



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE POSGRADO EN ECONOMÍA

FACULTAD DE ECONOMÍA ♦ DIVISIÓN DE ESTUDIOS DE POSGRADO

**Generación de información económica oficial
a través de fuentes alternativas de obtención
de datos**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctora en Economía

PRESENTA:

Paloma Merodio Gómez

TUTOR:

Dr. Igor Lugo Olmos

CENTRO REGIONAL DE INVESTIGACIONES MULTIDISCIPLINARIAS

MIEMBROS DEL JURADO:

DR. CLEMENTE RUÍZ DURAN

FACULTAD DE ECONOMÍA

DR. MANUEL SUÁREZ LASTRA

INSTITUTO DE GEOGRAFÍA

DR. JOSÉ GASCA ZAMORA

INSTITUTO DE INVESTIGACIONES ECONÓMICAS

DRA. NAXHIELLI RUÍZ RIVERA

INSTITUTO DE GEOGRAFÍA

Ciudad Universitaria, Cd.Mx. 2023



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

1. Introducción	3
1.1. Planteamiento del problema	4
1.2. Objetivos generales y particulares	6
1.3. Preguntas	8
1.4. Hipótesis de trabajo	8
2. Marco Teórico	10
2.1. Marco conceptual para la medición de la pobreza	11
2.1.1. Definición de áreas desfavorecidas	19
2.1.2. Dominios de estudio	23
2.2. El concepto de vulnerabilidad ante desastres	28
2.2.1. Contexto internacional para la sistematización de la información de desastres	32
2.3. Teoría de percepción remota	35
3. Método	40
3.1. Los datos	40
3.1.1. Datos estadísticos	40
3.1.2. Datos raster	43
3.2. La Retícula	46
3.3. Selección de estudio de caso	53
3.4. Observaciones de agua desde el espacio	59
3.4.1. El código	62
3.5. Clasificación	63

3.6. Método del trabajo de campo	68
3.6.1. Criterios de selección de los entrevistados	69
3.6.2. Cuestionario	69
4. Resultados	71
4.1. Validación cuantitativa	74
4.1.1. Índice de Moran	76
4.1.2. Datos con mayor desagregación	78
4.1.3. Validación con modelos nulos	79
4.2. Efectos temporales en el ingreso	83
4.3. Validación en campo	85
5. Discusión	90
5.1. Teoría Económica	90
5.2. Herramientas para la evaluación de daños y pérdidas ante los efectos del cambio climático	92
5.3. Reproducibilidad de la Ciencia	96
6. Conclusiones	97
A. Guía de Entrevista de Campo	101
B. Indicadores del Índice de Rezago Social	104
C. Código	106
Referencias	114

«El recurso más valioso del mundo ya no es el petróleo sino los datos.»

Semanario The Economist, mayo 2017.

1. Introducción

La motivación de esta investigación proviene de la necesidad de contar con información oportuna para la toma de decisiones. A menudo, los hacedores de política pública se ven obligados a actuar con un acceso limitado a datos. Esto puede deberse a que existen datos que no se han actualizado en un largo periodo de tiempo o a que los datos disponibles no cuentan con la desagregación necesaria. Por ejemplo, para hacer política de género, necesitamos estadísticas desagregadas por sexo ya que si no tenemos la desagregación hombre-mujer no podemos conocer las brechas de género. Otro ejemplo podría ser el caso que, debido a que no se cuenta con la desagregación geográfica necesaria para observar diferencias regionales, se limita la focalización de las políticas de desarrollo regional.

La necesidad de contar con más y mejores datos se está dando en un contexto en donde la economía global se está viendo afectada por un avance tecnológico acelerado. El futuro de la humanidad se verá afectado por el desarrollo de la tecnología. Cada vez más, utilizamos herramientas digitales e inteligencia artificial en nuestra cotidianidad.

Por ello, en esta investigación planteo investigar cómo el uso de datos a

través de técnicas innovadoras para la toma de decisiones. A continuación, se especifica el planteamiento del problema, las preguntas que la investigación deberá responder, el marco teórico que se utilizará. Empleo metodologías innovadoras en el manejo de datos para el cálculo empírico de un índice de rezago social y su integración con observaciones de la Tierra para el mapeo de áreas desfavorecidas. Posteriormente, muestro los resultados para el estudio del caso en la Cuenca Hidrológica Río Grijalva-Villahermosa. Por último planteo las conclusiones del trabajo de investigación.

1.1. Planteamiento del problema

La localización de la población más pobre y la focalización de asistencia ante emergencias, como ha sido latente con la reciente pandemia por COVID-19, es un reto mayor que requiere información precisa y confiable. Las encuestas y los censos nos proveen elementos para dicha encomienda. Tradicionalmente, el ingreso de los hogares y el acceso a servicios ha sido medido a través de encuestas. Muchas veces estos programas estadísticos resultan costosos. A manera de ejemplo, la Encuesta Nacional de Ingreso Gasto de los Hogares ENIGH, el principal insumo para la medición de la pobreza en México costó aproximadamente 500 millones de pesos en 2020. Mientras que, al Censo de Población y Vivienda del mismo año se le asignaron alrededor de 8 mil millones de pesos. Estos instrumentos estadísticos se vuelven poco eficientes para proveer información oportunamente, por lo que fuentes no tradicionales que han surgido en la Era Digital proveen una gran oportunidad para generar información oportuna.

La información proveniente de fuentes no tradicionales se obtiene de la digitalización de trámites gubernamentales, el uso de aplicaciones móviles y los avances tecnológicos. Algunas fuentes no tradicionales de datos que empiezan a ser utilizadas en la generación de información estadística oficial son:

- Registros Administrativos. Datos que ya existen en la Era Digital pero no son homogéneos, no se encuentran sistematizados, y deben de ser tratados para poder ser usados estadísticamente. Algunos ejemplos en México de registros administrativos son: registro de empleo del Instituto Mexicano del Seguro Social; información de salud de la población como el registro de diabéticos; facturas del Servicio de Administración Tributaria, entre otros. Para el caso específico de la población en vulnerabilidad, los padrones de los programas sociales son cruciales. El reto para países en desarrollo como México, consiste en contar con dichos padrones ordenados con un identificador único, que permita hacer cruces e identificar a las personas.
- Datos móviles. El uso de teléfonos celulares ha hecho posible el consumo de datos móviles para aproximar el estado socioeconómico de los hogares [1]. Un considerable número de la población mundial cuenta con un teléfono a través del cual realiza operaciones financieras, recibe consultas médicas y navega aplicaciones de mapas.
- Observaciones de la Tierra. El acceso a datos provenientes de satélites ha aumentado significativamente en años recientes. Programas como

Landsat de la NASA, o el programa Copernicus de la Agencia Espacial Europea, proveen imágenes de satélite gratuitamente. Esto ha permitido que la investigación que utiliza las imágenes como fuente primaria se haya incrementado exponencialmente en los últimos años. Algunos ejemplos son: 1) Avances recientes en aprendizaje automático con imágenes satelitales de alta resolución para identificar pobreza regionalmente [2] 2) Imágenes nocturnas. Algunos estudios han mostrado relación entre luminosidad y desarrollo económico [3]; 3) A medida que los satélites han avanzado en proveer imágenes de alta resolución, la literatura sobre mapeo de áreas vulnerables o “slums” ha aumentado [4].

Esta investigación está dirigida a todos los académicos y generadores de información interesados en el uso de fuentes de datos no convencionales. En especial, para aquellos investigadores sociales que estudian el fenómeno de la pobreza y requieren información más oportuna a un menor costo. La investigación es relevante para los productores de información como las Secretarías de Estado a efecto de que conozcan las posibles fuentes de información y los diferentes métodos que se pueden desarrollar con el fin de ampliar el abanico de insumos utilizados para el diseño y evaluación de políticas públicas.

1.2. Objetivos generales y particulares

Los objetivos generales son:

- Estudiar la relevancia del uso de datos no tradicionales en las ciencias económicas.
- Plantear soluciones para el manejo de datos no tradicionales en la producción de información espacial sobre vulnerabilidad.

Objetivos particulares:

- Sensibilizar sobre las limitantes actuales en la generación de información a través de fuentes tradicionales como son las encuestas y los censos.
- Acercar las técnicas de inteligencia artificial a las ciencias sociales, en específico en la generación de información estadística y geográfica.
- Ejemplificar el uso de metodologías de aprendizaje de máquina y otras técnicas innovadoras para la localización de zonas desfavorecidas.
- Analizar las diversas metodologías de percepción remota para la generación de información oficial, en específico el seguimiento a la pobreza urbana.
- Proponer soluciones innovadoras para habilitar el uso de imágenes de satélite a través de inteligencia artificial para el mapeo de áreas deprimidas.
- Ejemplificar el uso de observaciones de la Tierra en la localización de áreas desfavorecidas afectadas por desastres.

