The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-3-2024
ISPRS TC III Mid-term Symposium “Beyond the canopy: technologies and applications of remote sensing”, 4–8 November 2024, Belém, Brazil

Assessing Chile's national expansion of urban and artificial coverages between 2000 and 2022

H. Jaime Hernández [1], Valentina González[1], Lissette Cortés[1] , Carlos Hormazábal[2], Diego Muñoz[3], Miquel Nieto[4], Alejandro Miranda[2], Rodrigo Fuentes[3], Adison Altamirano[2], Cristian Echeverría[3]

[1] Geomatics and Landscape Ecology Laboratory (GEP), Forestry and Nature Conservation Faculty, Universidad de Chile, jhernand@uchile.cl, v.gonzalez.1@uchile.cl, lissette.cortes@uchile.cl
[2] Landscape Ecology and Conservation Laboratory (LEPCON), Universidad de la Frontera,
[3] Landscape Ecology Laboratory (LEP), Universidad de Concepción,
[4] Universidad de Magallanes.

Key words: Urban expansion, Infrastructure, Chile, Regional analysis.

Abstract

This study investigates the national expansion of urban and artificial coverages in Chile between 2000 and 2022 using remote sensing techniques. The research employed Google Earth Engine and Landsat imagery to classify land use and land cover (LULC) changes. By analyzing annual cloud-free composites and employing random forest algorithms, the study produced detailed LULC maps, focusing on the 'infrastructure' class, which includes urban areas and major roads. Results indicate a significant 33% increase in infrastructure coverage, growing from approximately 380,000 hectares in 2000 to 510,000 hectares in 2022. This expansion was predominantly observed in seven regions, which accounted for 68% of the total infrastructure area and 83% of the national increase. A notable correlation (0.6) between urban growth and population growth was identified, suggesting that urban expansion is largely driven by demographic changes. However, regions with higher population growth exhibited more vertical urban development. The study highlights the complex interplay between urbanization, population dynamics, and regional planning, emphasizing the need for sustainable development strategies to manage urban growth effectively. Validation efforts showed high classification accuracies, with infrastructure class accuracies averaging 82% for the User and 87% for the Producer across the evaluated years.

1. Introduction

Urban expansion, especially when it occurs in sprawling patterns, has been extensively linked to various environmental issues. These include land pollution, deforestation, loss of agricultural lands, threats to biodiversity, and adverse effects on human health. Additionally, it exacerbates spatial inequalities through urban segregation (Viana et al., 2019). Latin America is one of the world’s most urbanised regions, with more than 80% of inhabitants living in urban areas and many cities with more than one million inhabitants (United Nations, 2014). Urban expansion in Chile has been significant over the past two decades, particularly in its major cities. Santiago, the capital, has experienced substantial urban sprawl in the last decades (Silva and Vergara-Perucich, 2021). The city's growth has led to increased land use and development, with new residential, commercial, and industrial areas emerging. Other Cities, like Valparaíso and Concepción have also seen growth, though not as pronounced as Santiago. There has been a shift in land use patterns with the conversion of agricultural and vacant lands into residential and commercial areas. This change has been more pronounced in Santiago and its surrounding areas.

There are many techniques for automatically mapping urban land cover (Jenerowicz and Bielecka, 2022). These can be broadly categorised into two types: input data classification (including pixel- and object-based methods) and index-based segmentation (like NDVI and NDBI). Recently, integrated artificial intelligence (AI) techniques have been used for change detection, improving data processing. AI, or machine learning, enables systems to interpret external data, learn from it, and adapt to achieve specific goals and tasks. This study had the objective of evaluating the increase in urban or artificial coverage, grouped in the "infrastructure" class at the national level between the years 2000 and 2002.

2. Study Area and Methods

2.1 Study Area

Chile is a long, narrow country located along the western edge of South America, stretching over 4,300 kilometres (2,670 miles) from North (17°30’ S) to South (55°59’ S). The country is administratively divided into 16 regions from North to South (Figure 2). Chile boasts diverse geography, including the Atacama Desert, The Andes Mountains running along its eastern border, and the Pacific Ocean to the west. The Central Valley, in the mid latitudes of the country, is an agricultural heartland known for its fertile soils and Mediterranean Climate. It is the place where most of the Chilean population resides and is where the major Cities of the Country are, like Valparaiso in Valparaiso Region, Santiago, the capital of the Country located in The Metropolitan Region and Concepción in the Biobio Region.

