esta extensión permite guardar capas alfanuméricas y geográficas con su referencia espacial asociada en tablas indexadas espacialmente, que agilizan el procedimiento de acceso a los datos espaciales.

OpenLayers:

OpenLayers es una interfaz de programación para aplicaciones web, que permite mostrar mapas interactivos desde distintas fuentes de datos geográficos en un navegador web. Entre las fuentes que soporta, están los servicios de mapas web (*WMS*) y servicios de mapas digitales comerciales como Google Maps, Bing y Yahoo!, ajustándose a las fuentes de datos disponibles y a los requerimientos de software.

2.7.2 Implementación.

La aplicación se ha diseñado con una estructura de 3 capas que se relacionan entre sí. La definición de cada capa es la siguiente:

1. Interfaz de Usuario: es la única capa de la arquitectura que tiene contacto con el usuario y en esta implementación fue programada con javascript⁴⁴, usando las librerías de *OpenLayers* e insertas en una página web mediante html⁴⁵.
2. Servidor de Aplicaciones: lugar donde se alojan los servicios de mapas publicados a través de *Geoserver* y los servicios que permiten responder a peticiones web.

⁴³ OGC: Open Geospatial Consortium, agrupación de entidades públicas y privadas cuyo fin es definir estándares para sistemas de información geográfica y la world wide web, que permitan la interoperabilidad de los sistemas de geoprocésamiento.

⁴⁴ Lenguaje de programación usado principalmente en desarrollos web del lado cliente, es decir, se interpreta en el navegador del usuario y no en el servidor.

⁴⁵ HyperText Markup Language, lenguaje de etiquetado de hipertexto usado en el desarrollo de páginas web.

3. Almacén de datos: capa que contiene los datos geográficos y alfanuméricos que son accesibles a través de los servicios web, y que en esta implementación es representada por un servidor *Postgres - Postgis*.

El esquema *Ilustración 6* del diseño anterior (con las relaciones entre las 3 capas) se muestra a continuación.

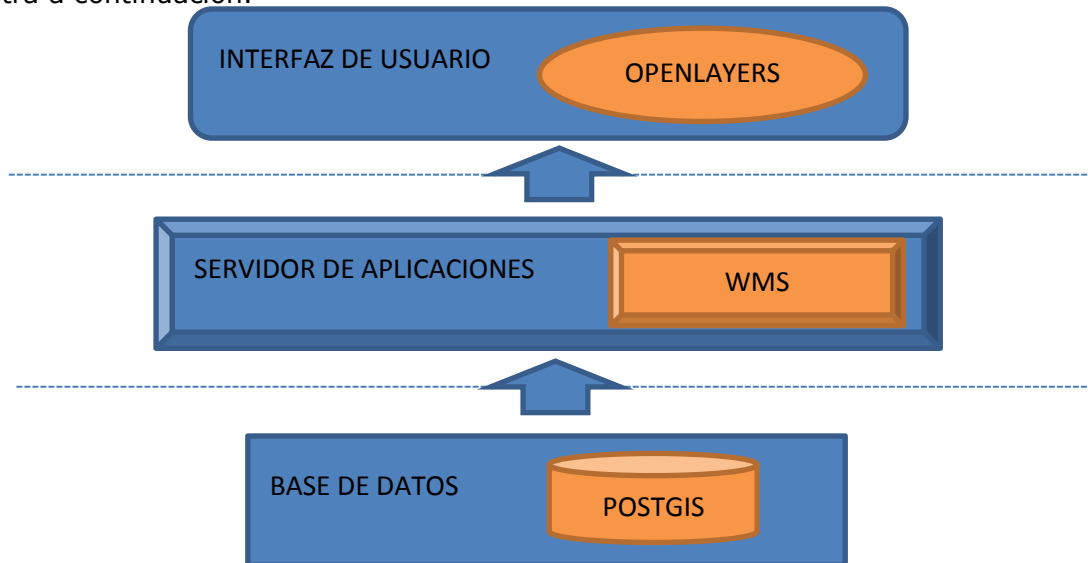


Ilustración 6
Fuente: Elaboración propia.

