



Universidad de Buenos Aires

2023

Maestría en Explotación de Datos y Descubrimiento del Conocimiento

**Atención Primaria de Salud en zonas rurales y
semi-rurales de la Provincia de Buenos Aires.
Análisis de cobertura y metodología de optimización**

Nicolás Rubinstein

Índice

1. Introducción	3
2. Marco teórico	4
3. Metodología	6
3.1. Datos y creación de variables	6
3.2. Métricas de evaluación	11
3.3. Optimización: algoritmos implementados	12
4. Resultados	14
5. Conclusiones y discusión	18
6. Bibliografía	20
7. Fuentes de Datos	22

Resumen

El trabajo se propone realizar una evaluación de la cobertura de la Atención Primaria de Salud (APS) en las zonas rurales y semi-rurales de la Provincia de Buenos Aires. Para esto, se consideran variables de análisis tenidas en cuenta por estudios similares para otras geografías, y se construyen tres métricas propias de cobertura, que tienen en cuenta la pobreza medida por Necesidades Básicas Insatisfechas (NBI) y la distancia geográfica a los Centros de Atención Primaria de la Salud (CAPS). Además, se realiza una simulación de selección de nuevas ubicaciones para CAPS, comparando cuatro métodos algorítmicos para maximizar la cobertura de acuerdo a las métricas construidas. Se señalan posibles mejoras en la disponibilidad de datos en particular respecto de prestaciones y capacidad de los CAPS, y la consideración de otros datos públicos como los de transporte para la aplicación de métricas más robustas. Entre otras observaciones, se concluye que la exploración de parámetros realizada para la simulación es insuficiente, pero quedan de manifiesto las potencialidades de los algoritmos genéticos para la optimización multi-objetivo o de métricas complejas como las que son habitualmente requeridas para la evaluación de cobertura sanitaria, si bien el costo computacional de este tipo de enfoques puede ser excesivamente elevado.

1. Introducción

Nos planteamos el objetivo de analizar la cobertura del subsistema de Salud de Atención Primaria en la Provincia de Buenos Aires. Este problema resulta de interés porque la atención primaria de la salud (APS) es una pieza clave del sistema sanitario. De acuerdo a la OMS (2022, 2-3), es una piedra angular de los “Objetivos de Desarrollo Sostenible” y en particular de la meta de Cobertura Universal de Salud. Entre los parámetros prioritarios de evaluación de su alcance se encuentran tanto la accesibilidad desde el punto de vista geográfico (ídem, 13) como la priorización de poblaciones vulnerables (ídem, 17). La APS cumple, de acuerdo a este marco de trabajo, un papel fundamental en la promoción, prevención, diagnóstico, tratamiento, rehabilitación y paliación.

Si bien la APS es un subsistema que debe ser ponderado en su conexión con los otros niveles de cobertura sanitaria, puesto que se trata de una primera entrada que debe ser articulada con organismos que permitan un tratamiento más complejo cuando la situación lo amerita (Testa, 1988), consideramos que a los efectos de este trabajo resulta un alcance suficiente.

A su vez, como detallaremos más adelante, los abordajes de la cuestión desde un punto de vista eminentemente geográfico han considerado conveniente la distinción entre zonas rurales o semi-rurales y urbanas para ponderar niveles de cobertura y distancias óptimas o aceptables entre los centros de asistencia y los hogares (Hong et al, 2023). En este sentido nos inclinamos por estudiar la situación en las zonas menos densamente pobladas de la Provincia, que presentan el desafío más significativo desde el punto de vista de maximizar la cobertura territorial y poblacional en una superficie sumamente extensa. Las áreas urbanas, en cambio, presentan índices de cobertura a simple vista más satisfactorios.

El interrogante que nos planteamos es doble:

- A. ¿Cómo se puede evaluar la cobertura actual de la APS en la Provincia de Buenos Aires?
- B. Y al mismo tiempo, ¿qué metodología puede aportarse desde la ciencia de datos de cara a la selección de ubicaciones geográficas para el emplazamiento de nuevos establecimientos sanitarios que permitan maximizar la cobertura?

En relación a lo primero construimos tres métricas de cobertura, que tienen en cuenta la pobreza medida por Necesidades Básicas Insatisfechas (NBI) y la distancia geográfica a los Centros de Atención Primaria de la Salud (CAPS) utilizando como unidad geográfica mínima los radios censales —entidades geoestadísticas producidas por el Instituto Nacional de Estadística y Censos de la República Argentina (INDEC) que agrupan alrededor de 300 viviendas en zonas urbanas y una cantidad menor en zonas rurales o semi-rurales (Cantar, 2019). En relación al objetivo B, generamos ubicaciones candidatas mediante un proceso aleatorio y comparamos distintos procesos de selección, que en algunos casos utilizan aprendizaje automático y en otros no, utilizando como criterio de evaluación las métricas construidas recién mencionadas.

A continuación, en la sección Marco teórico (2), llevaremos adelante un breve repaso de algunos trabajos previos que han abordado problemáticas similares en distintos contextos, y estableceremos el enfoque general con el que llevaremos adelante la tarea. En Metodología (3) haremos descripciones detalladas sucesivas de los datos utilizados, cómo fueron obtenidos, los criterios de filtrado y la creación y transformación de variables (3.1); de las métricas generadas para la evaluación de coberturas —tanto la existente como las resultantes de las simulaciones— (3.2) y de los procedimientos implementados para la optimización de selección de ubicaciones —problema B— (3.3). Luego presentaremos los Resultados (4), y, en la última sección, las Conclusiones y algunas discusiones (5) que consideramos pertinentes a la luz de nuestros hallazgos.

2. Marco teórico

La cuestión de la planificación territorial orientada por las herramientas provistas por los Sistemas de Información Geográficos (GIS, por su sigla en inglés) reconoce un largo recorrido y una relevancia creciente a la luz de los desarrollos en materia de capacidad de cómputo y accesibilidad a la tecnología y datos geográficos. Como señalaran Bosque Sendra y García (2000, 49), los problemas que este tipo de abordaje permite encarar son de dos tipos, íntimamente relacionados: descriptivos (“dónde están las cosas”) y prescriptivos (“dónde deben estar”).

Siguiendo a Mokgalaka (2014), la planificación de los sistemas de salud resulta particularmente desafiante, por la escasez de recursos y la multiplicidad de metas simultáneas a cumplir; especialmente en contextos de gran dispersión poblacional. El autor pone de manifiesto la necesidad de ponderar múltiples factores, como la accesibilidad territorial, estimando distancias de viaje y las tasas de utilización de los establecimientos. En base a ello modeló escenarios de demanda y de áreas de captación para establecimientos de la ciudad de Johannesburg. El autor realiza un análisis mayormente visual, lo que resulta habitual para este tipo de información, que plantea la necesidad y oportunidad de una sólida representación.

Otros aportes se han centrado en apuntalar con datos empíricos la noción de que la proximidad geográfica tiene mayor relevancia para las poblaciones de menos recursos (Paköz et al, 2014). Un ejemplo notable del uso de múltiples variables para el diagnóstico acerca de la accesibilidad a los servicios de Salud es desarrollado por Hong et al (2023) en torno a las áreas rurales de West Virginia, Estados Unidos. Los autores estimaron tiempos de viaje utilizando datos de la API de Google, y recogieron datos acerca del financiamiento de las instituciones de salud y sus capacidades de atención, datos socioeconómicos de la población, entre otras. El trabajo se hace eco del 2SFCA (Two-Step Floating Catchment Area), una métrica de uso extendido en el diagnóstico de accesibilidad espacial a los servicios médicos, que consolida factores de oferta, demanda y distancia. Fue desarrollada en el año 2000 (Radke et al), aunque su nombre fue acuñado posteriormente, y perfeccionada sucesivamente desde entonces (y también rebautizada como E2SFCA, con E de “Enhanced”, entre otras variantes).

