

Índice de Vulnerabilidad a COVID-19 en Centroamérica

Vulnerability Index to COVID-19 in Central America

Paloma Merodio Gómez¹
Andrea Ramírez Santiago²
Gabriela García Seco³
Sandra Liliana Moreno Mayorga⁴
Lelio Alejandro Arias Vizcaino⁵

Fecha de recibido: 1 de abril de 2022 Fecha de aceptado: 21 de septiembre de 2022

Resumen

Desde la declaración del COVID-19 como pandemia en marzo de 2020, se han llevado a cabo una gran cantidad de análisis de datos estadísticos que han permitido dar seguimiento a la pandemia. Sin embargo, para gestionar de forma efectiva su propagación, es necesario analizar estos datos desde un enfoque geoespacial, por lo que se ha vuelto fundamental la integración de datos estadísticos y geoespaciales a distintos niveles de desagregación geográfica que permitan desarrollar e implementar nuevas metodologías, como la construcción y cálculos de índices, para obtener datos a un nivel

- ¹ Instituto Nacional de Estadística y Geografía (INEGI), México. Correo electrónico: paloma.merodio@inegi.org.mx. ORCID: https://orcid.org/0000-0002-6086-7023
- 2 Instituto Nacional de Estadística y Geografía (INEGI), México. Correo electrónico: andrea.santiago@inegi.org.mx. ORCID: https://orcid.org/0000-0002-6201-1907
- 3 Instituto Nacional de Estadística y Geografía (INEGI), México. Correo electrónico: gabriela.seco@inegi.org.mx. ORCID: https://orcid.org/0000-0002-3974-6219
- 4 Departamento Administrativo Nacional de Estadística (DANE), Colombia. Correo electrónico: slmorenom@dane.gov.co. ORCID: https://orcid.org/0000-0003-3469-7785
- 5 Departamento Administrativo Nacional de Estadística (DANE), Colombia. Correo electrónico: laariasv@dane.gov.co. ORCID: https://orcid.org/0000-0003-4706-2244

espacial común a fin de respaldar la toma de decisiones. En ese sentido, el obietivo de este artículo es presentar la metodología diseñada por el Departamento Administrativo Nacional de Estadística (DANE) e implementada en la región de Centroamérica por UNGGIM: Américas y el Instituto Panamericano de Geografía e Historia (IPGH), para la integración de datos geoespaciales y el cálculo del "Índice de Vulnerabilidad a COVID-19", así como mostrar los resultados obtenidos, los cuales están encaminados a dar una respuesta efectiva a la pandemia de COVID-19.

Palabras clave: Índice de Vulnerabilidad; Integración de Datos; Información Geoespacial; COVID-19; Centroamérica.

Abstract

Since the declaration of COVID-19 as a pandemic in March 2020, a large amount of statistical data analysis has been carried out to monitor the pandemic. However, to effectively manage its spread, it is necessary to analyze these data from a geospatial approach, which is why the integration of statistical and geospatial data at different levels of geographical disaggregation has become essential, which allows the development and implementation of new methodologies, such as index construction and calculations, to obtain data at a common spatial level to support decision making. In this way, the objective of this article is to present the methodology designed by DANE and implemented in the Central American region by UNGGIM: Americas and the PAIGH, for the integration of geospatial data, and the calculation of the "COVID-19 Vulnerability Index", as well as showing the results obtained, which are aimed at providing an effective response to the COVID-19 pandemic.

Key words: Vulnerability Index; Data Integration; Geospatial Information; COVID-19; Central America.

1. Introducción

1.1. Información Geoespacial en respuesta al COVID-19

En diciembre de 2019 se identificó en Wuhan, China el brote del coronavirus SARS-CoV-2, causante de la enfermedad llamada COVID-19. Derivado de su alta transmisibilidad, en enero de 2020, la Organización Mundial de la Salud (OMS) declaró al COVID-19 como emergencia sanitaria internacional. Posteriormente, el 11 de marzo de 2020, el brote fue declarado oficialmente por la OMS como pandemia (Whitelaw et al., 2020; Kandel et al., 2020; Casanova et al., 2021).

Las características de esta enfermedad, tales como su largo período de incubación, alta infectividad y dificultad de detección, han llevado a la demanda del conocimiento de expertos en diferentes áreas de estudio para realizar análisis que permitan diseñar e implementar medidas para controlar

la transmisión y propagación del virus (Franch-Pardo et al., 2020; Casanova et al., 2021). Si bien, el análisis de los efectos de un virus compete a la ciencia biomédica, su propagación involucra principios geográficos, razón por la cual desde el inicio de la pandemia, la información y las herramientas geoespaciales han sido fundamentales para mapear, visualizar, analizar y comprender su distribución, factores de riesgo y recursos disponibles para su tratamiento y prevención, a fin de respaldar la toma de decisiones, formular medidas para una respuesta efectiva y evaluar la efectividad de prevención y control (García de León, 2020; Murugesan et al., 2020; Boulos & Geraghty, 2020).

De esta manera, se han utilizado constantemente para monitorear la propagación y los puntos críticos de COVID-19, ayudando así a: conocer las incidencias en tiempo real; hacer una asignación y distribución efectiva de los recursos y servicios; determinar los factores de riesgo de transmisión tales como las variables socioeconómicas y ambientales; y la creación de conciencia social para prevenir y abordar la pandemia (Franch-Pardo et al., 2020; Mbunge, et al., 2021).

Adicional a esta proliferación, durante la última década, el incremento en la producción, disponibilidad y uso de información geoespacial incluyendo la información en salud, ha aumentado considerablemente, adquiriendo un papel muy importante en la toma de decisiones, por lo que diversas organizaciones han generado una gran cantidad de conjuntos de datos geoespaciales provenientes de múltiples fuentes de datos. Esta disponibilidad de datos combinada con técnicas de análisis geoespacial y estadístico tiene un potencial significativo en temas de salud pública, sin embargo, se ha vuelto difícil administrarlos, integrarlos y compartirlos, debido a que se recopila en diferentes unidades de análisis, resoluciones espaciales y formatos, complicando su interoperabilidad (Merodio et al., 2019; Chen & Wang, 2016).

La necesidad de una integración rápida de datos estadísticos y geoespaciales útiles para dar una respuesta efectiva al COVID-19, ha resaltado la falta de interoperabilidad de datos a nivel mundial, la cual es fundamental para la realización de los análisis espaciales, convirtiéndose en un desafío crítico de la capacidad analítica en la epidemiología, la salud pública y la investigación en salud (O'Reilly-Shah et al., 2020; Chen & Wang, 2016).

Por ello, para aprovechar de manera efectiva el volumen y los diferentes tipos de datos geoespaciales disponibles a nivel mundial, es necesario estandarizar los datos a fin de permitir una integración e interoperabilidad entre los diferentes sistemas de datos e información, asimismo, es necesario el desarrollo de metodologías comunes y de fácil replicación, que permitan obtener información útil para una mejor respuesta a la pandemia. En ese sentido, el objetivo de este artículo es presentar la metodología diseñada por el DANE e implementada en la región de Centroamérica por UNGGIM: Américas y el IPGH para la integración de datos geoespaciales, y el cálculo del "Índice de Vulnerabilidad a COVID-19", así como mostrar los resultados obtenidos, los cuales están encaminados a apoyar y respaldar la toma de decisiones en la región y ayudar a dar una respuesta efectiva a la pandemia de COVID-19.

El documento se divide en cinco secciones: la primera sección describe las características principales del proyecto "UN-GGIM: Américas Comité Regional de las Naciones Unidas sobre la Gestión Global de Información Geoespacial para las Américas. Integración de Información Estadística y Geográfica en América Central (Proyecto Centroamérica)" para la medición del "Índice de Vulnerabilidad a COVID-19" en la región; la sección dos presenta la metodología desarrollada para la integración de los datos geoespaciales y para el cálculo del "Índice de Vulnerabilidad a COVID-19" en la región de Centroamérica; la sección tres muestra los principales resultados obtenidos del cálculo del Índice de Vulnerabilidad a COVID-19 para cada uno de los países participantes, así como el mecanismo de comunicación utilizado en la región; la sección cuatro enlista los desafíos y lecciones aprendidas; y finalmente la sección cinco presenta las conclusiones del proyecto.