1.3. Preguntas

¿Qué son los datos no tradicionales y cómo se relacionan con la generación de información?

¿Los datos provenientes de fuentes no tradicionales se pueden traducir en información oficial y específicamente en la ubicación de zonas desfavorecidas?

¿Cuál es el marco teórico utilizado en el uso de datos no tradicionales para la generación de indicadores de pobreza?

¿Qué componentes de un índice de pobreza multidimensional pueden ser estimados a partir de imágenes de satélite?

¿La información generada a través de algoritmos de aprendizaje de máquina obtenidos de imágenes de satélite puede aproximar índices de pobreza?

¿Es posible identificar áreas desfavorecidas en las grandes urbes a través de fuentes de datos alternativos como son las imágenes de satélite clasificadas a través de técnicas de percepción remota?

¿Las imágenes de satélite pueden detectar oportunamente los hogares vulnerables afectados por un desastre?

¿Cuáles son los límites de las fuentes no tradicionales?

1.4. Hipótesis de trabajo

- Técnicas innovadoras de manejo de fuentes de datos no tradicionales como son las imágenes de satélite permiten generar información confiable, veraz, pertinente y oportuna a partir de datos generados por las nuevas tecnologías.
- El uso de imágenes de satélite ofrece una alternativa en la localización

de la población desfavorecida que reduce los costos y provee información de una manera más oportuna.

2. Marco Teórico

La investigación aquí propuesta es de carácter multidisciplinaria. Por un lado, se analiza la teoría económica relativa a la pobreza y la medición del bienestar. Por otro, se responden las preguntas que planteamos en la sección anterior con el método propuesto para identificar áreas desfavorecidas, compuesto por una interacción de distintas disciplinas: la teoría económica, la geoespacial, la ciencia computacional y la estadística. La transversalidad del método utilizado permite analizar el fenómeno de la pobreza, desde una visión en la que se considera una colectividad de agentes heterogéneos, que al interactuar entre sí, producen un fenómeno sofisticado como lo es la pobreza [5].

Dicho fenómeno, no puede ser deducido del comportamiento de agentes individuales. Tampoco se puede explicar por una sola dimensión o variable. Por lo anterior revisamos, primero, modelos económicos de las ciencias sociales y analizamos el pensamiento económico; en específico la teoría social y la relacionada con la teoría del bienestar. Adicionalmente, la teoría de la complejidad, la pieza del rompecabezas que embona a través de las ciencias computacionales (el aprendizaje de máquina) y el uso de fuentes no tradicionales. Por último, analizamos la teoría relativa a técnicas de percepción remota y metodologías de clasificación de imágenes para ejemplificar el uso de imágenes de satélite en la medición y el mapeo de áreas desfavorecidas.

2.1. Marco conceptual para la medición de la pobreza

Desde el punto de vista conceptual, la medición del ingreso ha sido el enfoque tradicional en la teoría económica en la medición de la pobreza y el bienestar. Es decir, un enfoque utilitarista de la medición de la pobreza, o como lo determinó Foster [6] una pobreza económica relacionada con el acceso a bienes de consumo.

Los análisis comparados del Banco Mundial iniciaron en 1940 y entendían la pobreza como una operación estadística relativa a los ingresos per cápita de las diferentes naciones. El Banco Mundial determina que una persona es pobre si se ubica por debajo de la línea de pobreza con base en el promedio de las líneas de pobreza de los 15 países en desarrollo más pobres la cual, se cuantifica en 1.90 dólares al día. Eso quiere decir que aquellas personas que estén por debajo de un ingreso de 1.90 serían consideradas pobres. Además, el Banco Mundial promueve acciones y recomendaciones para reducir la pobreza estableciendo metas de largo plazo, en 2030 se estableció como meta reducir la pobreza al 3 por ciento de la población y aumentar los ingresos del 40 por ciento más pobre de la población [7].

Por otro lado, la Organización para la Cooperación y el Desarrollo Económicos (OCDE) proporciona información para promover una comparación internacional con base en la distribución del ingreso de diversas economías, así como, la medición del consumo y de la riqueza. Con el fin de proporcionar

más información, e incorporar el mayor número de dimensiones posibles. La OCDE publicó una medición [8] que en conjunto incorpora el ingreso, el consumo y la riqueza. En el caso de la distribución conjunta, se trata de una medición experimental cuya metodología no es aplicable a todos los países, debido a la falta de disponibilidad de información de la variable “riqueza”, ya que muchos países no pueden recabar la información tal y como lo establece la metodología de la OCDE. Sin embargo, el uso de metodologías estandarizadas permite la comparabilidad entre los diferentes países y de ese modo aporta una visión global de la problemática y de las recomendaciones que se pudieran generar, para la implementación de políticas públicas que incidan en reducir la pobreza.

Como indica Ravallion [9], es muy relevante seleccionar detenidamente lo que se quiere investigar ya que dependiendo del concepto que se quiere estudiar, existen distintas metodologías para su análisis. Si hablamos de pobreza, la teoría nos lleva a las definiciones de pobreza absoluta y relativa, la perspectiva objetiva y la subjetiva. Adam Smith [10], establece la siguiente definición de pobreza:

“...una falta de aquéllas necesidades que la costumbre de un país hace que sea indecente, tanto para la gente acomodada como para la de clase más baja, carecer de ellas”.

Un marco teórico más reciente utiliza distintas dimensiones para medir el bienestar. Sen [11], argumenta que el nivel de vida de un individuo lo determinan sus capacidades y no los bienes que posea ni la utilidad que experimente (enfoque de las capacidades) [12]. Así, según Sen, se pueden entender las “ca-

pacidades” como aquellas actividades que distintos objetos permiten realizar. Sen propone un conjunto de criterios que establecen las propiedades que un indicador debe cumplir para medir pobreza. A tales propiedades les llama axiomas [12].

De la misma forma, Alkaire [13], contribuye con la investigación de la medición multidimensional de la pobreza. Algunas de las dimensiones utilizadas en los enfoques multidimensionales, además de las medidas de ingreso y riqueza, incorporan el acceso al empleo, educación, tipo de vivienda, acceso a salud, la calidad del medio ambiente, involucramiento en su comunidad y conexiones sociales, entre otras. Éstas últimas medidas, tienden a tener el nombre de carencias, debido a que se miden en razón de si, una persona cuenta o no con ciertos bienes y servicios.

En un paso intermedio entre el enfoque utilitarista y el de capacidades, Ravallion [9] conecta la utilidad con el consumo de bienes [12]. Las capacidades dependen del consumo de bienes, q , y de las características del hogar, x ; así, si c representa a las capacidades, se puede expresar como función de q y x [12]:

La utilidad es una función de las capacidades [12] que denota por

$$U = U(c). \tag{1}$$

Si reemplazamos el término c en la función de utilidad por la función

$$c(q, x) : \tag{2}$$

Con ello, Ravallion [9] complementa el enfoque utilitarista sin descartar

el enfoque de capacidades.

El premio Nobel de economía Josef Stiglitz hace la siguiente analogía: si manejas un coche, tienes dos indicadores básicos, la velocidad y la cantidad de gasolina que queda. Sin embargo, en el tablero pueden encenderse indicadores adicionales en caso de ser necesarios, como el de la falta de aceite. Esta analogía nos dice que el nivel de ingreso o riqueza, puede ser un indicativo, pero cuando algo falla, la necesidad de contar con más información y precisión sobre lo que está sucediendo, es fundamental para encontrar cómo encender el auto y lograr la velocidad deseada. Es por ello que, la teoría económica, se ha concentrado en la medición del bienestar, que incluye estas variables que van más allá del ingreso.

El método del bienestar estudia funciones de utilidad individual que dependen del ingreso de la unidad en cuestión [12]. Gini, por su parte, pondera una función que indica la posición relativa que dicha unidad de análisis ocupa en la población. Este índice, ha sido muy utilizado y conocido como el índice de Gini que se emplea para medir la desigualdad y hacer comparativos internacionales.

Por su parte, Atkinson [14], establece la función de bienestar a través de la introducción de índices de desigualdad [15]. Los indicadores de desigualdad miden la equidad, a través, de la distribución del ingreso. El concepto de ingreso distribuido equitativamente está estrechamente relacionado con el de una prima de riesgo o certeza equivalente en la teoría de la toma de

decisiones bajo incertidumbre. Por ello, el concepto de desigualdad está vinculado al ingreso propio y el de los demás como una medida relativa de la población [12].