Uno de los aportes más significativos al respecto es el de Wang, Gu y McGregor (2010), que además intentan resolver un problema de optimización en la selección de ubicaciones para nuevos centros de salud. Los autores se proponen entrenar un algoritmo bi-objetivo utilizando una implementación de Interchange Algorithm. Este tipo de algoritmos pueden plantearse, por ejemplo, maximizar eficiencia y cobertura simultáneamente, como en este caso.

Tomando en consideración los dos problemas que nos planteamos (evaluación de cobertura de APS en PBA, por un lado; y simulación de selección de ubicaciones para nuevos establecimientos, por el otro) intentaremos emular algunas metodologías de trabajo recogidas en este sucinto estado de la cuestión. A saber,

respecto del objetivo A (evaluación): desarrollar métricas de evaluación que combinen datos de distinto tipo (en nuestro caso, de pobreza y espaciales), ponderar los factores socioeconómicos (incidencia de la pobreza), desarrollar una representación visual lo más elocuente posible. Respecto del problema B (optimización), llevaremos adelante una comparación de múltiples modelos, uno de los cuales pertenece a la clase de Algoritmos Genéticos, que resultan populares entre otras cosas por su capacidad, también, de llevar adelante tareas multi-objetivo (si bien en nuestro caso planteamos el desafío de una métrica a la vez, pero sí aprovechamos la ventaja de poder optimizar una métrica personalizada).

Los algoritmos genéticos están inspirados en los procesos evolutivos naturales. Tienen aplicaciones en tareas de optimización en espacios de búsqueda extensos, mediante la emulación de procesos de cruce, mutación, y selección. Algunos de los parámetros más importantes para su entrenamiento son la definición del tamaño de la población (cantidad de combinaciones), cantidad de generaciones, factor de mutación y factor de cruce. Con todo, no representan una garantía, por lo que algunas indagaciones recientes se han centrado en mitigar la tendencia a la convergencia lenta y subóptima mediante la ponderación de distintos parámetros (Lim et al, 2017).

Otro de los métodos explorados es el encuadre del desafío como un caso del Problema de Ubicación de Cobertura Máxima (MCLP, por su sigla en inglés: Maximum Coverage Location Problem). Este enfoque es más limitado, ya que no tiene en cuenta factores de optimización no geográficos. Típicamente, tiene en cuenta una serie de potenciales ubicaciones, geometrías de demanda (ya sea puntos o polígonos), rango de cobertura y número de ubicaciones a establecer. Una implementación previsible para solucionarlo es la Programación Lineal Entera (PLE) (Chauan et al, 2010).

3. Metodología

3.1. Datos y creación de variables

Utilizamos datos de múltiples fuentes para componer nuestro dataset. Tomamos los datos abiertos de la Provincia de Buenos Aires para obtener las ubicaciones de los establecimientos de salud del distrito [Fuentes de datos, 1]. En estos datos encontramos múltiples categorías de establecimientos que enumeramos

a continuación: "C.A.P.S", "Hospital", "Salud Mental - Adicciones", "Geriatrico", "C.I.C.", "Unidad Móvil", "U.P.A.", "Departamento", "Centro de día", "Región Sanitaria", "Dirección", "Hemoterapia", "Instituto", "Casa de Admisión", "Hogar de Tránsito", "Emergencias", "Laboratorio", "Casa de Adolescentes", "Ablación", "Oficina", "Centro de Rehabilitación", "Posta Sanitaria", "C.A.P.S de instituciones cerradas". El dataset carece de una descripción acerca de las características, responsabilidades o alcance de cada tipo de establecimiento, pero mediante una indagación acerca de cada categoría, y si bien en algunos casos puede haber dudas respecto de su pertenencia al subsistema de atención primaria, optamos por seleccionar los dos tipos que remiten directamente a la APS: "C.A.P.S" (Centros de Atención Primaria de Salud) y "C.A.P.S de instituciones cerradas". Los primeros son 1985 y los segundos son 21. Si bien el dataset público tiene múltiples variables, como medios de contacto, distritos de pertenencia, entre otros, de cada establecimiento, a nosotros nos interesó únicamente el punto geográfico asociado y un identificador. Se trata de instituciones públicas en su totalidad. El dataset de establecimientos contiene 2840 en total (de los cuales muchos son oficinas burocráticas), lo cual refleja el gran peso del subsistema APS en el sistema de Salud de la provincia (un 70% de los establecimientos).

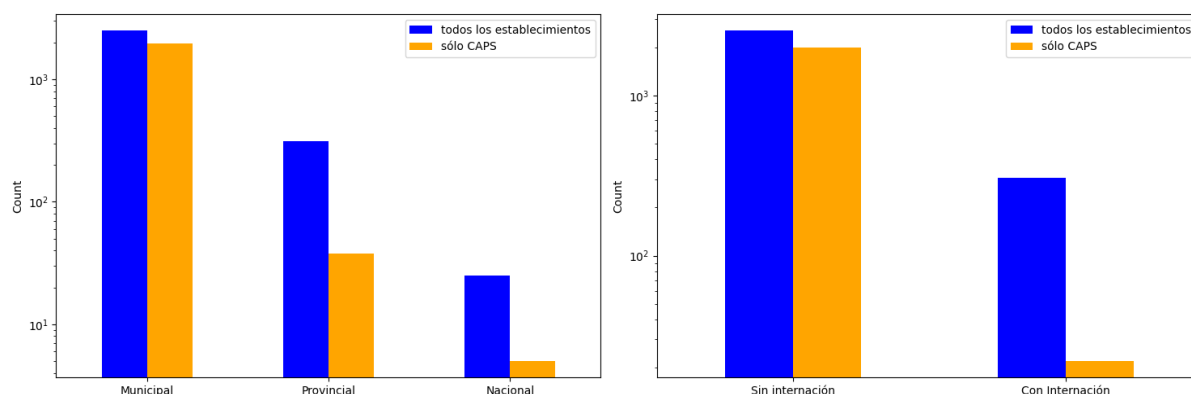


Fig. 1: Características de CAPS vs. todos los establecimientos. A la izquierda jurisdicción, a la derecha presencia de camas de internación en los establecimientos

La figura 1 muestra una comparación simple entre el conjunto de establecimientos de Salud de la provincia (barras azules) y los CAPS en particular (barras naranjas). Como se ve allí, los CAPS en su gran mayoría carecen de recursos para internación (gráfico a la derecha), y son municipales en su mayoría,

en una proporción mayor al total de los establecimientos (gráfico a la izquierda). Además, también se aprecia que la gran mayoría de los establecimientos municipales son, en efecto, centros de atención primaria.

Por otro lado, obtuvimos de la misma fuente de Datos Abiertos de la PBA los radios censales [Fuentes de datos, 2]. Este dataset geográfico tiene, además de la geoshape asociada a cada radio censal, las siguientes variables: “toponimo_i”, “link”, “varon”, “mujer”, “totalpobl”, “hogares”, “viviendasp”, “viv_part_h”, “geometry”.