En este caso particular, tanto el servidor web, los servicios de mapas y el servidor de datos se encuentran en una misma máquina, con las siguientes características:

- CPU Intel Xeon ® X6670 @ 2.93GHz
- 8 GB ram
- Sistema Operativo Ubuntu Server 10.04 LTS de 32 bits
- Postgis 1.5.2
- Geoserver 2.1.3
- OpenLayers 2.11
- Styler 1.7.3 (usado en equipo de pruebas para generar archivos de estilo para los mapas)

A partir de los datos de localidades almacenados en Postgis (capa de datos), se genera una vista⁴⁶ que considera solo los atributos relevantes para el usuario⁴⁷

Tabla 12

⁴⁶ Vista: en base de datos se define como una consulta que puede incluir una o varias tablas. A pesar que en la base de datos solo se guarda la definición de la vista y no sus datos resultantes, estos pueden ser leídos como si se tratara de una tabla común

⁴⁷ Postgis agrega además un identificador de la geometría del elemento (GID) y OpenLayers agrega un identificador del elemento geográfico (FID).

Tabla 12

TERRITORIO	Identificador de la localidad
MACROZONA	Macrozona
REGION	Identificador de la región
COMUNA	Nombre de la comuna
LOCALIDAD	Nombre de la localidad
POBLACION	Número de habitantes de la localidad
I_INTEGRAC	Índice de integración
I_AISLAM_E	Índice de aislamiento estructural
I_AISLAMIE	Índice de aislamiento
RANK_MACRO	Ranking de la localidad dentro de la macrozona
RANK_REGIO	Ranking de la localidad dentro de la región
COMUNA_AIS	Atributo para identificar si la localidad es aislada según el estudio comunal ⁴⁸
COMO_LOC_AI	Atributo para identificar si la localidad es aislada por estudio comunal o por localidad

Fuente: Elaboración propia.

A través de *GeoServer* se realiza la publicación de la capa en un wms, de forma que pueda ser consultada a través de *OpenLayers*. Se publican también las capas de red de interconexión, límite comunal y sede comunal, que en conjunto con la capa de localidades y los mapas comerciales (Google, Bing, Yahoo, etc), forman el mapa de localidades aisladas que muestra el visor.

La definición de la simbología, colores, filtros por escala y por valores de atributo para cada mapa, se realiza a través de código SLD⁴⁹ (Styled Layer Descriptor) mediante reglas de estilo. Se ha tomado como base el uso de la aplicación *Styler* de *Geoserver* que entrega una interfaz para facilitar la creación de un archivo de estilos (con extensión SLD). *Geoserver* administra dichos estilos, permitiendo su edición (en código) y asociación a una capa en particular.

⁴⁸ Estudio de Identificación de Territorios Aislados 2011

⁴⁹ SLD: codificación estándar para simbolización y colorido de elementos geográficos o coberturas definidos por el usuario.

Las capas publicadas en el WMS son leídas desde una página web, usando javascript y las librerías de *OpenLayers*. En este punto se cargan dentro de la página web todas las capas requeridas para generar el mapa, añadiendo, entre otros parámetros, las coordenadas de la extensión espacial del mapa y zoom inicial, la proyección del mapa, el formato de salida del mapa⁵⁰, posibilidad de consulta sobre el mapa, y leyendas.

La ventana que muestra el mapa es definido a través de html, pudiendo modificarse su ubicación en la pantalla, tamaño y estilos asociados a un CSS⁵¹.

La *Ilustración 7* muestra el visor web con los resultados del estudio de localidades aisladas, sobrepuestos a mapas de Google.