1.2. Proyecto Centroamérica

El Proyecto Centroamérica fue un proyecto de asistencia técnica enfocado en la integración de información estadística y geoespacial, aprobado el 23 de octubre de 2018, durante la 48^a Reunión del Consejo Directivo del Instituto Panamericano de Geografía e Historia (IPGH) en Bolivia y finalizó en 2021. El proyecto se desarrolló por el Comité Regional de las Naciones Unidas sobre la Gestión Global de Información Geoespacial para las Américas, en colaboración con la Comisión Económica para América Latina y el Caribe (CEPAL), financiado por el IPGH y liderado por el Buró de Censos de los Estados Unidos, la Secretaría Ejecutiva de UN-GGIM: Américas y el Instituto Nacional de Estadística y Geografía de México (INEGI) (IPGH & UNGGIM, 2022a).

El enfoque del proyecto giró en torno a cuatro objetivos base (Figura 1): 1) Coordinar la colaboración entre los países de Centroamérica participantes en el proyecto: Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica, y



Figura 1. Flujo de trabajo para el logro de los cuatro objetivos base del proyecto.

Panamá; 2) Desarrollar un plan para la integración de información estadística y geoespacial en Centroamérica que pueda ser utilizado para futuros proyectos de integración; 3) Identificar problemas y prioridades de la región y métodos para analizar y visualizar información estadística y geoespacial integrada; y 4) Llevar a cabo talleres técnicos y crear productos estadísticos y geoespaciales integrados para informar a los tomadores de decisiones (IPGH & UNGGIM, 2022a).

Con el advenimiento de la pandemia de COVID-19 y los problemas a nivel global, el equipo técnico tomó la decisión de enfocarse en apoyar a los tomadores de decisiones en Centroamérica con su respuesta al COVID-19. El COVID-19 se detectó por primera vez en Centroamérica, cuando Costa Rica reportó su primer caso conocido. Posteriormente, el 10 de marzo de 2020, se reportó en Panamá la primera muerte conocida por COVID-19 y un día después, el 11 de marzo de 2020, la OMS declaró pandemia mundial al brote de COVID-19, desde entonces, los casos y las muertes por COVID-19 se extendieron por toda Centroamérica. Dada la gravedad de la pandemia, era fundamental para cada país en la región identificar a sus poblaciones vulnerables para los esfuerzos de prevención y control (IPGH & UNGGIMb, 2022b).

Dentro de los entregables del provecto se propuso: 1) crear una plataforma geoespacial que permitiera el trabajo colaborativo de integración estadística y geoespacial en la región; 2) identificar formas innovadoras de visualización y análisis de datos (mapas digitales, dashboards y storymaps); y 3) desarrollar el índice de vulnerabilidad de COVID-19 para el respaldo de la toma de decisiones de salud pública en la región (IPGH & UNGGIM, 2022B).

1.3. Índice de Vulnerabilidad a COVID-19

Los efectos de la pandemia de COVID-19 no se distribuyen uniformemente entre la población, ya que dependen de sus características sociales, económicas y de salud, por lo que existen grupos sociales más vulnerables que otros. En este sentido, es clave contar con un marco que proporcione una comprensión detallada de la dinámica espacial de la vulnerabilidad a COVID-19 para la toma de decisiones informada, relativas a las intervenciones en el ámbito sanitario, económico y social (Macharia et al., 2020; Suárez Lastra et al., 2021).

La vulnerabilidad con un enfoque social la define Wisner et al. (2014) como el conjunto de características de un grupo o individuo que les permiten anticipar, hacer frente y recuperarse del impacto de un peligro (Ress et al., 2020). Uno de los métodos para caracterizar y cuantificar los factores que hacen que algunas comunidades o poblaciones sean más o menos vulnerables a determinados peligros es el Índice de Vulnerabilidad (Amram et al., 2020).

El Índice de Vulnerabilidad muestra las características demográficas, socioeconómicas y culturales de la población en una unidad espacial, y se utiliza para dos propósitos: a) reflejar condiciones de desventaja estructural que generan mayor susceptibilidad al daño; y b) ayudar a medir y predecir los daños probables ante la exposición a amenazas específicas (Amram et al., 2020; Suárez Lastra et al., 2021).

El concepto de usar índices de vulnerabilidad es común en tiempos de desastre, por ello, dada la situación de la pandemia de COVID-19, se comenzaron a ver como una herramienta clave para la toma de decisiones de mediano y largo plazo. Uno de los principales "Índices de Vulnerabilidad a COVID-19" conocido en la región de las Américas fue el desarrollado por el Departamento Administrativo Nacional de Estadísticas (DANE) en colaboración con el Ministerio de Salud y Protección Social y el Instituto Nacional de Salud. El objetivo de este Índice es identificar, en el mayor grado de desagregación posible (manzana) a la población que por sus características demográficas y sus condiciones de salud pueden tener más complicaciones en caso de contagio de COVID-19 (DANE). El objetivo de este Índice es identificar, en el mayor grado de desagregación posible (manzana) a la población que por sus características demográficas y sus condiciones de salud pueden tener más complicaciones en caso de contagio de COVID-19 (DANE).

Este Índice se construyó con variables agregadas a nivel manzana, de tipo demográfico obtenidas del Censo Nacional de Población y Vivienda 2018, así como variables de salud aportadas por el Registro Individual de Prestaciones de Salud que contienen información de las comorbilidades de la población, las cuales se integraron a través de las variables de identificación de las personas (DANE).

Esta metodología sirvió como base para desarrollar, con apoyo del DANE, el "Índice de Vulnerabilidad para Centroamérica: Respuesta a COVID-19" como parte del "Proyecto Centroamérica". Para ello, en 2020, el equipo técnico del proyecto realizó una serie de talleres en los que se discutió con los participantes la metodología y los cálculos utilizados para la construcción del Índice, así como la adecuación de la metodología, cálculos, alcances y limitaciones para cada país en función de la disponibilidad de datos y sus niveles de desagregación geográfica (IPGH, 2020a).

El papel de los Marcos Internacionales

Marco Estadístico y Geoespacial para las Américas (MEGA)

Con el fin de facilitar la coordinación y colaboración institucional entre las oficinas nacionales de estadística, las agencias nacionales de información geoespacial, los ministerios de salud y otros actores relevantes para apoyar la armonización e integración de datos en la región en el contexto de la pandemia de COVID-19, el equipo técnico al frente de esta investigación, eligió utilizar los estándares de estructura de datos, formato y carga de datos del Marco Estadístico y Geoespacial para las Américas (MEGA),1 el cual define una propuesta para integrar de manera consistente la información estadística y geoespacial mediante el uso de estándares y niveles de desagregación geográficos comunes, así como para el acceso y uso de esta información

Estructura de datos del proyecto.

integrada, a fin de que los datos puedan analizarse y compararse entre países (IPGH, 2021a).

El MEGA, tiene una estructura de datos estandarizada y abierta, la cual proporciona tres niveles de desagregación, datos de población (desagregados por sexo) y recuentos de viviendas en cada nivel de desagregación. Además, brinda orientación sobre: la convención de nomenclatura de archivos; la referencia espacial (Proyección SIRGAS); y los formatos de archivo (shapefile de Esri). Durante el desarrollo del proyecto, el equipo técnico identificó que la estructura de datos del MEGA necesitaba ampliarse para integrar niveles adicionales de desagregación y variables de datos necesarias para desarrollar el "Índice de Vulnerabilidad para Centroamérica: Respuesta a COVID-19" (IPGH, 2021a).

Para ello, el equipo técnico del proyecto trabajó en conjunto con el DANE y el INEGI para asegurar una estructura de datos coherente y ajustada al marco de MEGA. La estructura de datos final del proyecto estuvo integrada por más de treinta variables de datos adicionales a las utilizadas por el MEGA para el cálculo del Índice de Vulnerabilidad desglosados en 5 niveles de desagregación, es decir, mayor detalle de lo permitido por el MEGA, los cuales una vez estandarizados se cargaron en ArcGIS Hub (IPGH, 2021a).

Marco Global Estadístico y Geoespacial (GSGF)

A fin de responder al objetivo central del proyecto sobre la integración de información estadística y geoespacial para establecer información armonizada y estandarizada, el proyecto adoptó el Marco Global Estadístico y Geoespacial (GSGF), el cual permite la producción de datos estadísticos y geoespaciales armonizados y estandarizados necesarios para su integración.

Esta estandarización de datos es guiada mediante la aplicación de cinco principios: 1) Infraestructura geoespacial y geocodificación; 2) Registros de datos geocodificados; 3) Geografías comunes; 4) Interoperabilidad de datos; y 5) Estadísticas geoespaciales, así como elementos claves de apoyo. Los principios básicos del GSGF presentes a lo largo del proyecto fueron: "Geografías comunes para la difusión de estadísticas", "Interoperabilidad estadística y geoespacial" y el ser "Accesible y utilizable", además de utilizar los cuatro elementos clave (colaboración institucional, infraestructura técnica, políticas y leyes nacionales y estándares y buenas prácticas) a lo largo de la vida del proyecto (IPGH, 2021a).