To carry out the classification, the Google Earth Engine platform and Landsat images from missions 5, 7 and 8 were used, as well as information in vector and raster format of the Chilean territory to help take training and other samples and processes. This classification was made with the objective of obtaining the first annual LULC map for Chile. The product and the full methodology can be found at chile.mapbiomas.org.

2.2 Landsat Mosaics and Feature Space

Annual Landsat cloud free composites obtained from images distributed along the whole year were constructed. The cloud/Shadow Removal script takes advantage of the quality assessment (QA) band and the Google Earth Engine Median Reducer.

When used, QA values can improve data integrity by indicating which pixels might be affected by artefacts or subject to cloud contamination (USGS, 2019). Due to a large part of the national territory approximately from the latitude of Santiago (33° 27' 33" S) to the south, the availability of the images from the Landsat program for years prior to 1998 is very low, it was decided to cover the period from 2000 to 2022. The feature space for digital classification of the categories of interest comprises a subset of 34 variables (MapBiomas Chile, 2024).

2.3 Samples and Stable Samples

Training samples were defined following a strategy of using random pixels for which the LULC remained the same throughout the entire period (23 years), so named “stable samples”. The stable areas were identified through annual preliminary classification made using random pixels selected from manually drawn polygons made by photo interpretation. For this, false colour composites of the Landsat mosaics for all the 23 years of backdrop and graphs with the temporal behaviour of spectral indices per pixel were used to establish the LULC class. The need for complementary samples was evaluated by visual inspection and by comparing the output of the preliminary classification with both Landsat and High-resolution images available in GEE.

We use a Stable Samples set collected for the original 13 classes of the LULC Chilean (MapBiomas Chile, 2024). The samples were reclassified in only two classes, the first ‘Infrastructure’ (id 24); Which corresponds to urban areas with a predominance of impermeable surfaces. This category also includes paved roads and primary roads, and the second “Others” (id 68); Which corresponds to all the others categories.

2.4 Classification

We restricted de classification geographical extent to a mask built using the Nighttime Day/Night Annual Band Composite V2.1 product available in GEE which is constructed from an annual time series of nighttime lights using monthly cloud-free radiance averages obtained from the sensor VIIRS (Elvidege et al., 2011). We used the “average_masked” that eliminates sporadic lights like northern lights, fires or isolated events. The mask was built by selecting only the areas that have a pixel value greater than 1. This threshold was chosen by iterations. The resulting area corresponds to the potential infrastructure area.

2.5 Post-processing

The final classification was improved by applying a set of filters, to correct missing data, “salt-and-pepper” effects and classification errors. First, a spatial filter was applied, which consisted of eliminating areas where the number of pixels connected is less than 6 (~ 0.5 ha) replacing the original value by the mode in the neighbourhood. The spatial filter avoids unwanted modifications to the edges of the pixel groups, this filter was built based on the “connectedPixelCount " function native to GEE, this function located connected components (neighbours) that share the same pixel value. This process is carried out independently each year.

To eliminate high-frequency variations in classification, temporal filters were applied. The temporal filter uses sequential classifications in a three-year unidirectional moving window to identify temporally non-permitted transitions. The temporal filter inspects the central position of three consecutive years (“ternary”), and if the extremities of the ternary are identical but the centre position is not, then the central pixel is reclassified to match its temporal neighbour class. Temporal filters were divided into two broad categories according to their functioning: regular and extremes (beginning and ending). While regular filters only use a 3-year window, extremes can employ a 3- or 5-years window. In this case que use filters for 3- and 4-years windows, all the filters work with the same logic. This filtering is applied interactive in order of priority, changing and modifying the noisiest class first and then, reviewing the most stable class. The filter does not modify the first and the last year of the time series.

The “Infrastructure” (id 24) map resulting from the previous steps was overlaid over full classification with 12 other classes for Chilean territory. This was necessary due to frequent confusion with non-vegetated use classes such as “other areas without vegetation” (id 25), “rocky soils” (id 29) and “sand, beach and dunes” (id 23). Before the overlay was performed, the infrastructure map was further processed. The vector map of main regional and national roads of Chile[1] was converted to raster and joined year by year with the infrastructure classification. Finally, the infrastructure final maps were superimposed with the general classified maps generating the final integrated maps.