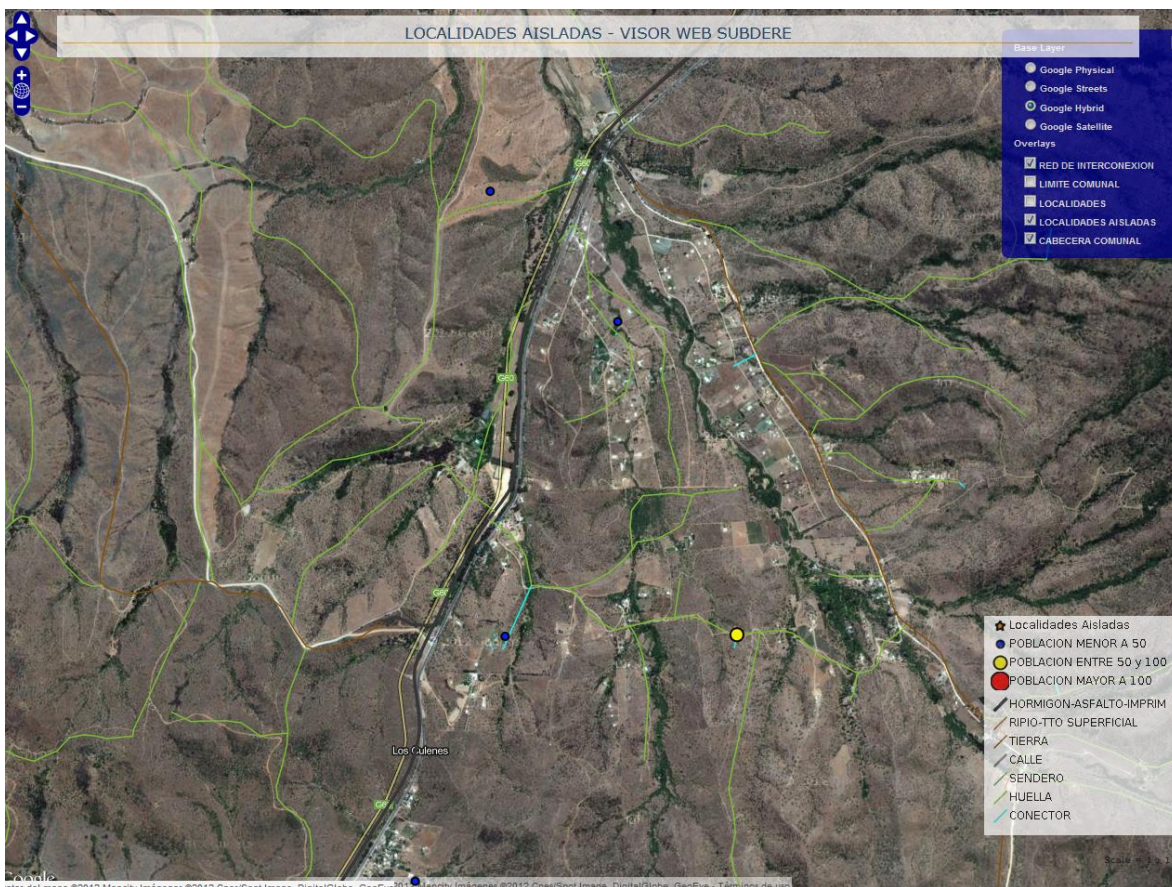


Ilustración 7

Fuente: Elaboración propia.

⁵⁰ El WMS responde a la petición desde el navegador del usuario con una imagen con la capa consultada. Las capas son superpuestas por OpenLayers, generando una imagen final que puede tener un formato distinto del definido por el WMS.

⁵¹ CSS (Cascading Style Sheets) es un lenguaje que permite definir el estilo y presentación de una página web escrita en HTML.

2.8 Mecanismo de Verificación.

De las 36.053 localidades, se procede a filtrar o seleccionar al conjunto de localidades que cuentan con un índice de aislamiento inferior a 0, que corresponde a aquellas que se encuentran en "condición de aislamiento".

Para la revisión de resultados se proponen 3 mecanismos de verificación de los resultados. Sin embargo, de ellos se aborda solamente el primero en este estudio. Se sugiere que los otros dos mecanismos deben ser asumidos por las instituciones regionales o locales.

El primer mecanismo de verificación de resultados se basa en la hipótesis que en base a la ponderación de acceso a sedes comunales (30%), debería haber un número reducido de localidades en condición de aislamiento en las zonas cercanas a las sedes comunales. Al encontrarse cerca de las sedes comunales, que cuentan con una ponderación muy alta en el componente de integración, podríamos estar en presencia de posibles errores. Por ello se genera un mapa digital con las localidades filtradas con un índice de aislamiento inferior a "0". Haciendo uso de Sistemas de Información Geográficos, se generan áreas de influencia en torno a todas las Sedes Comunales de 3 y 5 km.