Como un complemento del Proyecto, el IPGH financió los cursos de e-Learning del GSGF tanto en español como en inglés, los cuales fueron desarrollados por los líderes del proyecto en colaboración con la empresa canadiense We Love Learning, así como por el trabajo del Grupo de Expertos en Integración de Información Estadística y Geoespacial (EG-ISGI) con el desarrollo del GSGF. El curso de aprendizaje electrónico brinda información para que los países y usuarios comprendan el GSGF: su valor, aplicación, infraestructura y requisitos de implementación (IPGH & UNGGIM_a).

2. Metodología

2.1. Construcción del Índice de Vulnerabilidad

Estructura de los datos²

Para lograr la gestión e integración de la información estadística y geoespacial fue necesario la armonización de los datos entre los diferentes cuerpos de trabajo de la región. Para ello, se optó por utilizar los estándares de estructura de datos, formato y carga de datos del MEGA a fin de que los datos en múltiples formatos de archivo y convenciones de nomenclatura de los países y agencias de la región pudieran recopilarse, organizarse y estandarizarse de manera consistente, permitiendo el análisis y la comparación de información entre países.

La estructura de datos del proyecto incluyó: variables; nombre de campo/ alias; tipo/ formato/ tamaño de datos; rangos de valor aceptables; unidades de medida; códigos de estado de salud; e información de cálculo. Una vez estandarizados los datos, se cargaron en un ArcGIS Hub centralizado, lo que permitió que se utilizaran para desarrollar los índices de vulnerabilidad en varios niveles geográficos sin variaciones en los mismos.

En cuanto a las variables, se obtuvieron 30 variables de datos adicionales a las utilizadas por el MEGA que incluyeron atributos demográficos, sociales, económicos y de salud, las cuales se proporcionaron en cinco niveles de desagregación geográfica (más que lo permitido por MEGA), permitiendo el desglose de datos desde el nivel nacional hasta los cinco niveles subnacionales. Los niveles más bajos de desagregación geográfica proporcionan un examen granular de la vulnerabilidad para ayudar a guiar la respuesta a un nivel local.

Reemplazo de Ceros por Nulos en conjuntos de datos³

Durante la elaboración de los archivos por Nivel, basados en la estructura de datos del proyecto técnico, se detectó un problema relacionado con las variables que contienen sólo datos parciales. El uso de valores cero para las variables con datos faltantes, causan un problema con el cálculo del índice lo que puede ser malinterpretado y analizado incorrectamente. Existen dos ejemplos diferentes de datos faltantes para atributos:

- 1. Cuando no se tienen datos disponibles para una variable para cualquiera de las geografías en el archivo por Nivel.
- 2. Cuando se tienen datos disponibles para una variable para algunas de las geografías en el archivo por Nivel, pero le faltan datos o no tiene datos para las geografías restantes.

Estructura de datos del proyecto.

Reemplazo de Ceros por Nulos en conjuntos de datos.

Para el caso 1) se determinó que el atributo (variable) debía ser nulo para todas las geografías, y para el caso 2) se determinó que se debía poner el valor cero, para aquellas geografías que no tenían datos disponibles. Sin embargo, se detectó que el uso de variables de valor mixto impacta los cálculos del índice de vulnerabilidad. Por ejemplo, los valores cero pueden ser malinterpretados y analizados incorrectamente como valores cero, causando confusión para aquellas geografías que tienen un valor de atributo real y verdadero de cero. Una práctica de gestión de datos más apropiada y necesaria es utilizar valores nulos cuando faltan datos.

Por otro lado, el uso de variables de valor mixto con cero para los valores faltantes reduce o aumenta incorrectamente la vulnerabilidad para algunas geografías, resultando en una categorización inexacta durante el cálculo del índice. En este caso, los ceros que se utilizan actualmente para representar los datos faltantes deben reemplazarse por nulo. Solo los valores cero reales y verdaderos deben permanecer en sus archivos, con el objetivo de eliminar cualquier posible confusión sobre el valor.

El efecto es que las geografías sin valores (nulo) tendrán "vulnerabilidad promedio", lo que permite que la agrupación de vulnerabilidades sea influenciada por los otros atributos usados en el cálculo del índice. La utilización de esta solución reduce el impacto de los ceros en el cálculo y la categorización inexacta de estas geografías durante el cálculo del índice.

Variables de no-proporción⁴

La herramienta desarrollada para el cálculo de Índice de Vulnerabilidad es una versión basada en Python del modelo de índice de vulnerabilidad COVID-19 del DANE.⁵ La metodología calcula el índice de vulnerabilidad a través del agrupamiento de k-medias, que asigna el valor de los datos a un grupo o K-cluster, lo que significa que el valor de cada dato determinado se agrupa junto con otros datos que pertenecen juntos a un solo cluster. Es importante mencionar que para determinar el número óptimo de grupos, se aplicó el método del codo, el cual determinó un K=5: 1) Vulnerabilidad Alta, 2) Vulnerabilidad Media alta, 3) Vulnerabilidad Media, 4) Vulnerabilidad Media Baja, y 5) Vulnerabilidad Baja (IPGH, 2021c).

Para que el agrupamiento de K-medias sea efectivo en la identificación de grupos de vulnerabilidad, es importante que todas las variables utilizadas tengan valores entre 0.0 y 1.0. Los valores más altos de las variables utilizadas para el cálculo del índice corresponden a una mayor probabilidad de vulnerabilidad (es decir, 0.0 es un indicador de vulnerabilidad más baja y 1.0 es un indicador de vulnerabilidad más alta) (IPGH, 2021c).

La estructura de datos del proyecto sirve para que los países normalicen sus variables para tener un valor entre 0.0 y 1.0, usando fórmulas matemáticas

- Calculando el Índice de Vulnerabilidad: Utilizando variables de no proporción.
- Nota Metodológica: Índice de vulnerabilidad del DANE.

básicas para crear una nueva variable que pueda utilizarse en el Índice de Vulnerabilidad. Esta normalización evita que una variable tenga más impacto en el cálculo del Índice de Vulnerabilidad, ya que la utilización de cualquier variable que no esté en proporción puede afectar significativamente los resultados del índice (IPGH, 2021c).

Para las variables que aún no están en forma de proporción, la metodología presenta dos fórmulas que permiten normalizarlas en una nueva variable útil para el cálculo del índice (IPGH, 2021c):

- 1. Variables donde los valores más altos corresponden a una mayor vulnerabilidad: Valor actual/valor máximo para ese nivel. Con este cálculo, el valor más alto se normalizará a 1.0 (más vulnerable) y los valores más bajos están más cerca de 0.0 (menos vulnerable).
- 2. Variables donde los valores más baios corresponden a una mayor vulnerabilidad: Valor mínimo para ese nivel/valor actual. Con este cálculo. el valor más bajo ahora se normalizará a 1.0 (más vulnerable) v los valores más altos están más cercanos a 0.0 (menos vulnerable).

En este sentido, la herramienta necesitó que los usuarios importaran sus feature layer o shapefiles con datos de manera normalizada, con el obietivo de coincidir con la estructura de datos del proyecto. Posteriormente se obtuvo el valor máximo de cada una de las variables por cada una de las geografías, y se contó el número de máximos en cada uno de los cinco grupos. Por consecuencia, el grupo con mayor vulnerabilidad es el grupo con mayor número de máximos, y en caso de presentarse un empate, se resolvió de forma aleatoria.

Finalmente, cuando se ejecutó la herramienta, los resultados del cálculo del Índice de Vulnerabilidad se agregaron al feature layer de entrada generándose un mapa digital de vulnerabilidad basado en las variables y los parámetros del modelo seleccionados por el usuario.

2.2. Notebook del Índice de Vulnerabilidad

Para facilitar el uso de la herramienta, el Buró de Censos de Estados Unidos desarrolló un ArcGIS Notebook⁶ que permite a los usuarios importar los datos necesarios y configurar los parámetros del modelo a través de una interfaz simple. El Notebook consta de 11 pasos, los cuales no requieren de conocimientos de Python. Los usuarios deben seguir claramente los 11 pasos definidos para ingresar sus datos, seleccionar las variables que se utilizarán en el cálculo de la vulnerabilidad y establecer algunos parámetros básicos del modelo a través de una interfaz simple. Los 11 pasos se describen a continuación:

- Paso 1: importar todos los paquetes de Python necesarios para que Notebook funcione correctamente.
- Paso 2: agregar el feature layer con el que desea trabajar en el Notebook.