2.6 Statistics

Zonal statistics of the mapped classes were calculated for spatial units, such as ecoregions, regions, as well as watersheds and protected areas. Once the surface areas were obtained, the annual rate of change was calculated using the following formula (Puyravaud, 2003).

$$r = \left(\frac{100}{t_2 - t_1}\right) \times ln\left(\frac{S_2}{S_1}\right) \quad (1)$$

Where *r* represents the percentage annual growth rate of the infrastructure class, with A_1 and A_2 being the class coverage areas at times t_1 and t_2, respectively.

2.7 Validation

Validation was performed for the classifications of the years 2002, 2012 and 2022. We used two databases of 1300 randomly generated points for each year with approximately 100 points per class. After a proper training step, two independent teams classified each point into one the classes via photo-interpretation on the Google Earth Pro platform using the available historical images. If images were not found for the year, nearby years before and after the target year were used. The classification was carried out in two rounds. In the first round, each team made up of two members classified a set of points independently. The points where the classifier's decision did not coincide were sent to the second round, where a third independent subject reclassified each point with a discrepancy.

Points where no agreement was reached were eliminated, points for which there were no images close to the target year were also eliminated.

[1] https://www.bcn.cl/siit/mapas_vectoriales

In total, 2044 points were used to validate the 2022 map, 2000 points to validate the 2012 map and 1770 points to validate 2002 map, reaching 5814 points in total. With this information, the confusion matrices were generated and the classification accuracies were calculated. All these processes were carried out in Google Earth Engine and R-project.

3. Results and Discussion

The results indicate that the surface area associated with infrastructure and urban uses increased from ~380,000 hectares in 2000 to ~510,000 hectares in 2022 (Figure 1). In other words, infrastructure land cover class increased by 33% in 22 years (127,000 ha) with an annual change rate of 1.3%.

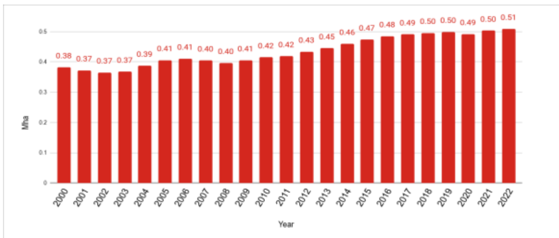

Figure 1: Total annual national area (ha) of the Infrastructure class.

The analysis at Regions level reveals that in 2022, 68% of this infrastructure is concentrated in seven of them (Table 1, Figure 2). These Regions, therefore, have been the main drivers of urban growth at the national level, representing 83% of the total national increase in urban area in the period analysed.

Region	Area (ha)		Increase		Annual rate of change (%)	
	2000	2022	ha	%	Area	Pop[2]
Metropolitana	91,072	123,217	32,145	35	1.4	1.4
Valparaíso	43,802	632,96	19,493	45	1.7	1.1
Coquimbo	38,162	57,223	19,060	50	1.8	1.6
O' Higgins	27,284	39,339	12,055	44	1.7	1.1
Maule	23,518	33,374	9,856	42	1.6	1.0
Antofagasta	20,757	28,637	7,880	38	1.5	1.72
Biobío	30,201	34,964	4,763	16	0.7	0.61
National	**382,300**	**509,040**	**126,740**	**33**	**1.3**	**1.17**

Table 1: Changes in the Infrastructure land cover class, and Population (Pop.) annual rate change from 2000 to 2022 in the seven regions with the highest proportion of change. Note that the national total includes 16 regions and not just the 7 presented in the table.

The annual rate of change values, in area and total population by Region, can be used to establish the relationship between the growth of urban areas and population growth. The simple correlation between both variables, close to 0.6, indicates that a significant part of the growth of urban areas and infrastructure is explained only by the growth in the number of inhabitants. When the population growth rate is lower than the growth rate of infrastructure areas, as in the case of the Regions of Valparaíso, Coquimbo, O'Higgins, Maule and Biobío, an expansion of urban areas is observed with a tendency to grow more horizontally than in vertically. The opposite occurs in the Antofagasta Region. For the Metropolitan Region, which contains the capital Santiago, a strong linear behaviour is observed between both rates of change, both above the average values at the national level. Nevertheless, the relationship between population growth and city size is influenced by factors such as urbanisation, economic opportunities, infrastructure, government policies, and cultural dynamics (Huang and Liu, 2021). While cities often grow as people move from rural areas seeking better opportunities and services, the ability of a city to support this growth depends on its infrastructure and planning. Economic strength and effective policies can attract more people and manage growth effectively, whereas poor planning can lead to challenges. Thus, while population growth can drive city expansion, the relationship is not always direct and varies by context.