En México, se definió a través de la Ley de Desarrollo Social que la medición de la pobreza fuera proporcionada por el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) y a partir de 2009 se aplica la medición multidimensional de la pobreza [16], la cual, está conformada por tres esferas, una que mide el bienestar económico, otra los derechos sociales y, por último, la esfera del contexto territorial.

La primera está definida por una línea de pobreza y pobreza extrema por ingresos, mismas que se establece con base en la conformación de una canasta básica alimentaria y no alimentaria. Para el caso de la línea de pobreza por ingresos, se utilizan tanto la canasta alimentaria como la no alimentaria mientras que para la línea de pobreza extrema solo se emplea la canasta alimentaria.

En cuanto a los derechos sociales, se trata de aquellos derechos mínimos que los ciudadanos pueden ejercer y se definen como carencias, cuando su ejercicio se encuentra mermado. El marco de los derechos humanos está basado en los siguientes rubros: salud, seguridad social, vivienda; servicios básicos de vivienda, alimentación nutritiva y de calidad y educación. Con base en estos elementos, se define si una persona carece o no de dichos atributos, mediante la construcción de 6 indicadores sociales a través de los cuales, de

forma dicotómica, se define si una persona carece o no de alguno de estos derechos.

Por otro lado, la esfera del espacio territorial se determina por el grado de cohesión social y por el grado de accesibilidad a carreteras. Las características analizadas son la distribución vial, la calidad de la red de caminos, la pendiente del terreno, características ambientales en términos de las barreras que pudieran existir para acceder a las vialidades, la identificación geográfica de las localidades con respecto a las carreteras, la disponibilidad de transporte público y por último, el tiempo de desplazamiento a los centros de servicio, para llegar de una localidad al centro de servicio deseado, ya sea un hospital, escuela, etc.

Es importante mencionar que, las variables mencionadas en la metodología de CONEVAL se obtienen de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) para las mediciones nacionales y estatales, y otras fuentes adicionales como el Censo Nacional de Población y Vivienda y la Encuesta Intercensal para la medición a nivel municipal. Estas fuentes son proporcionadas por el Instituto Nacional de Estadística y Geografía (INEGI).

CONEVAL reconoce que aún hay trabajo por hacer con respecto a las variables incluidas en la esfera de las carencias sociales como, por ejemplo, en la incorporación de la calidad de los servicios y no solo de la disponibilidad de éstos para la ciudadanía. También, sugiere incorporar variables como lo son, el acceso a la infraestructura y al empleo, e incluir el tema de discriminación,

los cuales, son aspectos que podrían incidir en el nivel de vida de la población y que actualmente no se toman en cuenta en la medición de la pobreza en México [17].

Todas estas metodologías requieren fuentes de información. Se obtienen de levantamientos mediante encuestas, ya sea en papel, o de forma electrónica, recopilando los datos de casa en casa. La disponibilidad de información para el cálculo de la pobreza se vio amenazada ante la pandemia mundial del SARS COV-2. Los Institutos de Estadística se vieron forzados a detener los levantamientos de información con el fin de evitar el contacto físico entre personas y así frenar el esparcimiento del virus. Esto provocó que muchas encuestas tuvieran que recalendarizar la publicación de una serie de productos de información o en su caso, buscar fuentes alternativas para obtener información.

Hasta aquí, se han mostrado algunas de las variables y métodos tradicionales más utilizados para medir pobreza. Sin embargo la pandemia obligó a suspender el levantamiento de encuestas lo cual evidenció la necesidad de innovar a través de nuevas técnicas y herramientas que nos permitan buscar alternativas para recolectar información. Es por ello que, esta investigación busca visibilizar las aportaciones del uso de la información geoespacial como fuentes alternativas para obtener datos en paralelo al marco estadístico, alrededor de los temas de pobreza, y que se permita una evaluación territorial a través del uso y aprovechamiento de imágenes satelitales.

Con base en la medición multidimensional de la pobreza, se han detectado algunas características de la población denominada como pobre, que podrían ser identificadas con el uso de técnicas espaciales. Es importante en este punto mencionar que la información geoespacial es complementaria y en ningún momento busca sustituir a las mediciones tradicionales. Es decir, se pretende aportar en la medida de lo posible información adicional.

Sobre la dimensión geográfica de la pobreza, es evidente su inyección en el marco teórico aquí expuesto. Ahora bien, en el caso de Sen [11], el ejemplo canónico de las capacidades es el siguiente:

Una bicicleta es un bien que posee distintas características, por ejemplo, ser un medio de transporte. Esta característica le da a la persona que posee la bicicleta la capacidad de transportarse, y a su vez esa capacidad le puede proporcionar a la persona utilidad; por lo que, según este razonamiento, los bienes no son los objetos que nos proporcionan el estándar de vida, ya que la posesión de un bien no implica por sí misma las actividades que un individuo pueda realizar, dichas actividades dependen de las facultades e impedimentos de cada individuo. Así, si bien “los objetos proveen la base para una contribución al estándar de vida, no son en sí mismos una parte constituyente de ese estándar” [11].

De nada sirve, tener la bicicleta si no se sabe andar en ella. Y de igual forma, de nada sirve una bicicleta si estamos en una isla y queremos llegar a territorio continental en cuyo caso necesitamos una balsa. Por ello, se puede

decir, que el lugar y el contexto son determinantes para el desarrollo de las capacidades.

En el marco teórico de Ravaillon [9], se concilian la teoría utilitarista con la teoría de capacidades, ya que las características del hogar determinan la utilidad. El techo, el acceso por terracería o por vía rápida, la presencia de áreas verdes, determinan entonces las características de la vivienda que definen el nivel de utilidad alcanzado. Por último, en el caso de Alkaire [13], es clarísimo el aspecto territorial y espacial como una dimensión de la multidimensionalidad de la pobreza.

Por tanto, se considera posible, el uso de imágenes satelitales para aproximar algunas de las carencias sociales enmarcadas por CONEVAL como lo son, las características de la vivienda, a través de la detección del material de los techos, tipo de suelo alrededor de la vivienda, etcétera, así como la accesibilidad a carreteras a través de la observación de la orografía. A continuación, definimos el concepto de áreas desfavorecidas de acuerdo a la revisión de la literatura y destacamos los dominios de estudio que nos permiten identificar las variables que serán empleadas en el análisis.

2.1.1. Definición de áreas desfavorecidas

La necesidad de contar con sistemas de información confiables y veraces que sustenten la localización y caracterización de la población que permita la correcta focalización de políticas públicas, es un reto en la mayoría de los

países en desarrollo.

En específico, la ubicación de las personas más vulnerables para focalizar apoyos tiene el desafío sobre la definición de “vulnerabilidad”. El término *barrio marginado* se ha utilizado para menospreciar y discriminar a grupos en algunos contextos, y es utilizado como marcador de identidad entre los residentes en otros entornos [18]. Las palabras *favela*, *gueto*, *asentamiento informal*, *barrio marginado* o *slum* también son términos comunes en algunas ciudades; sin embargo, cada una de estas etiquetas viene con una historia política y social específica [19].

Naciones Unidas define un hogar marginado como:

“...un grupo de individuos viviendo bajo el mismo techo en un área urbana, que carece de uno o más de los siguientes condiciones: (1) vivienda duradera que protege contra el clima extremo; (2) suficiente espacio habitable, que significa no más de tres personas compartiendo la misma habitación (3) fácil acceso a agua segura en cantidades suficientes a un precio asequible; (4) acceso a saneamiento adecuado en la forma de un baño privado o público compartido por un razonable número de personas; y (5) seguridad de tenencia que evita los desalojos forzosos [20]”.

Por otra parte, la Declaración de Pretoria de las Naciones Unidas [21] define los asentamientos informales como:

“Áreas residenciales en las cuales: a) los habitantes no tienen seguridad

de la tenencia respecto de las tierras o viviendas en las que habitan, con modalidades que van desde la ocupación ilegal de una vivienda hasta el alquiler informal; b) los barrios suelen carecer de servicios básicos e infraestructura urbana o no tienen acceso a ellos; y c) las viviendas podrían no cumplir con las normas edilicias y planificación vigentes y a menudo están ubicadas geográfica y ambientalmente en áreas peligrosas. Además, los asentamientos informales pueden ser una forma de especulación inmobiliaria para los residentes urbanos de todos los niveles de ingresos, ricos y pobres”.