Notebook: Manual de usuario para el cálculo del índice de vulnerabilidad.

- Paso 3: Identificar la capa dentro del feature layer con la que se trabajará. El feature layer seleccionado durante este paso debe contener más de una geografía para que este Notebook funcione correctamente.
- Paso 4: Abrir el feature layer para análisis, la cual no requiere entrada de usuario
- Paso 5: Seleccionar el valor de k-Means.
- Paso 6: Seleccionar el campo de geocodificación único mediante la selección de un valor del menú desplegable, el cual debe ser uno de los códigos de nivel integrados de la estructura de datos.
- Paso 7: Seleccionar los campos que contienen las métricas que se utilizarán para calcular el índice de vulnerabilidad.
- Paso 8: Proporcionar el nombre del nuevo campo para contener el índice de vulnerabilidad. El nombre es de libre elección, sin embargo, debe ser un nombre de campo único que aún no exista en el feature layer.
- Paso 9: Realizar el análisis de *k-Megns* utilizando las entradas de usuario adquiridas previamente.
- Paso 10: Guardar el índice de vulnerabilidad en la columna definida por el usuario en la capa de características para que pueda usarse en análisis futuros fuera de Notebook.
- Paso 11: Mostrar los resultados del análisis, permitiendo al usuario ver rápidamente si los resultados son los que esperaba sin necesidad de crear un mapa web utilizando el feature layer. También ayuda a determinar si se debe usar un valor diferente de k-Means

2.3. Visualización de los datos⁷

Mapas Digitales y Dashboard

Como parte de la visualización de datos, se crearon mapas digitales de los índices calculados por cada uno de los niveles de desagregación geográfica, los cuales fueron la base para el desarrollo de los dashboards, creados para la publicación de los resultados obtenidos sobre el cálculo del Índice de Vulnerabilidad en cada uno de los países.

Los dashboards constituyen una plataforma geoespacial innovadora para visualización y análisis de información geoespacial integrada, que permitieron a lo largo del proyecto el trabajo estadístico y geoespacial colaborativo en la región. Estos dashboards además de tener los mapas digitales creados, contienen gráficas y cuadros que muestran el comportamiento de las variables y estadísticas utilizadas para el cálculo de los índices.

StoryMaps

Como parte de la difusión de los resultados, cada uno de los países realizó un Storymap con cinco apartados sobre: 1) la pandemia de COVID-19, 2)

Resultados del índice de vulnerabilidad por país Centroamericano.

el Proyecto Técnico Centroamérica, 3) el Índice de Vulnerabilidad, 4) las herramientas para la evaluación v 5) la serie de indicadores. Cada apartado contiene información que compila fuentes oficiales, el abordaje de la pandemia en el país, la situación de las fases de vacunación, metodología aplicada al proyecto, resultados de los cálculos de los indicadores, así como información sobre generalidades del país.

Informe final

Finalmente, en el marco de la documentación del Proyecto Centroamérica y sus resultados en cada uno de los países, se realizaron informes finales por país, los cuales son el documento de referencia sobre las actividades más significativas ejecutadas bajo el proyecto, además de que presenta el proceso para la elaboración de indicadores asociados a la pandemia del COVID-19.

3. Resultados v discusión

A nivel regional se elaboraron cuatro dashboards8 (Figura 2). Dos sobre las condiciones de salud de Centroamérica para el Nivel de desagregación 1 (nacional) y para el Nivel de desagregación 2 (primer nivel subnacional), para las variables de datos comunes de salud: hipertensión, obesidad, diabetes, cardiopatías, condiciones del sistema respiratorio, cáncer y condiciones del sistema inmunológico. Y dos sobre otros factores de riesgo, para las variables de datos comunes de población: población de 60 años y más, pobreza, acceso a agua pública, acceso a saneamiento público, densidad de población.



Figura 2. Dashboards de resultados, correspondientes a los factores de riesgo en Centroamérica para cada uno de los cuatro niveles de desagregación.9

Dashboards de los datos de factores de riesgo para Centroamérica.

Panorama general de los datos de factores de riesgo de Centroamérica.

3.1. Índice de Vulnerabilidad por país

Costa Rica¹⁰

Costa Rica determinó su Índice de Vulnerabilidad en tres niveles de desagregación geográfica subnacional, además del Nivel 1. País: Nivel 2. Provincias; Nivel 3, Cantones y Nivel 4, Distritos. Para cada uno de estos tres niveles de desagregación geográfica, se utilizaron las mismas 11 variables: cinco variables correspondientes a datos de salud (factores de riesgo) actualizadas al año 2019 y seis variables correspondientes a datos de población actualizadas al año 2011 (dos de acceso a servicios públicos, uno de condición de las viviendas, nivel de pobreza, población total y densidad demográfica). Además, se determinaron cinco niveles de vulnerabilidad (alta, media alta, media, media baja y baja) para cada uno de los niveles de desagregación.

- Nivel 2: Provincias. El comportamiento del Índice de Vulnerabilidad difiere de la situación y distribución de casos de COVID-19, debido a factores que escapan de las variables utilizadas para el cálculo, como son el acceso a la educación, el nivel cultural, el compromiso con la salud, el acceso a información confiable y a los problemas económicos que enfrenta el país, los cuales obligaron a sectores de la población a salir a trabajar y a exponerse al riesgo del contagio. En este sentido, las provincias costeras son las que presentan menos casos y menos muertes en números absolutos, pero una vulnerabilidad más alta. En el caso de las provincias centrales con mayor densidad de población, presentan una leve relación directa entre la vulnerabilidad presentada de media a media alta y contagios de COVID-19 con los números más altos de muertes.
- Nivel 3: Cantones. A nivel general la vulnerabilidad predominante es la media alta. Sin embargo, este nivel de desagregación permite distinguir casos más puntuales de vulnerabilidad alta hacia el centro del país. Por otro lado, hacia las costas sobresale una mezcla de unidades territoriales con vulnerabilidad entre media alta v media.
- Nivel 4: Distritos. En este nivel sobresalen distritos con vulnerabilidad alta distribuidos a lo largo de todo el país, asociados a áreas de alta densidad demográfica en el centro del país y a condiciones de pobreza y exclusión económica y social en la periferia. La vulnerabilidad baja se presenta en zonas aisladas y de muy baja densidad de población, mientras que la vulnerabilidad media alta y alta domina el resto del país, donde se encuentra la mayoría de la población económicamente activa, pero también donde se presenta el mayor rezago educativo, falta de servicios de calidad, desinformación respecto a una alimentación saludable, descuido de la salud y bajo nivel de ingresos. Por otro lado, existe una alta vulnerabilidad en la frontera con Panamá, ocasionada por el alto movimiento fronterizo de mercancías y de personas, sin embargo, la

Resultados de Costa Rica.

incidencia de COVID-19 es baja, ya que las campañas de salud fueron más fuertes ante el peligro inminente.

En función de la visualización de datos, en Costa Rica desarrollaron cinco dashboards: un dashboard total, el cual contiene los datos conjuntos de los niveles de desagregación y cuatro correspondientes a cada uno de los niveles de desagregación, incluyendo el Nivel 1 País (Figura 3), los cuales contienen los mapas del Índice de Vulnerabilidad, así como gráficas y cuadros que muestran el comportamiento de las variables y estadísticas utilizadas para el cálculo del índice.

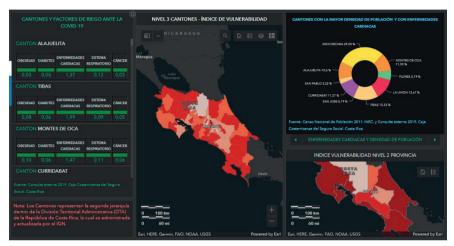


Figura 3. Dashboard de Costa Rica correspondiente al Nivel de desagregación 3, Cantones.11

El Salvador¹²

El Salvador determinó su Índice de Vulnerabilidad únicamente en el Nivel de desagregación 2 Departamental, lo que dificultó las comparaciones con la información sobre los casos COVID-19 del gobierno central. Este índice se calculó con información proveniente de dos fuentes de datos: por un lado, la Encuesta de Hogares de Propósitos Múltiples (EHPM) actualizada a 2018 con cobertura hasta Nivel Departamental, para el caso de las variables de población, y por el otro, la información del Ministerio de Salud (MINSAL) para el caso de las variables de salud, en este caso, se proporcionaron datos sobre la primera consulta referida a cada uno de los padecimientos, ya que no se cuenta con un registro clínico único para cada paciente. Además, se determinaron cinco niveles de vulnerabilidad (muy alta, alta, media, baja y muy baja).