Todos los datos geográficos fueron transformados al Sistema de Referencia de Coordenadas (CRS) EPSG 22176. Este sistema pertenece a la faja 6 de POSGAR (Posiciones Geodésicas Argentinas)¹. La transformación se realizó para asegurar la consistencia de los datos. Además, se seleccionó un sistema proyectado porque se buscó garantizar que la unidad de medida sea el metro —en oposición a sistemas geográficos de referencia cuya unidad es el grado, correspondiente a las coordenadas angulares de latitud y longitud— ya que esto simplifica cálculos espaciales como el cómputo de los buffers o radios de cobertura.

De aquí tomaremos la geometría, las variables “link”, “totalpobl” y “hogares”. “link” es el identificador del radio censal. “totalpobl” es la cantidad de población censada en dicho radio censal. “hogares” es la cantidad de hogares que hay en el radio. Entre las variables descartadas “toponimo_i” representa una “identificación única propia del archivo” (Cantar, 2019), “varon” a la cantidad de varones censados, “mujer” a la cantidad de mujeres censadas, “viviendasp” el total de viviendas particulares, y “viv_part_h”, el total de viviendas particulares habitadas.

Como hemos dicho, la APS está orientada mayormente a la población vulnerable. Por este motivo, consideramos fundamental contar con datos geográficos de pobreza. Para esto extrajimos del sistema REDATAM [Fuentes de datos, 1] -mediante el cual el INDEC disponibiliza información desagregada- los hogares con necesidades básicas insatisfechas (NBI) agregados por radio censal en la provincia de Buenos Aires. El dato de NBI resulta de gran relevancia porque comprende características de los hogares que están fuertemente relacionadas con

¹ El EPSG es un conjunto de códigos de referencia estandarizados, desarrollado por el European Petroleum Survey Group. POSGAR es el marco de referencia geodésico nacional de Argentina. Es un sistema de coordenadas específico para representar puntos en la superficie terrestre de Argentina, y tiene distintas versiones y fajas (zonas geográficas) para representar con precisión diferentes partes del territorio argentino, entre las cuales las fajas 4 a 6 incluyen a la Provincia de Buenos Aires. Muchos marcos de referencia geodésicos nacionales están codificados en el EPSG, de ahí la correspondencia entre la faja 6 del POSGAR y el EPSG 22176.

la exposición a condiciones sanitarias de riesgo (como la falta de retrete o situación de hacinamiento). Con el porcentaje de hogares con NBI y la población total construimos una variable *poblacion_nbi_estimada*, que supone uniformidad de población por hogar al interior de cada radio -independientemente de que se trate de hogares pobres o ricos- y traspola el porcentaje de hogares con NBI al porcentaje de población con NBI, infiriendo en consecuencia el valor absoluto. Como indican las notas aclaratorias publicadas junto a las Definiciones Conceptuales asociadas a los radios censales de la PBA, existen 25 radios censales de la provincia, todos correspondientes a zonas rurales, para los cuales no existen datos en la base de REDATAM, “debido a problemas en el relevamiento de los datos” (Indec, 2015).

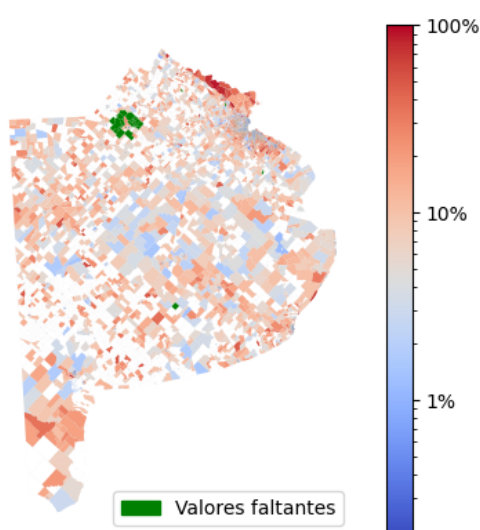


Fig. 2: Porcentaje de hogares con necesidades básicas insatisfechas por radio censal

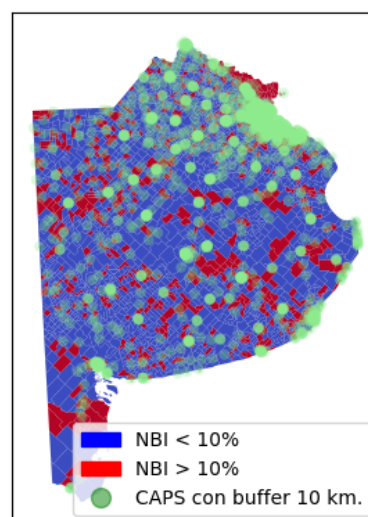


Fig. 3: Hogares con NBI y CAPS

En la figura 2 podemos ver los datos faltantes a los que ya nos referimos, además de las proporciones de hogares con necesidades básicas insatisfechas (NBI) en los radios censales de acuerdo a la escala cromática indicada en el gráfico. En la figura 3 se visualizan los valores de porcentaje de hogares con NBI segmentados en dos categorías, por encima o por debajo de un cierto umbral. Determinamos este umbral de manera heurística en un 10% y esto permite identificar, de acuerdo a nuestra metodología, zonas (radios censales) con alta incidencia de pobreza estructural (a las que por economía de lenguaje llamaremos en adelante “zonas pobres” o “zonas críticas”). Este gráfico también muestra las

ubicaciones de los CAPS con un “buffer”, esto es un área en torno del punto con un radio determinado, en este caso de 10 km (también definido arbitrariamente, aunque teniendo en cuenta una distancia que consideramos razonable para un servicio de salud que pueda considerarse “de cercanía”). Puede verse con claridad la distribución radial de los mismos y su concentración en el conurbano. A su vez puede verse su proximidad o superposición con las zonas pobres, y la falta de cobertura de muchas de estas.

Dado que la literatura, de acuerdo a lo señalado en la Introducción y el Marco teórico, da cuenta de la conveniencia de un tratamiento diferenciado de la cobertura de APS en áreas urbanas y rurales (ver especialmente Hong et al, 2023, ya citado), y que a primera vista el Conurbano parece razonablemente cubierto desde el punto de vista puramente geográfico, decidimos concentrarnos en las áreas menos densamente pobladas (que a groso modo pueden denominarse rurales y semi-rurales). Como los radios censales están asociados a partidos, seleccionamos los partidos con mayor población y excluimos los radios censales asociados del dataset. Seleccionamos los 24 partidos del Conurbano, el de la Ciudad de Mar del Plata (General Pueyrredón), que tiene 512 mil habitantes, de La Plata y Gran La Plata (más de 200 mil habitantes), de Bahía Blanca (258 mil habitantes) y de las tres siguientes ciudades en importancia de la provincia (San Nicolás, Olavarría y Tandil, de más de 100 mil habitantes cada una). También excluimos los radios censales para los que carecemos de datos.