Es decir, los barrios marginales son los asentamientos informales más desfavorecidos y excluidos, y se caracterizan por la condición de pobreza y las grandes aglomeraciones de viviendas en mal estado, ubicadas, por lo general, en las tierras urbanas más peligrosas. Además de no tener seguridad de la tenencia de las tierras, los habitantes de los barrios marginales no disponen de acceso formal a la infraestructura y los servicios básicos, espacios públicos y zonas verdes, y están constantemente expuestos a desalojos y enfermedades [22].

Los hogares marginados se encuentran, muy a menudo, situados en zonas de riesgo como son las colinas o las orillas de los ríos. Es por lo que, los hogares en condiciones vulnerables sufren constantemente de las afectaciones ocasionadas por desastres, como inundaciones o deslizamientos. Así, las áreas desfavorecidas son aquellas compuestas de hogares marginados que se encuentran situados en zonas de riesgo.

En México, el concepto sobre la tenencia y la ocupación informal de viviendas tiene algunas limitaciones, ya que la Ley General de Asentamientos Humanos, Ordenamiento Territorial y Desarrollo Urbano [23] no contempla la definición de asentamientos irregulares. Sí encontramos, en las Reglas de Operación del Programa para Regularizar Asentamientos Humanos Irregulares, para el ejercicio fiscal 2018 de la Secretaría de Desarrollo Agrario, Territorial y Urbano, SEDATU la definición de Asentamiento Humano Irregular como sigue:

“corresponde a los núcleos de población ubicados en áreas o predios fraccionados o subdivididos sin la autorización correspondiente, e independientemente de su régimen de tenencia de la tierra” [24].

Sin embargo, por alguna razón desconocida, en las Reglas subsecuentes ya no se considera el concepto. Por su parte el INEGI, en el Censo Nacional de Gobiernos Municipales y Demarcaciones Territoriales 2019 [25] define asentamiento humano irregular como el conglomerado demográfico, ubicado en áreas o predios fraccionados o subdivididos sin la autorización correspondiente, cualesquiera que sea su régimen de tenencia de la tierra o se ubique en zonas de restricción, riesgos o de amortiguamiento.

Con el reconocimiento de estas limitaciones, y las múltiples acepciones de los términos antes mencionados, en esta investigación definimos un área desfavorecida como un polígono que presenta un conjunto de factores de riesgo sociales, ambientales y ecológicos a la salud y al bienestar más allá de las características del hogar e individuales. De acuerdo con Thomson [26], las

áreas urbanas desfavorecidas son aquellas caracterizadas por su morfología en el entorno urbano.

Las variables físicas de áreas desfavorecidas incluyen el tamaño, la forma y la altura del edificio; carreteras y otras vías de acceso; densidad de edificios; forma de asentamiento; y ubicación del asentamiento con respecto a espacios públicos verdes, pendientes pronunciadas, zonas de inundación y proximidad a ferrocarriles y alta tensión líneas eléctricas [27]. Las variables sociales pueden incluir niveles de delincuencia; cobertura y calidad de los sistemas de residuos sólidos, agua, saneamiento y energía; proximidad y accesibilidad a escuelas, instalaciones de salud, tiendas, empleo e infraestructura pública; y capital social derivado de organizaciones comunitarias y entre vecinos [28].

La morfología de las áreas desfavorecidas varía de una ciudad a otra y de contexto en contexto. Además, las áreas son dinámicas [29] [30]. En la siguiente sección abordaremos con mayor detalle las variables utilizadas para el mapeo de dichas áreas.

2.1.2. Dominios de estudio

A fin de identificar las variables de estudio, establecemos los dominios que se analizan de acuerdo con los indicadores encontrados en la revisión de literatura que están estrechamente relacionados con la definición de áreas desfavorecidas, lo que nos permite identificar los polígonos que serán caracterizados como marginados. Siguiendo la metodología establecida por Abascal [31] podemos identificar tres distintos dominios de estudio, es decir, niveles en los



Figura 1: Cuatro tipos de hogar deprimido. a) no desfavorecido en área no desfavorecida, b) desfavorecido en área no desfavorecida, c) no desfavorecido en área desfavorecida, d) desfavorecido en área desfavorecida.
Fuente: Thomson, 2020 [26]

cuales se identifican variables que denotan marginación.

El primer dominio, es el relativo a los indicadores a nivel de vivienda, que incluye indicadores socio-demográficos, como por ejemplo, tasas de educación individual, estado de salud, empleo y propiedad familiar de bienes. Dentro de este dominio también se encuentran los indicadores a nivel de vivienda que refleja la calidad de materiales de construcción, el acceso a servicios de agua y saneamiento, y si los ocupantes tienen derechos de tenencia de la tierra.

Dentro del dominio a nivel entorno, se encuentran los indicadores relativos al área, se incluyen cuatro privaciones que se encuentran dentro de los asentamientos: riesgos sociales y activos, peligros físicos y activos, urbanización no planificada y contaminación. Los riesgos sociales incluyen riesgos como la delincuencia y la falta de seguridad, mientras que los activos sociales incluyen la identidad social o cohesión comunitaria. Los peligros físicos contemplan alta probabilidad de inundaciones, deslizamientos de tierra y otras amenazas naturales como terremotos, mientras que los activos físicos incluyen mitigación de recursos y estrategias tales como materiales resistentes a terremotos o presencia de áreas verdes y un aire más limpio.

Los indicadores dentro del dominio de área, relativos a la urbanización no planificada, están asociados con una inmigración rápida y no planificada a un municipio o colonia que podría resultar en viviendas no planificadas, espacios verdes limitados y falta de carreteras. La contaminación refleja acumulación de basura, contaminación del agua, contaminación del aire o altos niveles o

ruido constante que afectan el bienestar de los residentes.

El tercer dominio es el relativo a la interconexión del área y se refiere a la conectividad con los asentamientos circundantes y la integración con el resto de la ciudad. Este dominio es relevante para esta investigación ya que planteo un enfoque complejo en donde no se puede pensar un área o una vivienda como un agente representativo sino, que se debe de analizar cada agente y su conectividad o interacción con el resto [5].

Dentro del dominio de la interconexión destacan indicadores relativos a la infraestructura, refiriéndose a agua, residuos, transporte, y otros sistemas típicamente administrados por el gobierno municipal, así como las instalaciones y servicios que reflejan la disponibilidad, accesibilidad y cercanía de escuelas, establecimientos de salud, establecimientos bancarios, tiendas, iglesias y servicios culturales, y otras instalaciones y servicios necesarios para una ciudad próspera. Los desafíos en las comunidades desfavorecidas son prevenidas y abordadas mediante una planificación y gestión transparentes y eficaces en toda la ciudad que hacen que la gobernanza de la ciudad también se incluya en este dominio final.

El marco aquí planteado es la base para la identificación de las áreas desfavorecidas. La novedad de esta investigación se centra en que el análisis de marginación o pobreza, no se queda en las variables a nivel hogar o vivienda, sino que va más allá, al incluir otras como del entorno y la complejidad de las áreas desfavorecidas.



Figura 2: Dominios de estudio

2.2. El concepto de vulnerabilidad ante desastres

Es relevante destacar que esta investigación integra la relación compleja al concepto de vulnerabilidad a través de su caracterización como un fenómeno que sucede en el hogar, pero que está intrínsecamente relacionado con el entorno y su interacción con las zonas colindantes. Esta visión, permite además, ejemplificar la vulnerabilidad de las zonas desfavorecidas ante desastres.

La vulnerabilidad tiene orígenes sociales, económicos y políticos. De acuerdo con el nivel de vulnerabilidad, las comunidades son susceptibles a daños sociales, ambientales y emocionales. La medición de la vulnerabilidad nos da una idea del grado de daño al que están expuestas las comunidades. Existen distintos métodos para evaluar la vulnerabilidad y no contamos con una metodología universal. Para medir vulnerabilidad se debe definir al sujeto, a qué se es vulnerable y por qué situaciones es vulnerable.

La vulnerabilidad social se define como la incapacidad de desarrollar mitigaciones para evitar el desastre, responder a las emergencias y adaptarse a las consecuencias de la materialización de los riesgos. La evolución del concepto de vulnerabilidad integra una visión en la que la vulnerabilidad no es estática. Es dinámica y es dependiente de la escala [32]. La comunidad, el hogar y la persona, no son vulnerables de manera intrínseca, es decir, los

enfoques normativos [33] marcan que la base de la vulnerabilidad está determinada por las características que acompañan a esa unidad de análisis. Los factores que condicionan la vulnerabilidad son, la educación, la salud, los usos y costumbres, entre otros. Las encuestas socioeconómicas, los registros administrativos, las fotografías, y la experiencia en campo son algunas de las fuentes utilizadas para estimar la vulnerabilidad ante riesgos.