En este caso se calcularon dos Índices de Vulnerabilidad. El primero, se calculó con cinco variables de población (pobreza, población mayor de 60

Dashboard Costa Rica, Nivel 1, País.

Resultados de El Salvador.

años, hogares sin acceso a saneamiento y agua, densidad de población y hacinamiento en dormitorios) y dos variables sobre condiciones de salud (diabetes e hipertensión). El segundo se calculó con siete variables de población (pobreza, edad de 30 a 40 años, hogares sin acceso a saneamiento y agua, densidad de población, hogares con características intergeneracionales, hacinamiento en dormitorios, ingresos familiares) y con variables sobre personas con dolencias en el sistema respiratorio.

- Nivel 2: Departamental (Índice de Vulnerabilidad 1). Los resultados del primer índice permitieron ver que existió una relación entre la vulnerabilidad de los departamentos y el número de contagios publicados por el gobierno. También se pudo observar que los departamentos con los valores más bajos de vulnerabilidad están relacionados con un mayor acceso a los servicios públicos y mayores ingresos familiares, mientras que los valores más altos de vulnerabilidad coinciden con los departamentos que presentan mayor pobreza y alta densidad poblacional.
- Nivel 2: Departamental (Índice de Vulnerabilidad 2). Por otro lado, los resultados arrojaron que el segundo Índice de Vulnerabilidad meioró al incluir más variables de salud poblacional y la variable de hogares intergeneracionales. En este índice se pudo observar que los departamentos donde se concentra la mayor pobreza, y donde hay poco acceso a agua potable y saneamiento, tienden a tener los valores más altos en el Índice de Vulnerabilidad. Por el contrario, donde hay más desarrollo urbano tienden a valores más bajos en el Índice de Vulnerabilidad, a excepción de San Salvador, donde se concentra una mayor densidad poblacional, elevando un poco el valor del índice.



Figura 4. Dashboard de El Salvador correspondiente al Nivel de desagregación 2, Departamental.13

Dashboard El Salvador, Nivel 2, Departamental.

Finalmente, como parte de la visualización de datos, el Centro Nacional de Registros (CNR) de El Salvador desarrolló un dashboard en la plataforma ArcGIS Online (Figura 4), el cual contiene los mapas digitales de los dos Índice de Vulnerabilidad para el Nivel 2 Departamental, además de gráficas y tablas con información de las variables utilizadas para el cálculo de los índices.

Guatemala¹⁴

Costa Rica determinó su Índice de Vulnerabilidad en dos niveles de desagregación geográfica subnacional, además de en el Nivel 1, Nacional: Nivel 2, Departamental y Nivel 3, Municipal.

Este índice se calculó con información proveniente de dos fuentes de datos: por un lado, el Censo de Población del Instituto Nacional de Estadística (INE) actualizado a 2018, para el caso de las variables de población, y por el otro, la información del Ministerio de Salud Pública y Asistencia Social (MSPAS) actualizada a 2020, para el caso de las variables. Además, se determinaron cuatro niveles de vulnerabilidad (alta, media alta, media baja y baja).

Para el cálculo del Índice de Vulnerabilidad para el Nivel 2 y 3, se utilizaron las siguientes variables de población: población total, densidad de población, población por rango de edades a partir de los 10 años, total de viviendas, población de 60 años y más que viven en hogares unipersonales, población de 60 años y más que viven en hogares familiares, hogares con hacinamiento en cuartos y dormitorios, hogares con riesgo intergeneracional, acceso a agua pública, acceso a alcantarillado público, pobreza, discapacidades e ingresos; y las siguientes variables sobre condiciones de salud: hipertensión, obesidad, diabetes, enfermedades cardíacas, sistema respiratorio, cáncer y sistema inmune.

En el caso particular de Guatemala, se realizaron cuatro diferentes Índices de Vulnerabilidad para cada uno de los niveles de desagregación (Nivel 2 y Nivel 3), los cuales combinaron distintas variables de población y de salud.

En general los resultados del Índice de Vulnerabilidad tanto para el Nivel 2 como 3 se encuentran en los niveles media baja y alta enfocándose en la región sur oriental del país, la cual se encuentra próxima a zonas fronterizas y son lugares con extrema pobreza. Además, se pudo observar que la proporción más vulnerable al COVID-19, en estado grave, se manifiesta en población con enfermedades de Hipertensión, Diabetes y del Sistema Respiratorio. Por otro lado, se puede considerar que a medida que incrementa la vulnerabilidad de los municipios, la letalidad por COVID-19 aumenta. Particularmente los casos de personas de 60 años y más, presentan mayor tasa de mortalidad por COVID-19, evidenciando que este grupo de edad presenta mayor riesgo y vulnerabilidad.

Es importante mencionar que no hubo ningún desafío o problema con la disponibilidad y la coherencia de los datos de Guatemala solicitados en este proyecto, ya que se obtuvo toda la información del censo del año 2018, así como los datos de enfermedades que generan mayores riesgos de contagio por COVID-19.

Finalmente, como parte de la visualización de datos, se desarrollaron en la plataforma ArcGIS Online tres dashboard, uno a Nivel 1 Nacional, y dos para los Niveles 2 y 3 (Figura 5), los cuales contienen los mapas digitales del Índice de Vulnerabilidad, además de gráficas y tablas con información de las variables utilizadas para el cálculo de los Índices.

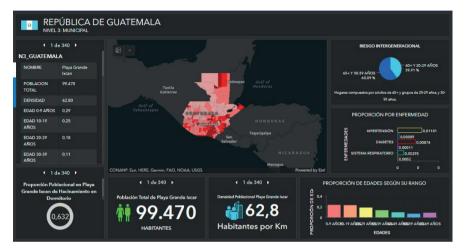


Figura 5. Dashboard Guatemala, Nivel 1, Nacional. 15

Honduras¹⁶

Honduras determinó su Índice de Vulnerabilidad en dos niveles de desagregación geográfica subnacional: Nivel 2, Departamentos y Nivel 3, Municipios. Este índice se calculó con información proveniente de tres fuentes de datos: el Censo Nacional de Población y Vivienda de 2013 y Proyecciones a 2018 del Instituto Nacional de Estadística de Honduras (INE), para el caso de las variables de población y vivienda; información geográfica a escala 1:300,000, del Instituto de la Propiedad (IP), para el caso de la geografía del territorio; y los datos de enfermedades preexistentes atendidas a nivel nacional para el año 2018 de la Secretaría de Salud para el caso de las variables de salud. Además, se determinaron cuatro niveles de vulnerabilidad (alta, media alta, media, media baja y baja).

Para el cálculo del Índice de Vulnerabilidad para el Nivel 2 y 3, se utilizaron siete variables de población y vivienda (población adulta mayor de 60 años y más, para el año 2013 y su proyección a 2018, hacinamiento para el

Dashboard Guatemala, Nivel 1, Nacional.

Resultados de Honduras

año 2013, densidad poblacional para el año 2013 y su proyección a 2018, vivienda sin servicio de agua para el año 2013, vivienda con problemas de saneamiento para el año 2013, población con alguna discapacidad y pobreza medida por NBI, y cinco variables de comorbilidades (hipertensión, diabetes, cardiopatía isquémica, pulmonares crónicas y cáncer), obtenidas del número de atenciones registradas para el año de 2018 por salud pública, las cuales no incluyen las atenciones en el sector privado, además que solo representan el número de atenciones o egresos durante el año ya que no se dispone de datos a nivel individual de la población atendida.

Derivado de esta información, se realizaron combinaciones de variables para el cálculo de cuatro distintos índices tanto para el Nivel 2 como para el Nivel 3, con el propósito de obtener distintos parámetros de análisis. Los productos finales que se obtuvieron fueron tres índices para los niveles administrativos departamental y municipal: Índice de Vulnerabilidad considerando características de la población y condiciones de la vivienda; Índice de Vulnerabilidad, considerando solo características de la población; e Índice de Vulnerabilidad, considerando solo las condiciones de la vivienda.

Los resultados generales arrojaron que el 44% de los departamentos del territorio están dentro del mayor grado de vulnerabilidad. Al subir el detalle a nivel municipal se identificó que el 2% de los municipios se encuentran en un alto nivel de vulnerabilidad. De la relación de los datos de COVID-19 y el Índice de Vulnerabilidad, se determinó que si existe una asociación entre el índice de vulnerabilidad calculado y los registros oficiales de COVID-19.