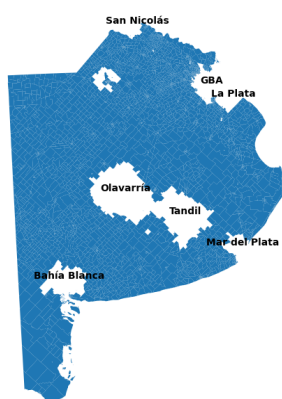


Fig. 4: Provincia sin ciudades²

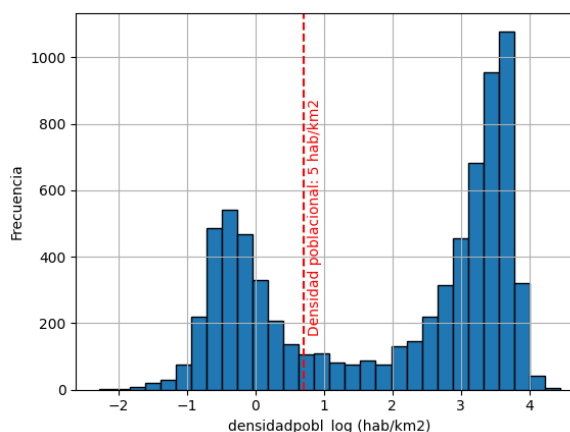


Fig. 5: Densidad de Prov. sin ciudades

² Las áreas urbanas filtradas resultaron incrementadas por la adyacencia de radios censales de baja densidad poblacional con las áreas formales de las ciudades (especialmente en Olavarría y Tandil).

Como podemos ver en la figura 5, todavía existen con este criterio zonas altamente pobladas (altas frecuencias en torno a 7500 habitantes por km²). En consecuencia establecemos un filtro por densidad de población, definiendo un tope de 5 habitantes por kilómetro cuadrado.

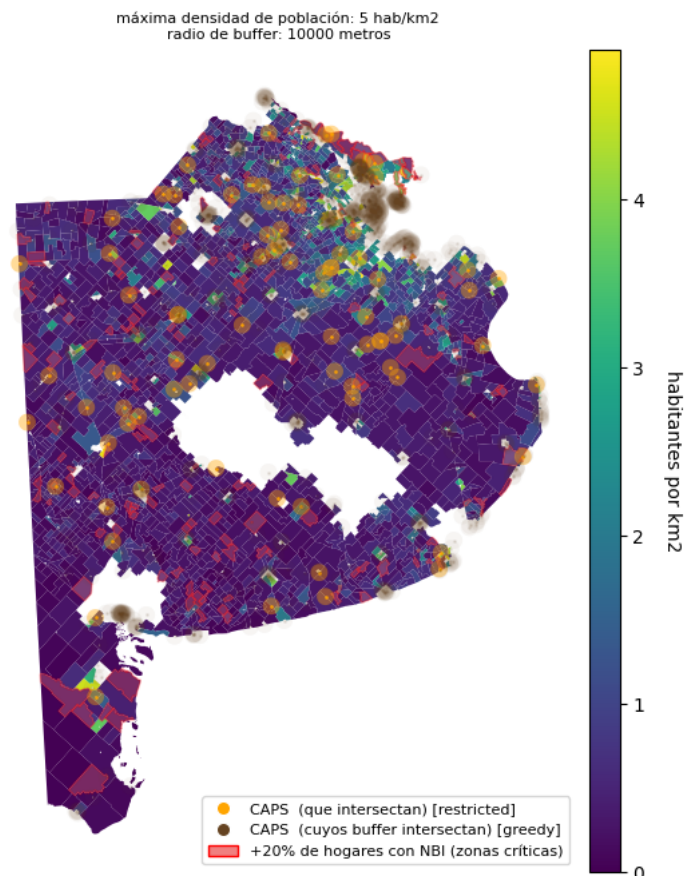


Fig. 6: Zonas de baja densidad poblacional, con CAPS y zonas críticas

En la Figura 6 observamos el universo definido para el problema planteado. Como se ve, hay dos tipos CAPS: *restricted* y *greedy*. Los CAPS del tipo *restricted* intersectan el área seleccionada (los radios censales de la provincia con densidad de hasta 5 hab./km²), es decir que el punto que los representa está dentro del conjunto de los radios. Los CAPS de tipo *greedy* son aquellos cuyos buffers intersectan al área seleccionada, es decir que se puede considerar que le brindan cobertura porque se encuentran a la distancia prudencial definida (pero este criterio más laxo no tiene en cuenta que, por tratarse de CAPS en zonas densamente

En una nueva repetición del experimento filtraríamos directamente por densidad poblacional de los radios, sin recurrir a las geoshapes de las ciudades.

pobladas, tienen más probabilidad de saturación por alta demanda). Nótese que, para los 2565 radios censales seleccionados, con el criterio restringido obtenemos 87 CAPS, y con el criterio laxo 1049 CAPS.

3.2. Métricas de evaluación

Definimos tres métricas para cuantificar la cobertura de APS en el área seleccionada (que por economía de lenguaje llamaremos en ocasiones “zonas rurales”, aunque en rigor se trata de zonas de baja densidad poblacional).

- *area_coverage* establece el porcentaje del área que está cubierto por los buffers de los CAPS, entendiendo cobertura por intersección u “*overlapping*”. Este indicador sigue un criterio puramente geográfico. Como el foco de nuestra metodología está puesto en los hogares vulnerables, esta métrica supone el establecimiento de un umbral de porcentaje de hogares con NBI (*NBI threshold*), que determina las áreas target a cubrir, sobre las cuales se determina la métrica.
- *coverage_weighted_population* pondera la métrica de cobertura de área por la cantidad de población estimada con NBI (proporción cubierta * población con NBI), por cada radio censal, obteniendo la población con NBI cubierta en total, y se realiza el cociente con la población total estimada con NBI. Esta métrica también supone distribución geográfica uniforme de los hogares con NBI al interior de los radios censales. En este caso, por la ponderación establecida, no es necesario definir un umbral de *hogares_nbi_pc*.
- *coverage_weighted_homes* es igual a *coverage_weighted_population* pero con los hogares (tiene un supuesto menos, el de uniformidad de población por hogar dentro de cada radio censal). Esta métrica y la anterior están fuertemente correlacionadas, de manera que las tomaremos indistintamente.

Para abordar el problema de optimización espacial, trabajamos con el supuesto de que se decide emplazar 10 nuevos CAPS, y se plantea la cuestión de dónde localizarlos para maximizar alguna de las métricas propuestas. Utilizaremos diferentes métodos que están detallados a continuación (generación y selección de candidatos), y para cada caso computaremos la mejora porcentual relativa obtenida

en la cobertura de cada una de los tres indicadores³. Es decir que si teníamos una cobertura del 40% y con las nuevas ubicaciones logramos una cobertura del 50%, decimos que obtuvimos una mejora del 25% en la cobertura. Como el punto inicial depende del criterio utilizado para filtrar los CAPS existentes (*greedy/laxo* vs. *restricted/exigente*) tenemos 6 métricas para cada criterio. Para la métrica de *area_coverage*, además, computamos su valor respecto de todas las áreas, y respecto de todas las áreas por encima de un umbral de incidencia de pobreza estructural, que definimos en 10%.

3.3. Optimización: algoritmos implementados

Generamos 4000 “candidatos” distribuyendo puntos de manera aleatoria en la geometría de la provincia. A su vez, filtrando estos candidatos por su pertenencia a las “zonas pobres” (dentro de los límites de densidad establecidos) tenemos 800 candidatos filtrados.