Diferentes conceptualizaciones reconocen la relación entre pobreza, riesgo, gestión del riesgo y vulnerabilidad [34]. En la mayoría de las disciplinas, el enfoque está en los riesgos (en un extremo) o en los riesgos subyacentes (o resultados) en el otro. Lo que algunos (p. ej., gestión de desastres, medio ambiente) llaman vulnerabilidad (o evaluación de vulnerabilidad) es realmente una evaluación de riesgos. El concepto utilizado por otros investigadores (por ejemplo, dinámica de la pobreza, medios de vida sostenibles, seguridad alimentaria, sociología/antropología, salud/nutrición) utilizan al hablar de vulnerabilidad es similar al empleado en evaluaciones de pobreza y se centra en el otro extremo de la cadena de riesgo: en los resultados. Es decir, se enfoca en la base de las condiciones e incapacidades de los hogares para responder a los riesgos.

Hoy en día, la literatura de vulnerabilidad establece el concepto de riesgo como un resultado, a diferencia de los conceptos anteriores donde el riesgo era definido como una distribución probabilística de la amenaza, en la mayoría de los casos, la amenaza física. Anteriormente, se hablaba de la vulnerabilidad como el resultado a la causa, que es el riesgo [32]. En este sentido, la

literatura reciente y su conceptualización desde el año 2010 aproximadamente, establece que la vulnerabilidad determina el riesgo. El shock es estudiado como un evento desencadenante que se divide en: 1) el peligro, representado con la distribución probabilística del evento (por ejemplo, una distribución de lluvias) y 2) la amenaza o peligrosidad definida como la distribución de los efectos de ese evento en el territorio (por ejemplo, cómo cae la lluvia en un terreno específico, no necesariamente la lluvia ocasiona inundación, depende del terreno y otros factores).

Natenzon [35] realiza un análisis a partir del siglo XX de los conceptos de riesgo, catástrofe y vulnerabilidad desde las diferentes disciplinas geográficas y ciencias sociales. Existen diferentes enfoques que se desarrollaron a lo largo de todo el siglo XX y que han dado origen a la pluralidad de visiones sobre la vulnerabilidad. La vulnerabilidad se define como la susceptibilidad al daño, pero tiene que ver en cómo ese hogar, persona o comunidad está en posición desventajosa con respecto a otros. El ser susceptible, hablamos de una condición potencial ante un posible desastre. Es por ello que, en esta investigación cuando nos referimos a la vulnerabilidad ante, nos referimos a la vulnerabilidad ante la amenaza [34]. El origen latín de la palabra vulnerabilidad, nos habla de un daño sobre alguien que ya tiene una susceptibilidad previa. En ese contexto, la vulnerabilidad es una condición que es visible y que se hace evidente ante ciertas fases críticas que pueden manifestarse como un shock o cambios forzados en un sistema estable. El cambio en los estudios de vulnerabilidad social como especificados por Turner [36] nos remontan a la teoría económica neoclásica y los modelos de equilibrio general donde a

partir de un estado estacionario un shock externo desestabiliza el sistema y genera una reorganización hacia un nuevo estado estacionario o simplemente genera un efecto temporal. Esto nos lleva a definir qué es un desastre en la literatura sobre vulnerabilidad social.

Un desastre se define como un estrés externo que desestabiliza un sistema y rompe su equilibrio. Enrique Quarantelli [37], define el concepto de desastre como una ruptura del orden social. A diferencia de Natenzon [38] quien define el desastre como una catástrofe. Es decir, Quarantelli, nos invita a ver el desastre no como una emergencia ni como un efecto aislado sino como un proceso. Gilbert [39] menciona que la idea de los desastres se origina con el paradigma de la guerra y de los términos con los que se conceptualiza. Para Smith [40], el desastre es un proceso, no es un evento aislado, se construye de patrones históricos. A saber, el desastre utiliza conceptos de peligrosidad, vulnerabilidad, exposición e incertidumbre.

La adaptación a un shock, implica combinar estructuras y procesos, renovar el sistema y generar nuevas trayectorias. Así, la capacidad adaptativa puede implicar un estado menos deseable. En este sentido, defino para efectos de esta investigación que las áreas desfavorecidas son áreas vulnerables al daño potencial. Establezco la amenaza o peligrosidad como la distribución de efectos en el territorio que afecta a una población con distintas condiciones normativas.

2.2.1. Contexto internacional para la sistematización de la información de desastres

La necesidad de contar con información de las afectaciones por desastres, los efectos del cambio climático y el cuidado del medio ambiente es cada día más urgente. Las convenciones internacionales han dado pie al establecimiento de marcos para monitorear los avances hacia el desarrollo sostenible.

Existen distintos marcos internacionales con el objetivo de preservar el medio ambiente y apoyar el desarrollo sostenible. La Convención Marco de las Naciones Unidas sobre el Cambio Climático, entró en vigor el 21 de marzo de 1994, con la ratificación de 197 países [41]. La reunión número veintisiete de la Conferencia de las Partes sobre Cambio Climático se llevó a cabo en Sharm El-Sheikh, Egipto (COP27) en 2022. En dicha reunión, se refrendó el compromiso de los estados miembros de detener el calentamiento global. Para ello, se requiere de financiamiento y disminución de las emisiones de gases de efecto invernadero así como la generación de estrategias en los países en el Sur Global. [41]. Por otra parte, el Convenio sobre la Diversidad Biológica (CDB), cuenta con la participación de 196 estados miembros con el objetivo de detener la pérdida de biodiversidad y restaurar los ecosistemas [42]. En la reunión de la Conferencia de las Partes sobre la Diversidad Biológica de Montreal (COP15), se fijó el compromiso de revertir la pérdida de la naturaleza, incluida la protección del 30 por ciento del planeta y el 30 por ciento de los ecosistemas degradados para 2030 [43]. Por su parte, en marzo de 2022,

la Plataforma Regional para la Reducción de Riesgos de Desastres en las Américas y el Caribe, aprobó la declaración de Punta del Este [44], en la cual se establecen acciones a implementar en la región, entre ellas, reducir el impacto negativo de los eventos extremos.

Es de interés para esta investigación señalar que en dichas declaraciones y convenciones se identifican los retos en la disponibilidad, accesibilidad, recopilación, desagregación, procesamiento y la generación de datos que permiten monitorear las metas establecidas. En específico, la declaratoria de Punta del Este, reconoce la necesidad de integrar la ciencia de la geo-información, datos y análisis geoespacial para la toma de decisiones antes, durante y después de los desastres [44].

Para lograr medir las pérdidas y daños asociados con los efectos adversos del cambio climático establecidos por los compromisos internacionales antes mencionados, son necesarios sistemas de información homogéneos, oportunos y de fácil alcance para los hacedores de política pública. Por ello, la resolución de la COP 27 destaca:

“...la necesidad de subsanar las lagunas existentes en el Sistema Mundial de Observación del Clima, en particular en los países en desarrollo, en un tercio del mundo, incluido el 60 por ciento de África, no tiene acceso a servicios de alerta temprana e información climática, por lo que existe la necesidad de mejorar la coordinación de actividades por parte de la comunidad de observación sistemática y aumentar la capacidad de proporcionar información climática útil y procesable para los sistemas de mitigación, adaptación y

alerta temprana, así como información para permitir la comprensión de los límites de adaptación y la atribución de eventos extremos” [41].

Sin información, sin instrumentos de alerta temprana y observación sistemática para la evaluación del riesgo de desastres, el combate del cambio climático y la prevención de desastres sería inútil. El Secretario General de las Naciones Unidas, António Guterres, ha hecho un llamado para contar con una cobertura universal de los sistemas de alerta temprana contra los fenómenos meteorológicos extremos y el cambio climático [41]. Como veremos más adelante en la sección de análisis cualitativo, esto coincide con los hallazgos encontrados en las entrevistas en campo de esta investigación, donde los entrevistados, destacan la necesidad de contar con un sistema de alerta temprana que apoye a la población a reaccionar ante el posible riesgo por desastres. Para ello, se requiere un sistema de información de calidad que permita reaccionar con oportunidad ante posibles amenazas.

Adicionalmente, la COP reconoce la importancia vital de contar con sistemas sólidos de observación de la Tierra y datos a largo plazo para una mejor comprensión del sistema climático global y sus cambios. Esto permitirá implementar acciones de adaptación y mitigación [41]. Además, hace hincapié en la necesidad de abordar las lagunas de observación sistemáticas, en particular en países en desarrollo y para las regiones oceánicas, montañosas, desérticas y polares para mejorar la comprensión del cambio climático, los riesgos relacionados con el clima, límites de adaptación y garantizar una mejor prestación de servicios climáticos y sistemas de alerta temprana. Es por

ello que, en la siguiente sección abordamos la teoría de percepción remota para utilizar las imágenes satelitales como un insumo en la generación de información en beneficio de la población en zonas desfavorecidas.