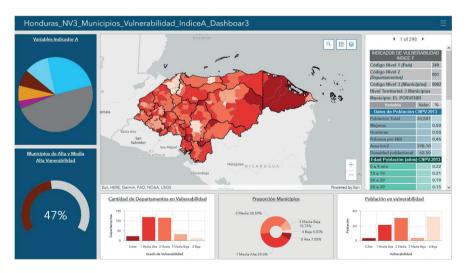


Figura 6. Ejemplo del Dashboard de Honduras correspondiente al Nivel de desagregación 3, Municipios, Índice A.17

Las cifras de mortalidad del COVID-19 están vinculadas a la población mayor de 60 años y más del país, por lo que es recomendable realizar una segunda versión de la metodología que se centre en la población entre 18 y 59 años, la cual incluye a la población económicamente activa y que en la actualidad representan un 83% de los infectados y un 61% de los decesos en el país.

Finalmente, con base a la visualización de los datos, se desarrollaron 10 dashboard: dos nacionales con datos generales y análisis de cobertura de la red de salud, cuatro dashboard con los índices de vulnerabilidad para el Nivel 2, Departamental, y cuatro dashboard con los índices de vulnerabilidad para el Nivel 3, Municipal (Figura 6).

Panamá¹⁸

Panamá determinó su Índice de Vulnerabilidad en tres niveles de desagregación geográfica subnacional, además del Nivel 1. País: Nivel 2 Provincias, Nivel 3 Distritos y Nivel 4 Corregimientos. Este índice se calculó con información proveniente de tres fuentes de datos: Censo de Población y Vivienda de 2010 del Instituto Nacional de Estadísticas y Censos (INEC), para el caso de las variables de población y vivienda; informes de pobreza multidimensional 2017 y Atlas Social 2010 del Ministerio de Economía y Finanzas para la información de pobreza y discapacidad; y boletines e informes anuales 2017 del Ministerio de Salud para el caso de las variables de salud. Además, se determinaron cuatro niveles de vulnerabilidad (alta, media alta, media baja y baja).

Para el cálculo del Índice de Vulnerabilidad para el Nivel 2, 3 y 4, se utilizaron las siguientes variables de población: población total, densidad de población, población por rango de edades a partir de los 0 años, población de 60 años y más que viven en hogares unipersonales, población de 60 años y más que viven en hogares familiares, discapacidades, pobreza, ingresos, total de viviendas, hogares con hacinamiento en cuartos y dormitorios, hogares con riesgo intergeneracional, acceso a agua pública, acceso a alcantarillado público; y las siguientes variables sobre condiciones de salud: hipertensión, obesidad, diabetes, enfermedades cardíacas, sistema respiratorio, cáncer y sistema inmune.

Las variables se utilizaron en función de la disponibilidad en las estadísticas del Instituto Nacional de Estadísticas y Censos, razón por la cual existe una mayor disponibilidad a nivel país y provincia, que a los en niveles más detallados de distrito y corregimiento. Para el caso del Nivel 1, País, se cuentan con todas las variables antes mencionadas, para el Nivel 2, Provincias, no se cuenta con información sobre hogares con hacinamiento en cuartos y dormitorios y hogares con riesgo intergeneracional, y para los Niveles 3 y 4, no se cuenta con información disponible de salud.

En respuesta al resultado de la combinación de variables, se crearon seis mapas de vulnerabilidad a Nivel 2, Provincias; tres a Nivel 3, Distritos; y tres a Nivel 4, Corregimientos. Adicionalmente, se crearon mapas de servicios de interés como aeropuertos, estaciones de policía y puertos.

Derivado de los análisis de comparación con datos reales de pandemia en distintas semanas a lo largo de ésta, se puede concluir que existe una correspondencia con los índices calculados principalmente en los indicadores a Nivel 2, Provincias, en los que se involucraron las variables de grupos de edades mayores a 40 años, pobreza, ingreso y servicios de alcantarillados y acueductos. Además, existe mucha más vulnerabilidad cuando se involucran las variables de grupo de edades mayores de 40 años. En los informes oficiales de la Organización Panamericana de la Salud (OPS), no se detalla información a nivel distrital, por lo cual no se pudo hacer una comparación a Nivel 3, Distritos. Por otro lado, para el Nivel 4, Corregimientos, la situación tiende a ser de mayor vulnerabilidad cuando se le suman las enfermedades crónicas: cáncer, diabetes y enfermedades asociadas al corazón.

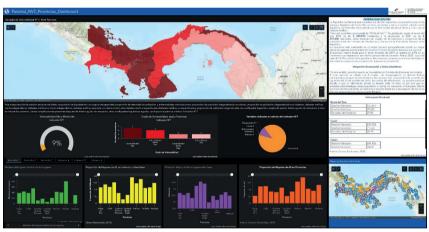


Figura 7. Dashboard de Panamá correspondiente al Nivel de desagregación, 2 Provincias.19

Finalmente, con el objetivo de que el usuario pudiera visualizar de forma rápida los datos más importantes del proyecto correspondientes a Panamá, se desarrollaron cuatro dashboard, uno a Nivel 1, País, con información general sobre la población, además de estaciones de policía, aeropuertos y puertos; e igualmente uno por cada uno de los niveles de desagregación 2 (Figura 7), 3 y 4 con información concerniente a los indicadores, cada uno con sus respectivas gráficas de vulnerabilidad y variables utilizadas para el cálculo de cada indicador, además de información general de los servicios de interés como centros hospitalarios, estaciones de bomberos, policía, puertos y aeropuertos del país.

Dashboard Panamá, Nivel 2, Provincias.

Nicaragua²⁰

Nicaragua determinó su Índice de Vulnerabilidad en dos niveles de desagregación geográfica subnacional: Nivel 2 Departamentos y Nivel 3 Municipios. Además, se determinaron cuatro niveles de vulnerabilidad (alta, media alta, media baja y baja).

- Nivel 2: Departamentos. Derivado de una combinación de variables, se calcularon tres índices diferentes:
 - El Índice de ancianos (60 años y más) utilizó tres variables de población (2 de rangos de edades de la población y 1 de acceso a agua pública) y 6 variables de salud (hipertensión, diabetes, cardiopatía, sistema respiratorio, cáncer y sistema inmunitario).
 - El Índice de edad laboral (20 años a 49 años) utilizó cuatro variables de población (3 de rangos de edades de la población y 1 de acceso a agua pública) y las mismas 6 variables de salud.
 - El Índice de jóvenes (10 a 19 años) utilizó tres variables de población (2 de rangos de edades de la población y 1 de acceso a agua pública) y las mismas 6 variables de salud.
- Nivel 3: Municipios. Derivado de una combinación de variables, se calcularon 2 índices diferentes:
 - El Índice de ancianos (60 años y más) utilizó tres variables de población (2 de rangos de edades de la población y 1 de acceso a agua pública) y 4 variables de salud (hipertensión, diabetes, cardiopatía y sistema respiratorio).
 - El Índice de edad laboral (20 años a 49 años) utilizó tres variables de población (rangos de edades de la población) y 4 variables de salud (hipertensión, diabetes, cardiopatía y sistema respiratorio).

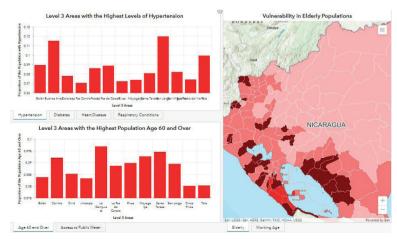


Figura 8. Dashboard de Nicaragua correspondiente al Nivel de desagregación 3 Municipios.

Resultados de Nicaragua.

Además, como parte de la visualización de datos, se desarrollaron dos dashboard, uno para el nivel de desagregación 2 y uno para el nivel de desagregación 3 (Figura 8), lo cuales contienen los mapas digitales de los Índice calculados para cada Nivel de desagregación, además de gráficas v tablas con información de las variables utilizadas para el cálculo de los índices. Es importante mencionar que los representantes de Nicaragua participaron únicamente en la primera fase del proyecto, por lo que no se cuenta con los resultados y los productos finales de visualización.

3.2. Desafíos metodológicos

El "Proyecto Centroamérica" permitió la coordinación y colaboración institucional entre las oficinas nacionales de estadística, las agencias nacionales de información geoespacial, los ministerios de salud y otros actores relevantes para apoyar la integración de datos en la región. Sin embargo, debido a la pandemia, cada país experimentó desafíos en el acceso, recopilación y agregación de los datos de las diferentes agencias gubernamentales, necesarios para el cálculo del "Índice de Vulnerabilidad para Centroamérica: Respuesta a COVID-19", debido a que muchas oficinas estaban cerradas y operando con personal limitado. Por lo que se requirió de tiempo adicional para la revisión y limpieza de datos antes de que los países pudieran seleccionar las variables necesarias para ejecutar los cálculos del índice (IPGH a. 2021).