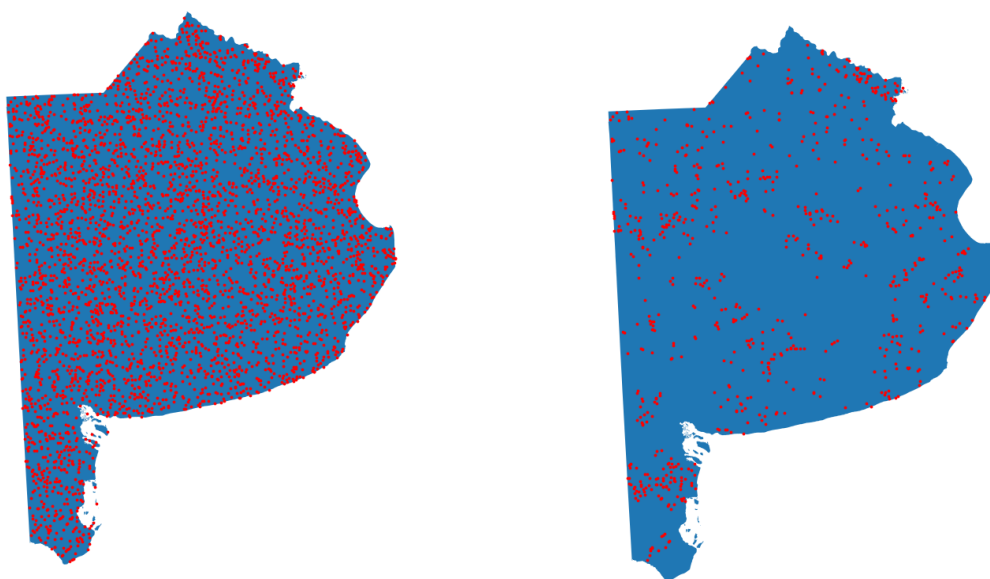


Fig. 7: Candidatos totales y candidatos filtrados

Utilizamos cuatro procedimientos de selección de candidatos:

³ La metodología implica el supuesto de que los CAPS (existentes y “nuevos”) son equivalentes en tamaño y prestaciones (ver comentario al respecto en Conclusiones y discusión).

- *random*. Selección aleatoria (control).
- *var_sorting*. Este procedimiento no utiliza los candidatos generados. Consiste en una serie de pasos:
 - filtrar las áreas que no tienen cobertura geográfica.
 - ordenarlas de acuerdo a una métrica (como área, o cantidad de población con NBI).
 - seleccionar las primeras n áreas (10, según el planteamiento del problema).
 - identificar sus centroides
 - Asignar a estos centroides los nuevos CAPS.
- *MCLP*. Maximum Coverage Location Problem. Encuadre típico para optimizaciones espaciales, resuelto mediante programación lineal entera (PLE).
- *gen_alg*. Algoritmo Genético. Familia de algoritmos populares para problemas de optimización.

Respecto del método tipo MCLP, desarrollamos una implementación utilizando la librería *pulp* de Python, especializada en PLE⁴. En cuanto al algoritmo genético, utilizamos la implementación de la librería *deap*, popular para algoritmos genéticos⁵. Usamos el algoritmo más simple provisto, “*eaSimple*”, optimizado para cada una de las métricas personalizadas que enumeramos previamente. Ambas implementaciones tienen en cuenta los CAPS pre-existentes.

Es pertinente establecer la salvedad de que la experimentación con los algoritmos fue limitada. Especialmente en el caso del algoritmo evolutivo, no llevamos adelante una exploración exhaustiva de hiperparámetros, sino que utilizamos un conjunto de valores fijos, establecidos mediante una breve exploración, de los principales parámetros del algoritmo: $pop = 100$, $cxpb=0.3$, $mutpb=0.7$, $ngen=10$. Cabe destacar que el tamaño de población seleccionado es muy limitado, pero aún en esas condiciones el tiempo de procesamiento fue muy extenso, como detallaremos a continuación, motivo por el cual optamos por no expandir esa variable. Es probable que parte del gran costo computacional, aparte

⁴ <https://pypi.org/project/PuLP/>

⁵ <https://deap.readthedocs.io/en/master/>

de características intrínsecas de los GA, esté relacionado con el elevado costo computacional de las métricas optimizadas.

En el contexto de nuestro problema, un individuo consiste en una selección de n (10) nuevas ubicaciones de CAPS, y una población, en la cantidad de individuos explorados en una generación. *mutpb* y *cxpb*, es decir los factores de mutación y cruza, definen el modo en que esas poblaciones se transforman de una generación a otra.

4. Resultados

Respecto del diagnóstico acerca de la cobertura actual, obtuvimos los siguientes resultados.

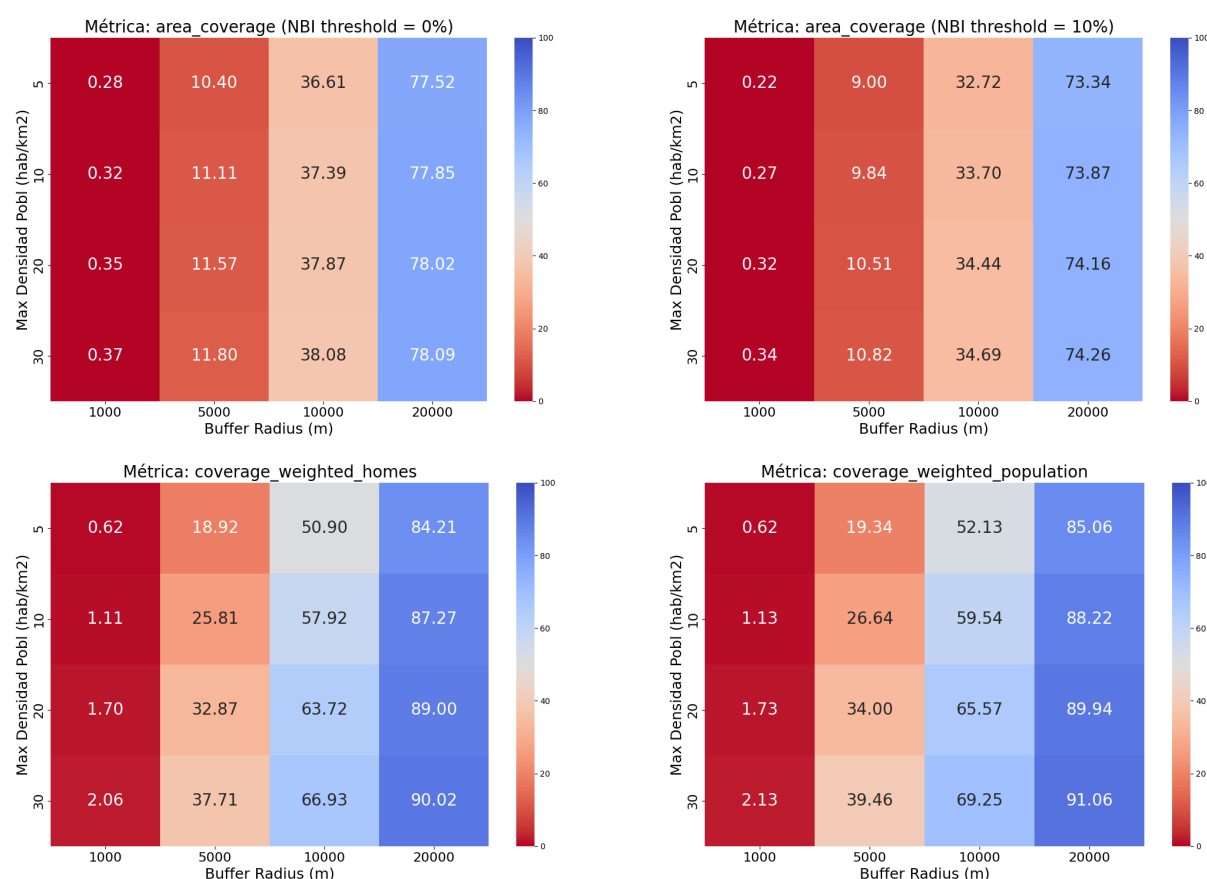


Fig. 8: Cobertura actual (zonas rurales) para cada métrica evaluada.