3.1 Validation

The average accuracy for the infrastructure class, for the three evaluated years, was 82% for the User and 87% for the Producer (see Table 2 for details).

CLASSES	ACCURACIES (%)					
	2002		2012		2022	
	U	P	U	P	U	P
Forest	79	80	83	81	88	70
Forest Plantation	84	70	92	75	83	87
Agricultural lands	74	88	81	91	81	87
Infrastructure	82	82	86	89	79	90
River, Lake or Ocean	98	86	98	89	97	87
Ice and Snow	85	89	76	81	74	83
Salt flat	88	87	86	86	84	89
Shrublands	68	59	62	61	62	58

Table 2: Accuracies for main classes for 2002, 2012 and 2022 land cover maps. **U** = User accuracy, **P** = Producer accuracy.

Confusion errors occurred mainly with classes related to sands and other non-vegetated areas (3,4%), shrublands (6.8%) and agricultural land (4.5%). On the other hand, omission errors occurred with the classes of grassland (1.2%), agricultural lands (2.4%), other non-vegetated land (1.2%) and shrublands (1.2%). Classification errors are more frequent in the northern part of the country, due to the characteristics of the natural coverage of those regions.

While Chile's vast latitudinal gradient and mountainous terrain present challenges for accurate land cover classification, the study provides valuable insights into national urban expansion. Geometric corrections address distortions, but radiometric effects remain a concern. Variations in illumination across slopes, atmospheric conditions, and topography-influenced spectral signatures can impact pixel brightness and spectral characteristics, potentially affecting classification accuracy.

Conclusions

From 2000 to 2022, the surface area of urban infrastructure grew by 33%, increasing from ~380,000 to ~510,000 hectares, with an annual growth rate of 1.3%. This growth was concentrated in seven regions, which accounted for 68% of the infrastructure and 83% of the total national urban area increase. A correlation of 0.6 between population growth and urban expansion indicates that population growth significantly influences urban area growth. Regions with slower population growth than infrastructure

[2] https://datosmacro.expansion.com/demografia/poblacion/chile

growth are experiencing higher horizontal expansion. In Santiago's Metropolitan Region, both population and urban growth rates are strongly linked and above the national average.

References

Breiman, L. (2001). Random forests. Machine learning, v. 45, n. 1, p. 5-32, https://doi.org/10.1023/A:1010933404324