A continuación, se mencionan los principales desafíos de datos experimentados durante el proyecto (IPGH a, 2021):

Naturaleza temporal de los datos

Los datos que conforman el conjunto de datos regionales presentaron variaciones entre países en la actualización de una misma variable debido a que cada país opera de forma independiente. Además, dentro de los países los conjuntos de datos también tuvieron variaciones en los años de actualización. ya sea porque cada una de las variables utilizadas estaban actualizadas en diferentes años, o bien porque una misma variable presentaba diferencias en su actualización según el nivel de desagregación geográfico, creando errores e inconsistencias en el cálculo y en la comparación de variables entre países.

Por otro lado, la estructura de datos del proyecto requería que algunas de las variables de datos estuvieran en formato de proporción (con valores entre 0.0 y 1.0), sin embargo, la mayoría de los países presentaba datos sin procesar y tuvieron que calcular el valor de la proporción utilizando dos puntos de datos diferentes, los cuales si eran de años diferentes creaba errores e inconsistencias en los datos de proporción.

Disponibilidad de los datos

La estructura de datos que se pidió que proporcionaran los países participantes estuvo conformada por aproximadamente 30 variables, sin embargo, dentro de cada país la disponibilidad de estas variables de datos varió mucho dependiendo el nivel geográfico de desagregación, siendo los niveles de desagregación más detallados (Nivel 5) los que presentaron mayores variaciones, debido a que algunos países no contaban con datos tan desagregados. Esto ocasionó que algunos países pudieran hacer un análisis más completo de la vulnerabilidad.

Por otro lado, la disponibilidad de variables de datos a nivel regional fue mucho más limitada, debido al bajo número de variables de datos comunes disponibles para el análisis entre los países participantes, principalmente en los niveles de desagregación más detallados. Esto ocasionó que fuera más difícil analizar la vulnerabilidad a nivel regional, a diferencia del análisis de la vulnerabilidad a nivel de país.

Completitud de los datos

La disponibilidad de datos completos no fue la misma para todos los niveles de desagregación geográfica. Generalmente, los niveles de desagregación geográfica menos detallados (Nivel 1 y 2) tenían datos disponibles más completos, y las variables disponibles en los niveles de desagregación más detallados (Nivel 3, 4 y 5) eran más limitadas.

Por otro lado, los países no siempre contaban con datos para cada área individual dentro de un mismo nivel de desagregación geográfica, es decir, podían tener datos disponibles para algunas pero no todas las áreas dentro de un nivel de desagregación geográfica específico, lo que causó dificultades al ejecutar el Índice de Vulnerabilidad. Esto ocasionó que se hicieran ajustes al Notebook del Índice de Vulnerabilidad para permitir la falta de datos en algunas áreas dentro de un nivel geográfico específico.

Además, existieron diferencias en la forma en que cada país informó los datos de las variables sobre la condición de salud, ya que algunos países solo informaron datos de los establecimientos médicos públicos, debido a la falta de disponibilidad de datos de los establecimientos médicos privados, lo que ocasionó dificultades en la comparación con precisión de los datos entre países.

Estandarización y normalización de los datos

La obtención de los datos de las variables de los países participantes en la forma proporcional requerida, con valores entre 0.0 y 1.0, fue complicada debido a que los participantes trabajaron con varias agencias gubernamentales para recopilar los datos, las cuales con frecuencia no contaban con los datos en el formato requerido, por lo que se necesitó realizar cálculos para obtener los datos en la forma de proporción solicitada, ocasionando retrasos y problemas en la integración de datos.

Sin embargo, estos cálculos no siempre resultaron exitosos para la estandarización de los datos en forma proporcional, debido a la falta de variables necesarias para el cálculo y a la diferencia de actualización de las variables para un mismo cálculo, ocasionando que se obtuvieron valores inexactos. En este caso, se permitió el uso de variables no proporcionales, las cuales debían normalizarse para que tuviera un valor entre 0.0 y 1.0, lo que requirió de más tiempo.

Errores e inconsistencias en los datos

Existió un gran número de errores e inconsistencias en los datos. El error más común fue usar valores 0 en lugar de valores nulos, lo que ocasionó que con frecuencia se requería la verificación para determinar si el valor era un valor 0 real y verdadero, o si los datos no estaban disponibles para esa área geográfica específica (en cuyo caso debería ser un valor Nulo). También hubo varios errores o inconsistencias debido a errores de cálculo o datos incompletos, lo que requirió de consultas con el país participante para revisar las posibles soluciones y hacer los ajustes apropiados.

Con el objetivo de ayudar con los desafíos de datos encontrados durante el proyecto, el equipo técnico implementó las siguientes acciones:

Revisión de la calidad de los datos

Se implementó un proceso de revisión de la calidad de los datos para identificar los problemas y brindar retroalimentación a los países, previo a la ejecución del índice. Este proceso incluyó comprobaciones para verificar que las variables proporcionadas tuvieran el formato de datos correcto y que se cumplieran los requisitos de tamaño de variable. Además, se verificó que los datos estuvieran dentro de los rangos permitidos.

Adicionalmente, se verificaron las conversiones de valores enteros a proporciones y la normalización de valores no proporcionales y se identificaron los errores para su revisión por parte del país participante. Finalmente, se verificó la consistencia, para asegurar que los valores de una variable no contradijeran los valores de otra variable.

Reuniones de consulta

Se llevaron a cabo reuniones de consulta para revisar los problemas identificados durante el proceso de revisión de calidad y discutir posibles soluciones. Este proceso consistió en una revisión inicial de los problemas asociados a los datos, y una o más reuniones de seguimiento, una vez que el país participante pudo recopilar información adicional de varias agencias con respecto a las preguntas de datos y los problemas identificados.

Revisiones de la estructura de datos

Aunque las variables de datos se solicitaron en formato de proporción, en ciertos casos se agregaron variables de datos sin procesar a la estructura de datos, para ayudar a solucionar problemas y verificar la precisión de los cálculos realizados por los países participantes. Además, se agregaron rangos de valores aceptables y se aclararon las unidades de medida en la estructura de datos para evitar errores y garantizar la coherencia entre países.

Adicionalmente, se agregaron aclaraciones a la sección de descripción de la estructura de datos para proporcionar detalles adicionales sobre la manera de calcular la variable. Finalmente, se agregaron los códigos de clasificación médica de la OMS, para enfermedades v condiciones de salud, a fin de garantizar una clasificación uniforme en todos los países.

Documentación técnica

El equipo del proyecto publicó documentación técnica sobre dos elementos que eran fuentes comunes de errores. El primero fue el uso de valores 0 en lugar de valores Nulos, y el segundo sobre el momento de utilizar cada valor y sobre la manera de convertir valores 0 a Nulo en sus archivos de ArcGIS cuando todas las áreas dentro de un nivel geográfico no tuvieran datos disponibles, y en circunstancias donde los datos estuvieran disponibles para algunas áreas dentro de un nivel geográfico, pero no para otras áreas dentro del mismo nivel geográfico. Adicionalmente se proporcionó documentación sobre la forma de normalizar los valores no proporcionales a una relación, con valores entre 0.0 y 1.0.

Ajustes al Notebook del Índice de Vulnerabilidad

Debido a que hubo problemas con la integridad de los datos para todas las áreas dentro de un nivel geográfico, se hicieron ajustes al Notebook del Índice de Vulnerabilidad para permitir que los países utilizaran los datos que tenían disponibles, en variables donde faltaban datos en algunas áreas dentro de un mismo nivel de desagregación geográfica.

Información de metadatos publicada

Debido a la variación en la actualización de los datos disponibles en los países participantes, se publicaron hojas de trabajo con el resumen de las variables del Índice de Vulnerabilidad, para proporcionar información sobre la fuente y la fecha de cada una de las variables que los países participantes seleccionaron para calcular su índice.

4. Conclusiones y recomendaciones

Sucesos de alcance internacional como la pandemia del COVID-19 y la necesidad de tomar decisiones para contrarrestarla, pusieron de manifiesto la necesidad de elaborar un Índice de Vulnerabilidad Georreferenciado para la región de Centroamérica que permitiera brindar información estratégica para la gestión de la emergencia sanitaria, facilitando la focalización de los programas de atención para la población más vulnerable y brindando información actualizada y oportuna al público en general.