Para distintas combinaciones de densidad poblacional máxima y tamaño de *buffer* se observan distintos valores en cada métrica (ver página 11).

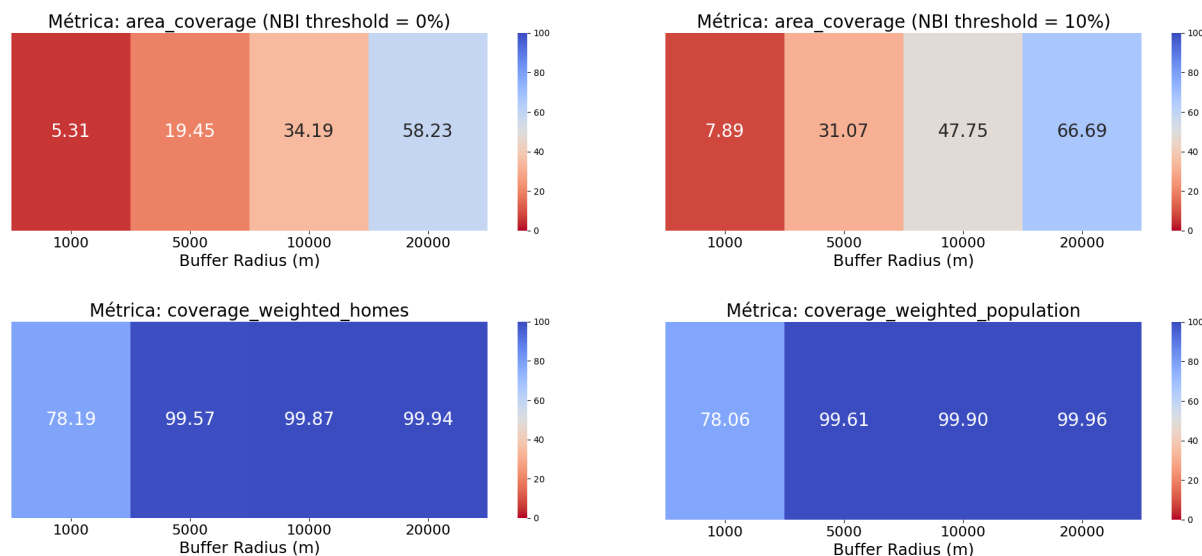


Fig. 9: Cobertura actual (zonas urbanas) para cada métrica evaluada.

En las zonas urbanas se registra una cobertura mayor en todas las métricas analizadas (ver página 11). En este caso no hay límite de densidad poblacional.

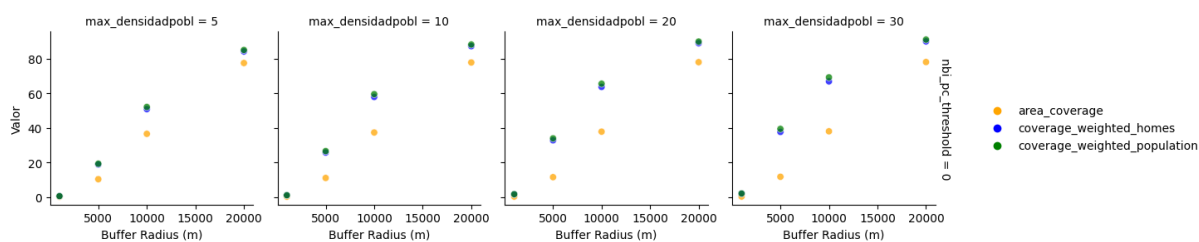


Fig. 10: Cobertura actual (zonas rurales) - dispersión.

Se observa la variación de la distancia entre las métricas de cobertura por área y las pesadas por pobreza para distintas configuraciones.

Las figuras 8, 9 y 10 ilustran algunos rasgos salientes. En primer lugar, que el diagnóstico es muy variable según la configuración elegida para la evaluación: cuál es el umbral de hogares con NBI definido para medir la cobertura geográfica, cuál es el tamaño del buffer de los CAPS, cuál es el límite de densidad considerado.

También se observa que en las zonas más urbanas la cobertura es mayor en casi cualquier métrica, en algunos casos prácticamente total. Sólo sale perdiendo en la comparación de cobertura de área con un buffer de 200 kilómetros, que en ámbitos urbanos es ciertamente excesivo (en ámbitos rurales consideramos que probablemente también).

Dentro de las áreas rurales o semi-rurales, queda claro que con un radio de cobertura de 200 kilómetros los indicadores son elevados, pero decidimos adoptar un criterio más exigente. A su vez, los radios de 1 y 5 kilómetros probablemente

resulten poco realistas para esos contextos. En este sentido optamos por un buffer de 10 kilómetros para el problema de optimización. También se advierte un valor de cobertura más elevado en las métricas pesadas por población/hogares, lo cual es consistente con las apreciaciones visuales que señalan una mayor concentración de CAPS allí donde hay más concentración poblacional, es decir, próximo a los centros urbanos.

En el mismo sentido, los cambios en el filtro de densidad poblacional no se traducen en cambios significativos para la métrica exclusivamente geográfica, pero en las otras dos se observa un incremento a medida que se amplía el criterio de selección. Consideramos que es apropiado focalizar la problemática de la cobertura allí donde se manifiesta de manera más aguda, motivo por el cual optamos por un filtro de densidad poblacional de hasta 5 habitantes por kilómetro cuadrado para modelar la selección de nuevas ubicaciones.