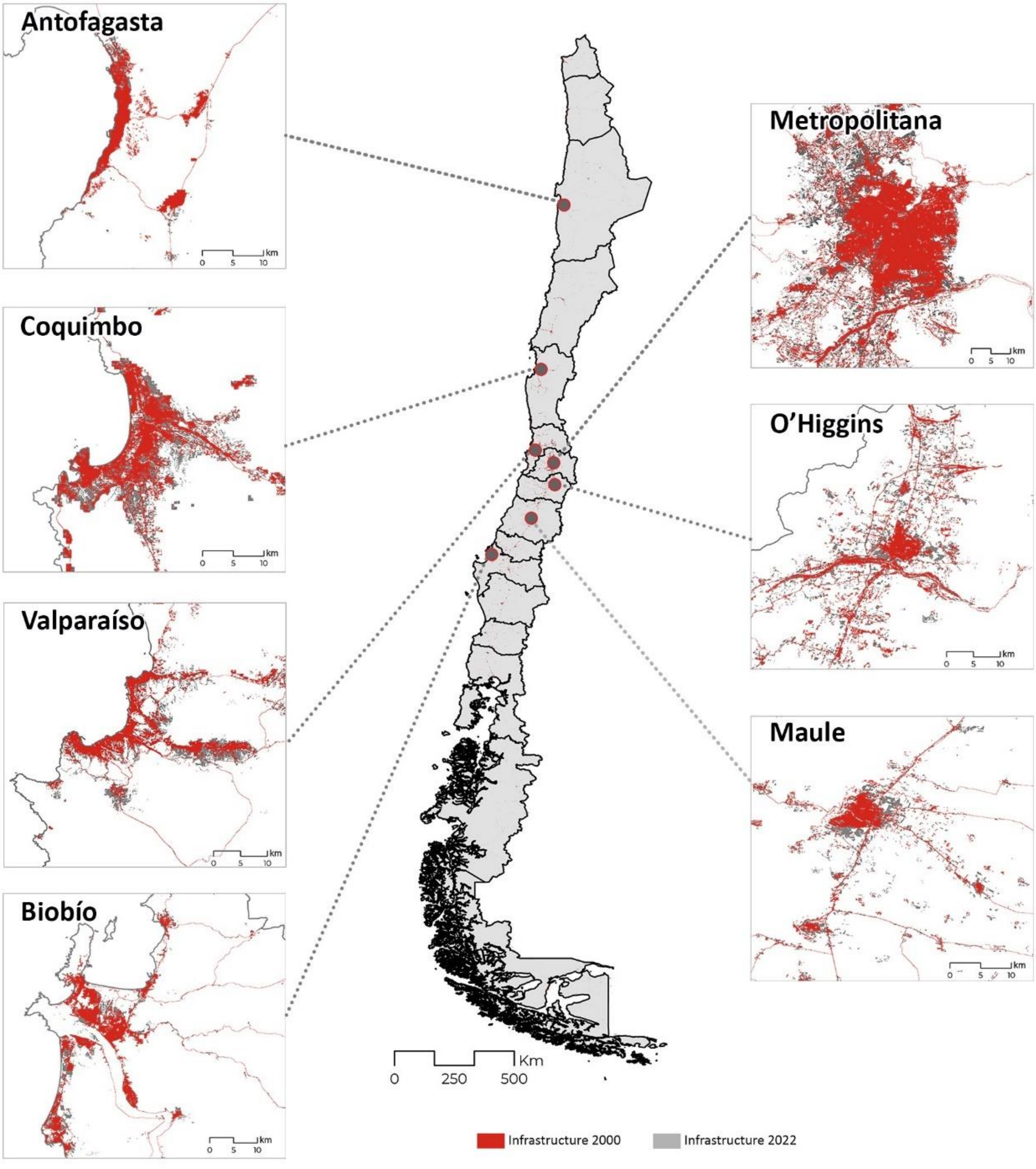

Figure 2: Regions with the highest contribution to infrastructure change between 2000 y 2022.

Elvidge, C.D, Zhizhin, M., Ghosh T., Hsu FC, Taneja J. (2021). Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. Remote Sensing, 13(5), p.922, doi:10.3390/rs13050922 doi:10.3390/rs13050922.

Jenerowicz, A. & Bielecka, E., 2022. Urban growth monitoring - Remote sensing methods for sustainable development, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B3-2022, 107–112, https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-107-2022.

Map Biomas Chile, 2024. Algorithm Theoretical Base Document and Results. Collection 1. https://chile.mapbiomas.org/wp-content/uploads/sites/13/2024/04/ATBD_Chile-Coll-1.pdf

Silva, C., Vergara-Perucich, F., 2021 Determinants of urban sprawl in Latin America: evidence from Santiago de Chile. SN Soc Sci 1, 202 (2021). https://doi.org/10.1007/s43545-021-00197-4

Viana, C.M., Oliveira, S., Oliveira, S.C., & Rocha, J., 2019. Land use/land cover change detection and urban sprawl analysis Hamid Reza Pourghasemi. In Spatial modelling in GIS and R for earth and environmental sciences. Elsevier Inc 621–651. doi:10.1016/b978-0-12-815226-3.00029-6.

United Nations (UN), 2014. World Urbanization Prospects: 2014 Revision. Department of Economic and Social Affairs, Population Division.

Puyravaud, J.P., 2003. Standardizing the calculation of the annual rate of deforestation. Forest Ecology & Management 177, 593–596. https://doi.org/10.1016/S0378-1127(02)00335-3

US Geological Survey (2019). USGS Landsat Collection 1 Level 1 Product Definition; USGS Landsat: Sioux Falls, SD, USA, 2017; Volume 26.

Huang, Q., Liu, Y. (2021). The Coupling between Urban Expansion and Population Growth: An Analysis of Urban Agglomerations in China (2005–2020). Sustainability.13, 7250. https://doi.org/10.3390/su13137250