2.3. Teoría de percepción remota

La nueva geografía económica de Paul Krugman establece la conexión entre sistemas económicos desde una perspectiva multidimensional. Krugman [45], muestra la influencia de la teoría de la complejidad en la economía y desarrolla un modelo que es ampliamente considerado por haber dado origen a la nueva geografía económica, y ciertamente ha estimulado el surgimiento de una nueva ola de teorización y trabajo empírico.

En este sentido, y siguiendo la narrativa presentada por Colander [46], el estudio de las urbes y la concentración de áreas desfavorecidas amerita un enfoque complejo en donde se estudien diversas interacciones [47]. La expansión global de los barrios urbanos plantea preguntas para la investigación económica con enfoque geográfico. Adicionalmente, en los años venideros, la agenda de política pública deberá contar con un importante contenido económico-geoespacial debido a la creciente concentración de la población en las urbes. Los retos de planeación urbana son enormes: barrios marginados, transporte deficiente, mala calidad del aire y acceso limitado a servicios básicos como agua y electricidad. Existe la necesidad de contar con herramientas cuantitativas que permitan diseñar y dar seguimiento a una agenda urbana

sostenible en favor de las comunidades más marginadas [48].

De acuerdo con el *Union of Concerned Scientists (UCS)* [49], al primero de abril de 2020, hubo un total de 2,666 satélites en el espacio. Estos satélites giran alrededor de la Tierra y van tomando fotografías de nuestro territorio en su trayecto. Dichas imágenes pueden ser utilizadas para dar seguimiento a distintos fenómenos que suceden en el territorio como son el crecimiento de las ciudades, cambios en la vegetación, entre otros.

La disponibilidad de literatura científica sobre mapeo de áreas desfavorecidas que emplea métodos de teledetección, es decir obtener la información del objeto a través de instrumentos de escaneo inalámbricos, ha aumentado, desde que existe mayor disponibilidad de sensores de muy alta resolución en los satélites que giran alrededor del planeta Tierra.

Por ello, diversos estudios han mapeado las zonas desfavorecidas de las ciudades a través de imágenes de satélite. Esto beneficia la generación de información más frecuente para el diseño y monitoreo de políticas públicas destinadas a reducir la brecha de desigualdad ya que provee información con mayor oportunidad y con mayor desagregación geográfica de los datos.

Varios investigadores han incursionado en combinar imágenes satelitales con datos de censos de población, utilizando el procesamiento de imágenes apropiado y las técnicas computacionales. Engstrom [50], utilizó datos del censo de Sri Lanka y mostró que los indicadores satelitales rastrean las

diferencias regionales en pobreza extremadamente bien, mostrando que las imágenes satelitales de alta resolución son un valioso complemento para los datos de las encuestas de hogares, con potencial para ayudar a generar mapas de pobreza locales más precisos y actualizados y refinar la focalización en las iniciativas de desarrollo.

Por otro lado, se han utilizado indicadores como material de pared, hacinamiento, acceso a agua entubada, conexión de saneamiento a alcantarillas y propiedad para construir un índice de barrios desfavorecidos basado en datos del censo de la ciudad de Medellín [51]. Este índice, en comparación con la información basada en imágenes satelitales sobre la cobertura del suelo, características estructurales y basadas en texturas, mostró que la información basada en imágenes podría explicar el 59 por ciento del índice de barrios marginales.

Así mismo, Xie [52] y Jean [2] detectan asentamientos vulnerables o marginados al entrenar una red neuronal para detectar luz nocturna con la cual predicen consumo per cápita. Por su parte, Roy [53] genera el índice de marginación en los barrios de la Ciudad de México. También Shi [3] desarrolla un índice de pobreza basado en datos de percepción remota e información sobre actividades humanas. Mientras que Mason y Fraser [54] demostraron la extracción automática de viviendas con condiciones precarias a partir de imágenes aéreas de muy alta resolución.

El mapeo de barrios marginados a partir de técnicas de percepción remota

se ha enfocado en cuatro dimensiones: factores contextuales, características físicas de los barrios, datos y requisitos, y métodos de extracción de barrios marginales [4]. Metodologías exitosas muestran diversidad en términos de niveles de información extraídos de las imágenes (basados en área u objeto), conjuntos de indicadores implementados (conjuntos individuales o grandes) y métodos empleados como, por ejemplo, análisis de imágenes basado en objetos a través de aprendizaje máquina.

Para el mapeo de barrios marginales locales, las metodologías basadas en aprendizajes de máquina muestran buenas capacidades de extracción de información de áreas y objetos. Estos métodos buscan establecer una relación sistemática entre los elementos de la imagen y las características de las áreas desfavorecidas para entrenar algoritmos capaces de analizar variaciones en sus morfologías y facilitar su monitoreo. Por ejemplo, Dorji [55], demuestra que, utilizando una metodología de clasificación aumenta la precisión en la estimación del ingreso.

Kuffer [4] hace un análisis de los últimos 15 años de publicaciones relacionadas con el mapeo de zonas desfavorecidas a través de imágenes de satélite. Encuentra que, los algoritmos de aprendizaje automático tienen las precisiones más altas reportadas y permiten trabajar con grandes conjuntos de indicadores de una manera computacionalmente eficiente. Para el mapeo de zonas desfavorecidas a nivel local, Kuffer encuentra que los enfoques OBIA¹ muestran buenas capacidades para extraer información basada en

¹Análisis de imágenes basada en Objetos [56] (OBIA siglas en ingles): técnica utilizada

áreas y objetos. En última instancia, es esencial establecer una relación más sistemática entre los elementos de la imagen y las características de los barrios marginados para entrenar algoritmos capaces de analizar variaciones en las morfologías de las áreas desfavorecidas [4].

También se ha utilizado la interpretación visual de imágenes de muy alta resolución para extraer concentraciones de pobreza urbana [58]. Baud et al, [59] emplea imágenes de muy alta resolución para identificar patrones precisos dentro de áreas marginadas aprovechando variables físicas como la densidad de construcción, el tamaño del edificio y la idoneidad del sitio.

Por lo tanto, la cuantificación y la localización de las áreas desfavorecidas requiere métodos para identificarlos y definirlos espacialmente [60]. Debido a la expansión de los asentamientos que ocurren con el crecimiento urbano, las técnicas automáticas favorecen el monitoreo de los barrios marginados [27].

A continuación, expongo los datos y el método utilizado en esta investigación para el mapeo de zonas desfavorecidas.

para analizar las imágenes satelitales, basada en la información de un conjunto de píxeles similares llamados objetos. Estos objetos de la imagen son los grupos de píxeles que son similares entre sí sobre la base de una medida de las propiedades espectrales (es decir, color), tamaño, forma y textura, así como el contexto de un barrio que rodean los píxeles [57].

3. Método

3.1. Los datos

Esta investigación es interdisciplinaria, en gran medida por los tipos de datos que utiliza. Por un lado, datos censales provenientes de ejercicios estadísticos tradicionales como los Censos y por otra parte el uso de datos geoespaciales. En esta sección describo los datos a utilizar para realizar la clasificación de áreas desfavorecidas.

3.1.1. Datos estadísticos

Los datos estadísticos que utilizo provienen del Censo de Población y Vivienda para el año 2020. El Instituto Nacional de Estadística y Geografía (INEGI), con fundamento en lo dispuesto por la Constitución Política de los Estados Unidos Mexicanos en su artículo 26, apartado B, y en apego a las atribuciones que la Ley del Sistema Nacional de Información Estadística y Geográfica le otorga, realiza el Censo de Población y Vivienda cada diez años. Los datos del Censo, son gratuitos y se pueden consultar en su página de internet.

El Censo se realiza para levantar, recopilar, monitorear, analizar y publicar datos demográficos, económicos y sociales, relativos a la población de nuestro país, así como información sobre las viviendas del territorio mexicano, sus características, los servicios a los que tienen acceso y materiales en la vivienda. Toda esta información se recaba en un momento determina-

do, reflejando así una fotografía del país en un momento específico del tiempo.

Los datos censales, permiten un desglose en distintas unidades geográficas. Desde los agregados nacionales hasta las localidades y manzanas.