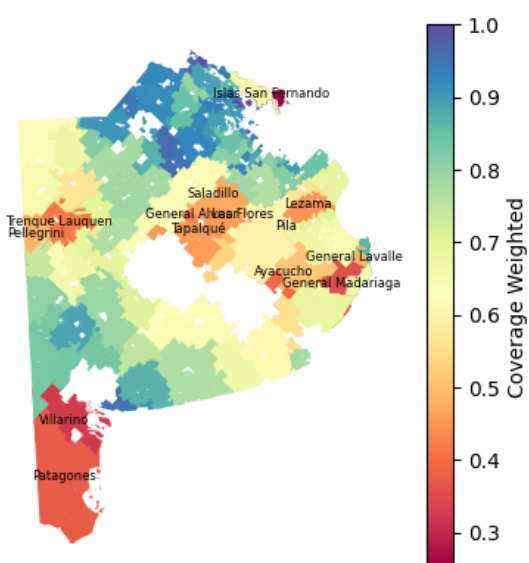


Fig. 11: Cobertura por municipio

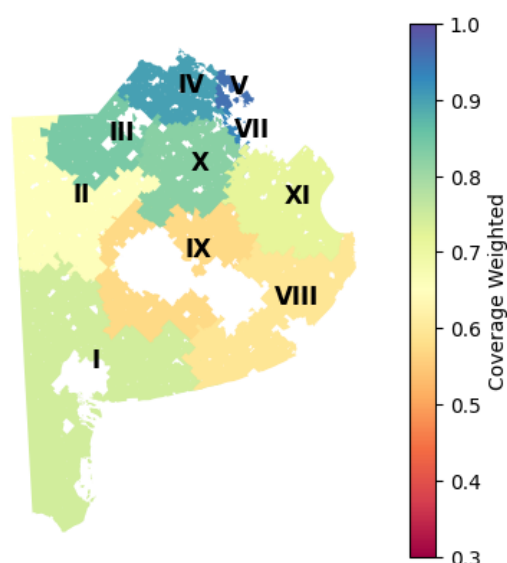


Fig. 12: Cobertura por región sanitaria

Las figuras 11 y 12 permiten apreciar la cobertura (medida según *coverage_weighted_population*) de las zonas de baja densidad poblacional agregada a nivel municipio y región sanitaria⁶. Si bien a continuación nuestro enfoque se centrará en optimizar la cobertura a nivel provincia (zonas rurales y

⁶ Las regiones sanitarias son las divisiones administrativas para la gestión del sistema de salud. Ver más en : https://www.gba.gob.ar/saludprovincia/regiones_sanitarias

semi-rurales), puede apreciarse una diferencia de cobertura significativa a nivel de las divisiones políticas y administrativas pertinentes. Esto abre otra ruta de exploración para futuros trabajos.

Veamos entonces los resultados de la optimización espacial. El algoritmo genético insume una cantidad de tiempo que está en otro orden de magnitud (más de 5 horas) para producir resultados aceptables, mientras que en los otros casos es cuestión de pocos segundos.

En la métrica de cobertura de área ningún algoritmo produjo una mejora significativa con respecto al procedimiento que llamamos *var_sorting* (ordenamiento por variable). Ya sea tomando el criterio restrictivo o el laxo (por separado) siempre produce la mejora más alta. Ahora bien, este método no resulta tan eficiente para optimizar las métricas que pesan la cobertura geográfica de acuerdo a los hogares o la población con NBI. Allí el mejor rendimiento lo produjo el algoritmo genético.

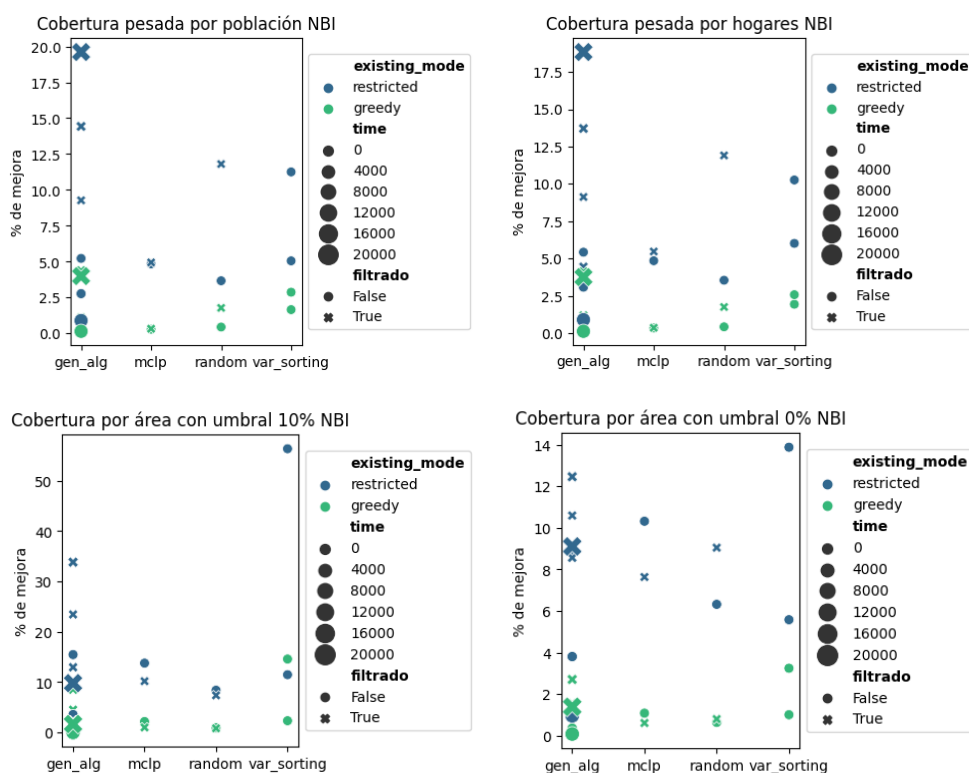


Fig 13: Resultados de la simulación, para cada métrica analizada.

existing_mode=restricted/greedy refiere a la cantidad de CAPS pre-existentes considerados para el cómputo de mejora de cobertura (ver detalle en Fig. 6 y pág. 10). El tamaño de los puntos (*time*) representa el tiempo consumido por el método de selección. *filtrado=false/true* refiere al conjunto de candidatos utilizados para la selección (ver detalle en fig. 7 y pág. 12). “Umbral x% NBI” refiere al área considerada para análisis (ver detalle en pág. 11)

La configuración del algoritmo genético en la experiencia con mejor resultado fue la que se expresa en la figura 14.

tipo de parámetro	parámetro	valor
del dataset	candidatos filtrados por pobreza (+10%)	True
del dataset	cantidad de candidatos	2613
del modelo	n_population	100
del modelo	cxpb (factor de cruza)	0.7
del modelo	mutpb (factor de mutación)	0.3
del modelo	ngen (cantidad de generaciones)	10
del modelo	indpb	0.05
del modelo	toursize	3
del modelo	métrica custom	coverage_weighted_homes

Fig. 14: Configuración del algoritmo genético

5. Conclusiones y discusión

El desarrollo de métricas específicas para el análisis de cobertura de servicios públicos, del sistema de salud y especialmente de la atención primaria tiene, como hemos visto, un cierto recorrido en la literatura. Esta tarea involucra el desafío de combinar múltiples factores, que básicamente se relacionan con la accesibilidad desde el punto de vista logístico por parte de los habitantes y la heterogeneidad de estos mismos habitantes en cuanto a sus necesidades de asistencia. La aplicación de indicadores bien sofisticados, como las distintas variantes de E2SFCA, excede los límites de este trabajo, y además resulta dificultoso en el ámbito analizado por múltiples razones, algunas de las cuales comentaremos a continuación.

En primer lugar, un déficit significativo en la disponibilidad de datos abiertos en las distintas jurisdicciones de la República Argentina, incluyendo la Provincia de Buenos Aires, refiere a la capacidad operativa de los distintos centros de salud. Este aspecto, que en otros niveles del sistema sanitario está mayormente centrado en la cantidad de camas para internación (si bien este dato tampoco suele estar disponible) en el caso de la atención primaria debe considerar otros aspectos de la infraestructura y también de personal profesional.

Otro aspecto que no hemos incorporado en nuestro análisis por motivos de alcance, a la hora de evaluar la accesibilidad, tiene que ver con la disponibilidad de rutas y medios de transportes públicos. Estos factores permitirían determinar distancias de viaje más precisas, mucho más apropiadas que distancias euclidianas que ignoran grandes barreras geográficas.