En ese sentido, fue fundamental para los países, contar con información estadística y geoespacial integrada, que permitiera la comprensión espacial de los fenómenos que reflejaban los datos estadísticos. La integración de la información estadística y geoespacial fortalece el análisis de la información, enriqueciendo el conocimiento desde una perspectiva espacial, y permitiendo identificar tendencias, patrones y distribuciones en el territorio de los diferentes fenómenos v eventos.

La existencia de una amplia cantidad de datos e información sin estandarizar en la región dificultó la interoperabilidad y la integración de los datos producidos por diferentes fuentes. Sin embargo, el trabajo coordinado de los países participantes y la alineación con proyectos e iniciativas globales y regionales mejoraron el proyecto, permitiendo avanzar en los hitos de manera ágil, basados en principios y estándares comunes y promoviendo la comparabilidad de datos entre los países participantes.

Los resultados del provecto en cada uno de los países permitieron sugerir las siguientes recomendaciones:

- Aunque inicialmente se limitaron los datos a las variables solicitadas en la estructura de datos, en el transcurso del desarrollo del provecto. se determinó que las variables de datos sin procesar adicionales eran útiles para solucionar problemas y realizar una revisión de la calidad de los mismos. Por lo que se recomienda que en futuros proyectos de colaboración, se incluyan en la estructura de datos, los valores brutos y proporcionales/normalizados para cada variable.
- Es importante proporcionar más claridad en la estructura de datos, ya que si bien se realizaron algunas mejoras para aclarar su estructura durante el transcurso del proyecto, continuaron presentes elementos que causaron confusión en la interpretación de las variables. Una manera de hacerlo es proporcionando en futuros proyectos de colaboración ejemplos de los cálculos, además de una descripción.
- Se presentaron varios problemas de calidad de datos durante el desarrollo del proyecto, por lo que se recomienda que para proyectos futuros, se proporcionen procedimientos/herramientas de control de calidad que los países participantes puedan ejecutar para la verificación de sus datos antes de cargar sus archivos. Adicionalmente se recomienda desarrollar herramientas que ayuden a automatizar algunos de los controles de calidad.
- Existen muchas variables de datos adicionales que podrían considerarse en el futuro; pero por la coyuntura actual se consideró que se podrían agregar variables de vacunación para COVID19, desnutrición y otras enfermedades de notificación obligatoria.
- Adicionalmente, existe la necesidad de contar con niveles de mayor desagregación geográfica, que permitan análisis más completos y detallados para la toma de decisiones basadas en evidencias.

Este proyecto es un ejemplo de cooperación regional e interinstitucional que permitió llevar a cabo los trabajos de integración de información estadística y geoespacial a fin de dar respuestas o posibles aproximaciones a soluciones para prevenir situaciones en el ámbito del riesgo social. Es importante resaltar que esta metodología del Índice de Vulnerabilidad puede

servir para analizar otros desastres como inundaciones, deslizamientos, terremotos, erupciones volcánicas, etc.

Cabe mencionar que hoy más que nunca se hace imperioso fortalecer la integración de la información estadística y geoespacial a nivel regional y global, así como propender a que las entidades productoras de la información adopten estándares que faciliten la integración de la información obtenida por diferentes fuentes.

Referencias

- Amram, O., Amiri, S., Lutz, R. B., Rajan, B., & Monsivais, P. (2020). Development of a vulnerability index for diagnosis with the novel coronavirus, COVID-19, in Washington State, USA. Health & Place, 64, 102377. 10.1016/j.healthplace.2020.102377
- Casanova, R., Gómez, P. M., Hernández, A. M., & Santiago, A. R. (2021). Americas' Geospatial Response to COVID-19. In COVID-19 Pandemic, Geospatial Information, and Community Resilience (pp. 245-254). CRC Press. DOI: https://doi.org/10.1201/9781003181590-22
- Chen, X., & Wang, F. (2016). Integrative spatial data analytics for public health studies of New York state. AMIA Annual Symposium Proceedings, 391
- DANE, https://geoportal.dane.gov.co/visor-vulnerabilidad/, [accessed: 15.01.2022].
- Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billa, L. (2020). Spatial analysis and GIS in the study of COVID-19. Science of The Total Environment, 140033. https://doi.org/10.1016/j.scitotenv.2020.140033
- García de León Loza, A. (2020). Indicadores básicos y tendencias espacio-temporales en 20 países por mortalidad COVID-19.
 - http://ri.unlu.edu.ar/xmlui/handle/rediunlu/745
- INEC & IGN (2021). Costa Rica's Development of a Vulnerability Index A COVID-19 Response. https://centralcollaboration.interamericangeoportal.org/pages/ costa-ricas-covid-19-response
- IPGH: Global Statistical and Geospatial Framework e-Learning Tools, https://www.ipgh.org/gsgf-e-learningtool.html
- IPGH (2020a). PAIGH Technical Assistance Project: The Integration of Geospatial and Statistical Information in Central America. Project Update.
- IPGH (2021a). PAIGH Technical Project: Data Challenges and Lessons Learned. https:// centralcollaboration.interamericangeoportal.org/documents/datachallenges-and-lessons-learned-report/explore
- IPGH & UNGGIM_a. Americas Project to Integrate Statistical and Geospatial Information in Central America. https://centralcollaboration.interamericangeoportal.org/
- IPGH & UNGGIM_b: Central American Vulnerability Index Work: COVID-19 Response. https://centralcollaboration.interamericangeoportal.org/pages/covid-19response-in-central-america
- Kamel Boulos, M. N., & Geraghty, E. M. (2020). Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around

- the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. International journal of health geographics, 19(1), 1-12. DOI: https://doi.org/10.1186/s12942-020-00202-8
- Kandel, N., Chungong, S., Omaar, A., & Xing, J. (2020). Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries. The Lancet, 395(10229), 1047-1053. DOI: https://doi.org/10.1016/S0140-6736(20)30553-5
- Macharia, P. M., Joseph, N. K., & Okiro, E. A. (2020). A vulnerability index for COVID-19: spatial analysis at the subnational level in Kenya. BMJ global health, 5(8), e003014. DOI: http://dx.doi.org/10.1136/bmjgh-2020-003014
- Mbunge, E., Akinnuwesi, B., Fashoto, S. G., Metfula, A. S., & Mashwama, P. (2021). A critical review of emerging technologies for tackling COVID-19 pandemic. Human behavior and emerging technologies, 3(1), 25-39. DOI: https://doi.org/10.1002/hbe2.237
- Merodio Gómez, P., Pérez García, M., García Seco, G., Ramírez Santiago, A., & Tapia Johnson, C. (2019). The Americas' spatial data infrastructure. ISPRS International Journal of Geo-Information, 8(10), 432. DOI: https://doi.org/10.3390/ijgi8100432
- Murugesan, B., Karuppannan, S., Mengistie, A. T., Ranganathan, M., & Gopalakrishnan, G. (2020). Distribution and Trend Analysis of COVID-19 in India: Geospatial Approach. Journal of Geographical Studies, 4(1), 1-9. DOI: https://dx.doi.org/10.21523/gcj5.20040101
- O'Reilly-Shah, V. N., Gentry, K. R., Van Cleve, W., Kendale, S. M., Jabaley, C. S., & Long, D. R. (2020). The COVID-19 pandemic highlights shortcomings in US health care informatics infrastructure: a call to action. Anesthesia and analgesia. DOI: https://doi.org/10.1213/ANE.0000000000004945
- Ressl, R., Martínez, C. L., Camacho, M. E. P., Hruby, F., Rosas, J. M. D., Huesca, M. E. R., & Jimenez-Rosenberg, R. (2020). Mapping Mexican COVID-19 vulnerability at municipal scale. Terra Digitalis.
- Suárez Lastra, M., Valdés González, C. M., Galindo Pérez, M. C., Salvador Guzmán, L. E., Ruiz-Rivera, N., Alcántara-Ayala, I., ... & Garnica-Peña, R. (2021). Índice de vulnerabilidad ante el COVID-19 en México. Investigaciones geográficas, (104). DOI: https://doi.org/10.14350/rig.60140
- UN-GGIM: Integrated Geospatial Information Framework (IGIF), https://ggim.un.org/IGIF/part1.cshtml
- Whitelaw, S., Mamas, M. A., Topol, E., & Van Spall, H. G. (2020). Applications of digital technology in COVID-19 pandemic planning and response. The Lancet Digital Health, 2(8), e435-e440. DOI: https://doi.org/10.1016/S2589-7500(20)30142-4
- Wisner, B., Blaikie, P., Cannon, T., & Davis, I. (2014). At risk: natural hazards, people's vulnerability and disasters. Routledge. DOI: https://doi.org/10.4324/9780203714775