Algunas características de los censos de población y vivienda de acuerdo con el INEGI [61] son:

- La enumeración individual. Se entrevista a cada persona y se enumera a cada vivienda con sus características de manera individual.
- La universalidad. Un censo debe llegar a todo el territorio nacional. A diferencia de las encuestas que visitan una muestra, el censo visita todo México para contar a todas las personas que residan en el país.
- La simultaneidad. A todos se les debe entrevistar en el mismo periodo de tiempo.
- La periodicidad definida. Con el objetivo de contar con información comparable en el tiempo, los censos deben levantarse en intervalos regulares y tener una secuencia fija. La CEPAL, recomienda que se levante un censo nacional por lo menos cada diez años. [62]

Para poder ubicar geográficamente a los habitantes del país se utiliza el Marco Geoestadístico que permite relacionar la información estadística con los lugares geográficos correspondientes. De acuerdo con INEGI [61], se divide al territorio nacional por medio de límites geoestadísticos creando tres niveles de desagregación: Área Geoestadística Estatal (AGEE), Área Geoestadística Municipal (AGEM) y Área Geoestadística Básica (AGEB) la cuál puede

Vivienda	Población	Migración internacional
<ul style="list-style-type: none"> • Materiales en paredes • Materiales en techos • Lugar donde cocinan • Combustible • Fogón con chimenea • Focos • Uso del sanitario • Separación y reutilización • Eliminación de basura • Tenencia • Escritura o título • Adquisición • Financiamiento • Deuda • Nombre y número de la persona dueña o propietaria 	<ul style="list-style-type: none"> • Municipio de asistencia escolar • Entidad o país de asistencia escolar • Tiempo de traslado a la escuela • Modo o medio de traslado a la escuela • Nombre de la carrera • Identificación de la pareja • Nombre de la ocupación • Tareas • Posición en el trabajo • Prestaciones laborales • Ingresos por trabajo • Horas trabajadas • Negocio, empresa o lugar de trabajo • Actividad del negocio, empresa o lugar de trabajo • Municipio de trabajo • Entidad o país de trabajo • Tiempo de traslado al trabajo • Modo o medio de traslado al trabajo • Hijas(os) sobrevivientes • Fecha de nacimiento • Supervivencia • Identificación de la última hija o hijo • Edad al morir 	<ul style="list-style-type: none"> • Condición de migración internacional • Número de personas • Personas migrantes • Lista de personas • Condición de residencia • Sexo • Edad al migrar • Fecha de emigración • Causa de la emigración • Lugar de origen • País de destino • País de residencia • Fecha de retorno • Causa del retorno • Condición de residencia actual
		Ingresos monetarios de fuentes distintas al trabajo
		<ul style="list-style-type: none"> • Otros ingresos
		Alimentación
		<ul style="list-style-type: none"> • Acceso a los alimentos en la vivienda • Alimentación de los adultos • Ingesta de alimentos de los adultos

Figura 3: Variables incluidas en el cuestionario ampliado. Fuente, CENSOPV2020 [61]

ser urbana o rural. Para establecer los límites geostadísticos, de acuerdo con el INEGI [61] se utilizan líneas divisorias convencionales. Estas líneas se trazan sobre rasgos naturales (ríos, arroyos, barrancas, cerros, litorales, etcétera) o culturales permanentes e identificables en el terreno (calles, vías de comunicación terrestre, líneas de conducción, cercas, ductos, límite de viviendas, linderos, etcétera) [61]. Las variables que provee el cuestionario ampliado del Censo de Población y Vivienda nos permite ligar la información con los indicadores del dominio a nivel vivienda.

3.1.2. Datos raster

Para esta investigación analizamos el uso de imágenes de satélite como una alternativa a los costosos censos y encuestas. La mayoría de los estudios sobre el mapeo de áreas urbanas desfavorecidas funcionan con imágenes comerciales de muy alta resolución (1 metro o menos de resolución). La resolución espacial se refiere al tamaño de un píxel en el suelo. Un píxel es el punto más pequeño que forma una imagen de satélite óptico y básicamente determina qué tan detallada es una imagen.

Los datos de los satélites Landsat de la NASA, por ejemplo, tienen una resolución de 30 metros, lo que significa que cada píxel representa un área de 30 por 30 metros en el territorio. Se considera una imagen de resolución media, aquella que puede cubrir toda el área de una ciudad por sí sola, pero el nivel de detalle no es lo suficientemente fino como para distinguir objetos individuales como casas o automóviles.

Esta subdivisión en baja, media y alta es provisional, ya que la tecnología de imágenes avanza todo el tiempo. Lo que se consideraba alta resolución en los años 80, por ejemplo, los datos de los satélites Landsat con 60 metros por píxel en ese entonces, se ha vuelto de resolución baja en los estándares actuales. La mejor resolución hasta ahora, es de 30 centímetros, proporcionada por satélites comerciales de muy alta resolución.

De acuerdo con Kuffer [63], quien analizó todos los estudios publicados en Scopus ($N = 28$) sobre áreas desfavorecidas vistas desde el espacio de los

últimos tres años, más del 50 por ciento trabaja con imágenes ópticas de muy alta resolución, seguidas por el 30 por ciento de los artículos que utiliza imágenes ópticas de alta resolución (por ejemplo del satélite Sentinel-2 del programa Copérnico de la Agencia Espacial Europea que tienen una resolución de 10m). Por último, menos del 20 por ciento utiliza imágenes de Google Earth con imágenes de SAR (de muy alta y alta resolución). El predominio de las imágenes de alta resolución también está contribuyendo al dilema, en el sentido de que la mayoría de los estudios desarrollan y prueban métodos en áreas muy pequeñas (generalmente pocos km²); lo cual es insuficiente para reflexionar sobre el potencial de integrar observaciones de la Tierra y datos estadísticos, ambos típicamente disponibles para cobertura de áreas grandes [63].

En la comunidad de investigadores que trabajan con observaciones de la Tierra, existe un entendimiento de que las imágenes de muy alta resolución son necesarias para revelar patrones urbanos complejos e información socioeconómica. Sin embargo, Abascal [31] al comparar las precisiones de clasificación y los costos de imagen, los agrupa en tres clases (donde 'Alto' significa más de 10 € / km², 'Moderado' significa 10 € / km² o menos, e imágenes 'Gratis' sin costo). No existe una fuerte relación entre alta precisión y alto costo de imágenes. Las variaciones dentro de las tres clases (del costo de la imagen) pueden explicarse por los diferentes algoritmos de aprendizaje automático utilizados y la diferente complejidad de las ciudades.

En general, las imágenes gratuitas (principalmente las de Sentinel y Goo-

gle Earth) logran una precisión menor, pero alcanzan más del 80 por ciento en promedio. Por tanto, son una alternativa prometedora para evitar los altos costos de los datos de imágenes de muy alta resolución.

Varios estudios [50] [64] han confirmado que las imágenes con muy alta resolución no tienen la resolución óptima para mapear barrios marginados, ya que se ven afectadas por detalles de nivel de objetivo demasiado altos, lo que agrega ruido a la clasificación. Por tanto, una resolución óptima depende de la morfología urbana y normalmente varía entre 2 y 5 metros en zonas marginadas. Sin embargo, esta conclusión difiere cuando se trata de mapear objetos (por ejemplo, edificios).

Las imágenes de muy alta resolución son fundamentales para mapear los objetos de construcción que se requieren para muchos propósitos de análisis morfológico urbano. Una resolución inferior a 0.5 metros es necesaria para mapear edificios en áreas marginadas. Sin embargo, en zonas urbanas muy densamente construidas, como las áreas desfavorecidas en Nairobi, Kenia [65] incluso 0.3 metros no son necesariamente suficientes para capturar todos los techos como objetos separados. Los techos en dichas áreas están formando manchas más grandes que, incluso para los intérpretes visuales, son difíciles o incluso imposibles de separar.

Para el caso de estudio planteado en esta investigación se utiliza la Geomediana Landsat. La Geomediana Landsat es un producto de acceso gratuito con la cual se presenta por año, una imagen sintética del territorio nacional

que representa un promedio estadístico de las imágenes que se tuvieron del territorio en ese año. Esta imagen tiene la ventaja de no tener nubes y preservar las relaciones multispectrales [66].

3.2. La Retícula

La propuesta de esta investigación se basa en la integración de información estadística y geoespacial, específicamente la integración de datos de percepción remota con datos censales. Para realizar dicha integración es importante que los datos sean interoperables y para ello construiremos una retícula de población que permite traducir los datos censales que se encuentran organizados de acuerdo con las demarcaciones territoriales político administrativas a datos a nivel pixel o celda para lograr la integración con las imágenes de satélite. Más allá de una tendencia para la representación geoespacial de datos relacionados con fenómenos sociales, biológicos o físicos, una retícula es una herramienta para el estudio y análisis de esos fenómenos —en su distribución, en su comportamiento, en sus combinaciones, en su localización, y en sus relaciones de conexión y causalidad.