Además, la generación de candidatos para nuevas ubicaciones podría haber seguido criterios más elaborados que el método aleatorio utilizado. Por caso, circunscripciones a tierras fiscales o dependencias públicas, entre otras opciones.

Aún con estas limitaciones, consideramos un acierto la incorporación del factor socioeconómico, en particular mediante la consideración de las Necesidades Básicas Insatisfechas de los Hogares por sobre una medida de pobreza habitual como la línea de pobreza por ingresos. Las NBI toman en consideración variables fuertemente asociadas a problemáticas sanitarias, como la falta de acceso a agua corriente, cloacas, o condiciones sanitarias de hacinamiento. Por otro lado, la utilización de los radios censales permite una granularidad adecuada teniendo en cuenta la considerable segregación geográfica en la provincia de Buenos Aires al interior de divisiones administrativas como los municipios o incluso más amplias como las regiones sanitarias.

En cuanto a las medidas de evaluación de cobertura, la identificación de zonas críticas mediante la determinación de un umbral de proporción de hogares o población con NBI ofrece algunas ventajas. En primer lugar, facilita la selección de prioridades (zonas críticas a cubrir), y es compatible con métodos de procesamiento más simples. En este sentido, el método por ordenamiento (*var_sorting*) obtuvo los mejores registros para la maximización de esta variable, y además en tiempos de procesamiento muy exigüos.

Sin embargo, y máxime teniendo en cuenta que se observa una distribución muy dispersa de hogares pobres en las zonas rurales de la provincia resulta interesante el uso de un indicador que permite evaluar la cobertura sobre toda el área analizada de manera ponderada. Para este tipo de medición, el algoritmo genético presenta resultados significativamente mejores, aunque a un alto costo computacional. Por otro lado, en la implementación utilizada se buscó maximizar un indicador compuesto por ambos factores (área y pobreza), pero otra ventaja de este tipo de algoritmos es su compatibilidad con problemas multi-objetivo, que permitiría

buscar un óptimo en simultáneo para estos factores y otros (como los tiempos de viaje) considerados cada uno de manera “pura”.

Es relevante señalar que, como mencionamos antes y si bien no lo comprobamos de manera concluyente, probablemente resulte importante asegurarse de que el cómputo de la/s métrica/s optimizadas sea computacionalmente poco costoso. De lo contrario, los tiempos de procesamiento resultan muy ineficientes.

En general, para la cobertura base que la optimización debe mejorar, consideramos más apropiado el criterio restrictivo, siguiendo el razonamiento que hemos expuesto previamente: que los CAPS cuyo radio de cobertura alcanza a las zonas rurales pero están ubicados en zonas densamente pobladas probablemente experimenten una saturación por la demanda de esa población urbana.

Consideramos que los resultados de los experimentos realizados no son categóricos. Por un lado, por las limitaciones de enfoque ya mencionadas. Por otro, por la insuficiente exploración de parámetros que permitirían una comparación más acabada de los métodos de optimización en la selección de nuevas ubicaciones. Esperamos que a modo de una primera aproximación a la problemática en el contexto de la PBA resulten estimulantes para el desarrollo de nuevas investigaciones.

6. Bibliografía

Bosque Sendra, J. & García, R.C. (2000). El uso de los sistemas de información geográfica en la planificación territorial. *Anales de Geografía de la Universidad Complutense*, 20, 49–67.

Cantar, L. (2019). “Qué son los radios censales y para qué sirven. En <https://lutarocantar.com/post/qu%C3%A9-son-los-radios-censales-y-para-qu%C3%A9-sirven/#:~:text=%C2%BFQu%C3%A9%20es%20un%20Radio%20Censal,la%20cantidad%20promedio%20es%20menor>.

Chauhan DR, Unnikrishnan A, Figliozzi M, Boyles SD (2020). Robust Maximum Coverage Facility Location Problem with Drones Considering Uncertainties in Battery Availability and Consumption. *Transportation Research Record*. En

https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1591&context=cengin_fac

Freidin, B, Ballesteros, M. S., Krause, M. Wilner, A. D., & Vinitzky A. (2020). Atención primaria de la salud en tiempos de crisis: experiencias de un equipo de salud en el conurbano de Buenos Aires. Ciudad Autónoma de Buenos Aires: Betina Freidin.

Gu, W., Wang, X. & McGregor, S.E. (2010). Optimization of preventive health care facility locations. Int J Health Geogr 9, 17. en <https://doi.org/10.1186/1476-072X-9-17>

Hong, I., Wilson, B., Gross, T. et al (2023). “Challenging terrains: socio-spatial analysis of Primary Health Care Access Disparities in West Virginia”. Appl. Spatial Analysis 16, 141–161. En <https://doi.org/10.1007/s12061-022-09472-0>.

Indec (2015). “Notas aclaratorias acerca de la Provincia de Buenos Aires”. En https://github.com/datos-provincia-abierta/metadatos/blob/master/Econom%C3%ADa/notas_aclaratorias_pba.pdf

Ministerio de Salud de la Nación (2018). “Análisis de situación de salud”, República Argentina / Ministerio de Salud de la Nación, Ciudad Autónoma de Buenos Aires. Libro digital, PDF Archivo Digital: descarga ISBN 978-950-38-0271-7 1. En https://bancos.salud.gob.ar/sites/default/files/2019-12/0000001392cnt-anlisis_de_situacin_de_salud_-_repblica_argentina_-_asis_2018_compressed.pdf

Mokgalaka Hunadi (2014). “Measuring Access to Primary Health Care: Use of a GIS-Based Accessibility Analysis”, Conferencia: Planning Africa 2014 at International Convention Centre, Durban, South Africa. En https://www.researchgate.net/publication/280041255_Measuring_Access_to_Primary_Health_Care_Use_of_a_GIS-Based_Accessibility_Analysis

Paköz, Muhammed & Yüzer, M.A.. (2014). Access to healthcare: A field survey in Istanbul. A/Z ITU Journal of the Faculty of Architecture. 11. 271-290. en https://www.researchgate.net/publication/287564860_Access_to_healthcare_A_field_survey_in_Istanbul

Lim Siew Mooi, Sultan Abu Bakar Md., Sulaiman Md. Nasir, Mustapha Aida, y Leong K. Y (2017). "Crossover and Mutation Operators of Genetic Algorithms," International Journal of Machine Learning and Computing vol. 7, no. 1, pp. 9-12.

Testa, M. (1988). Atención ¿primaria o primitiva? de la salud. Segundas Jornadas de Atención Primaria de la Salud.

World Health Organization and the United Nations Children's Fund (UNICEF) (2022).

"Primary health care measurement framework and indicators: monitoring health systems through a primary health care lens". En

<https://apps.who.int/iris/rest/bitstreams/1412305/retrieve>

7. Fuentes de Datos

[1] Establecimientos de Salud de la Provincia de Buenos Aires

<https://catalogo.datos.gba.gob.ar/dataset/establecimientos-salud>

[2] Radios Censales de la Provincia de Buenos Aires

<https://ri.conicet.gov.ar/handle/11336/149711>

[3] Datos completos del Censo 2010, por radio censal

<https://redatam.indec.gob.ar/argbin/RpWebEngine.exe/PortalAction?BASE=CPV2010B>