Posteriormente, se sobreponen las localidades filtradas (localidades en condición de aislamiento) con las áreas de influencia de las Sedes Comunales (3 y 5 km). Aquellas localidades que se encuentran en el área de influencia son analizadas individualmente, con la ayuda de los servidores de mapas que pueden mostrar imágenes satelitales, para determinar a qué motivos se debe que su índice de aislamiento sea inferior a "0".

El segundo mecanismo de verificación consiste en la generación de un área de influencia en torno a la red de caminos principales que son enrolados de 1 km, basados en la hipótesis que cerca de buenas vías de acceso no deberían haber localidades aisladas, porque esa localización asegura menores tiempos de traslado. Una vez obtenido el mapa de área de influencia de los caminos principales, se procede a sobreponer este mapa digital, con la selección de localidades con índices de aislamiento inferiores a 0 (se sugiere que sea realizado por instituciones de nivel local o regional).

El tercer procedimiento de verificación es la observación de la ubicación de las localidades aisladas, en comparación con la experiencia empírica de los profesionales de las instituciones a nivel regional o local. Esta observación con criterio experto permite determinar el grado de confiabilidad de los resultados obtenidos, usando imágenes satelitales que provee Google Earth, y visitas en terreno, en la medida que sea posible. Para este tipo de estudio, de escala local, pero a nivel nacional, es imprescindible contar con la retroalimentación de parte de los profesionales de Gobiernos Regionales y Municipalidades, que cuentan con un conocimiento mucho más acabado de la realidad.

El primer método se realiza principalmente a través de un servidor de mapa web, desarrollado por la SUBDERE, utilizando tecnologías de código abierto como *Open Layers* y tecnologías pagadas, como *Autodesk Infraestructure Map Server 2012*.

Cabe señalar que en este estudio se ha optado por no eliminar casos anómalos o extraños, pero que serán informados, para que los equipos técnicos de los diferentes gobiernos regionales, municipalidades o instituciones pertinentes, revisen las situaciones informadas para validar los datos.

2.8.1 El problema de nombres y tipo de localidades.

Como se ha mencionado en la metodología, el objeto de este estudio son las localidades. La base cartográfica que se utilizó para identificarlas es un trabajo que proviene de la Dirección de Planificación del MOP en conjunto con el INE, donde se identifican puntos representativos en el territorio de las entidades censales. A esto se agregaron puntos representativos de las capitales comunales, provinciales, regionales y las principales ciudades.

El universo total del estudio es de 36.053 localidades a lo largo de todo Chile continental (se excluye Isla de Pascua, Juan Fernández y Antártica Chilena). Sin embargo, la metodología aplicada arroja un conjunto de nombres que no representan el objeto de estudio de esta publicación (el de identificar localidades aisladas), como por ejemplo hoteles, refugios, camping, cabañas, posadas, regimientos, retenes, batallones, salmoneras, mineras, canteras, capillas, carreteras, puestos fronterizos, aserraderos y aeródromos, entre otros.

En el listado se ha identificado 209 casos de este tipo. Pero el número puede ser mayor, debido a que los nombres de los puntos identificados en muchas ocasiones no permiten distinguir adecuadamente cuándo se está en presencia de una localidad como tal o de una empresa, por ejemplo.

Con el fin de aprovechar al máximo el trabajo realizado, se ha decidido mantener la totalidad de los puntos identificados. Esto, porque puede ser de utilidad para las instituciones cuyos objetivos no sean únicamente el de localidades aisladas y requieran información sobre todos aquellos puntos donde habitan personas, independiente de la razón por la que están ahí. Sin embargo, los 209 casos descritos están claramente identificados en las tablas que acompañan los resultados del estudio, en el campo *Prob_nom*, identificados con un número "1". El número "2" corresponde a las localidades que no representan este problema.