Tal reticulado consiste en una subdivisión del territorio en unidades de área denominadas celdas, que tienen en lo general forma cuadrada, y cuyas dimensiones son elegidas en función del fenómeno sujeto a estudio; ya que, por ejemplo, no es lo mismo un estudio de deforestación regional, que un análisis de epidemia nacional, o una investigación de mercado local.

Dichas retículas también abonan a la generación de información con mayor nivel de detalle. La escasez de datos en países de ingresos medios y bajos es una preocupación para los gobiernos, la sociedad y los organismos internacionales. Por ello, han surgido esfuerzos para generar información a nivel global basada en las nuevas tecnologías y métodos que coadyuvan a la generación de información para conocer las condiciones socioeconómicas de los habitantes. Este es el caso del esfuerzo reciente basado en el uso de imágenes de satélite, el uso creciente del poder computacional, ciencia ciudadana como OpenStreet Map, entre otros. Anteriormente, para generar mapas de población se requería conectar una tabla estadística a un vector de coordenadas. En 1990, surge un nuevo enfoque para mapear población que consiste en generar retículas estandarizadas de celdas o píxeles. [67] [68].

Estos grids o retículas han sido impulsadas por la comunidad científica para crear bases de datos globales georeferenciadas que facilitan la integración con otras bases de datos reticuladas como son los productos de percepción remota. Han probado ser útiles para analizar fenómenos que no respetan fronteras político-administrativas. Los datos socioeconómicos, generalmente tienen una estructura basada en los distintos niveles geográficos: estatal, municipal, y por localidad. Sin embargo, los fenómenos que suceden en el territorio no siempre respetan dichas fronteras. Tal es el caso de una inundación, deforestación, crecimiento poblacional. Por ello, el análisis de los fenómenos pasados en celdas en lugar de municipios o localidades puede resultar de gran utilidad.

La proliferación de distintas fuentes de información ha creado iniciativas

como, por ejemplo, Pop Grid [69], que se puede consultar en línea y permite analizar los estimados de población por área. Las aplicaciones son ilimitadas. Pop Grid cuenta con 3,300 citas en SEDAC (Socioeconomic data and applications center de la NASA). De acuerdo con el Dr. Robert Chen, el director del SEDAC, de 2007 a 2016, PopGrid se utilizó en sesenta y tres distintos campos de conocimiento, entre ellos: agricultura, medicina, geografía, teledetección, meteorología, entre otras. El proceso de mapeo de población en una retícula requiere asignar espacialmente los datos a través de modelaje estadístico. Las técnicas de ponderación de área redistribuyen equitativamente los datos en las celdas [70] [71]. Los datos de ingreso o “inputs” pueden ser datos censales, o de registros administrativos mientras que las celdas objetivo, usualmente representan una superficie menor a la proveniente de la fuente. Este método asigna de manera uniforme la población de la fuente a las celdas. Este es un supuesto relevante ya que, generalmente, la población no se distribuye uniformemente. Sin embargo, la distribución proporcional es computacionalmente eficiente para informar políticas públicas que no requieren una muy alta resolución espacial [72], o para realizar análisis de correlación en los que se plantean problemas de endogeneidad (por ejemplo, Cohen y Small, 1998 [73]). El proyecto Gridded Population of the World (GPWv4) [74] es un ejemplo que utiliza el método de ponderación de áreas.

Otra técnica utilizada para generar retículas de población es la que utiliza datos auxiliares como pueden ser datos vectoriales de carreteras y caminos, cartas de uso de suelo y vegetación, datos topográficos, cuerpos de agua, entre otros. Leyk, et.al, [75] realiza un análisis de los tipos de variables au-

xiliares empleadas para generar retículas de población. A dichos métodos se les llama mapeo dasimétrico. El mapa dasimétrico utiliza información complementaria para dividir las áreas estadísticas en áreas homogéneas. Existen distintos métodos para asignar la población a cada celda desde los métodos binarios dasimétricos tradicionales, hasta los que utilizan pesos específicos para la asignación.

Recientemente, han aumentado los enfoques híbridos que combinan los métodos más tradicionales de mapeo dasimétrico con análisis estadístico. Estos enfoques a menudo se basan en técnicas de aprendizaje de máquina que permiten realizar una estimación robusta de los pesos de la población y, en un segundo paso, realizar un proceso de redistribución dasimétrica [76] [77].

La figura 4 ejemplifica visualmente los distintos métodos para distribuir la población en una retícula de población. En a) se observa la distribución proporcional en área que no utiliza variables auxiliares y distribuye proporcionalmente. En b) mapeo dasimétrico binario utiliza variables auxiliares y asigna pesos de manera binaria. En c) mapeo dasimétrico con pesos asignados. Y en d) mapeo dasimétrico con modelaje estadístico. Los sombreados con tonos de gris denotan distintas variables auxiliares que informan la distribución.

Existen distintas iniciativas globales que han generado retículas de población como se muestra en la figura 5. Las retículas aquí mencionadas utilizan distintos modelos para su generación. Los modelos de arriba hacia abajo (top

(a) Distribución Proporcional en área
Unidad A, n=54 Unidad B, n=36

6	6	6	4	4	4
6	6	6	4	4	4
6	6	6	4	4	4

(b) Mapeo dasimétrico binario
Unidad A, n=54 Unidad B, n=36

0	0	0	0	0	0
0	13.5	13.5	18	0	0
0	13.5	13.5	18	0	0

(c) Mapeo dasimétrico con pesos asignados
Unidad A, n=54 Unidad B, n=36

4	4	4	3	3	3
4	8.5	8.5	7.5	3	3
4	8.5	8.5	7.5	3	3

(d) Mapeo dasimétrico con modelaje estadístico
Unidad A, n=54 Unidad B, n=36

5	5.5	5.5	3	2	2
5	7	7	9	3	2
5	7	7	9	3	3

Figura 4: Métodos para desagregar población a nivel celda [75].

-down) están basados en datos de enumeraciones censales. Generalmente, los modelos de arriba hacia abajo son dasimétricos, es decir, la desagregación de la población se realiza de unidades de enumeración hacia las celdas a través de datos auxiliares [78].

Los modelos de abajo hacia arriba (bottom-up), utilizan microdatos censales al menor nivel de agregación posible que generalmente es aquel que no revela la identidad del informante para estimar la población en cada celda directamente [74]. Thomson [79], analiza los métodos empleados por nueve retículas disponibles [ver Figura 4.3] y encuentra que todas subestiman la población asentada en barrios marginados y asentamientos informales, especialmente en zonas con mayor población.

Las retículas ofrecen métodos relativamente sencillos, transparentes y fáciles de usar. Sin embargo, se recomienda usarlas con cautela debido a la subestimación de población en zonas altamente pobladas. Algunas recomendaciones de Thomson [79] para mejorar la estimación a nivel celda son: categorizar zonas habitadas como residenciales o no residenciales; clasificar zonas desfavorecidas; utilizar variables como caminos o carreteras y datos de elevación. Es importante recordar que las retículas son construidas a nivel global, por lo que si se busca realizar un análisis local con mayor nivel de granularidad, se necesita generar una base con mayor nivel de detalle.

Iniciativa	Productor	Año	Resolución	Cobertura	Método
Gridded Population of the World	CIESIN, Universidad de Columbia	2015, 2020	30 arc sec (~1km ²)	Residencial	Asignación uniforme a celdas (areal weighting on edge cells)
Global Human Settlement	Comisión Europea, Joint Research Centre (JRC)	2015	9 arc sec (~250m ²)	Residencial	Dasimétrico binario, asignación proporcional
High Resolution Settlement Layer	Facebook Connectivity Lab y CIESIN	2018	1 arc sec (~30m ²)	Residencial	Dasimétrico binario, proporcional a hogares
WorldPop-Unconstrained	WorldPop, Univ. de Southampton	2015, 2018	3 arc sec (~100m ²)	Residencial	Dasimétrico con bosque aleatorio
LandScan	Oak Ridge National Laboratory	2015, 2018	30 arc sec (~1km ²)	Ambiente (24-hour average)	Dasimétrico multivariable con 4 covariable
WPE	ESRI	2016	162m	Residencial	Algoritmo dasimétrico
World-Pop-Constrained	WorldPop, Univ. de Southampton	2020	3 arc sec (~100m ²)	Residencial	Modelo Bosque Aleatorio con 24 covariables y redistribución dasimétrica
GRID3	CIESIN, WorldPop, Flowminder, UNFPA, DFID	2016	3 arc sec (~100m ²)	Residencial	Modelo Bayesiano Jerárquico

Figura 5: Retículas de Población Globales [79]

3.3. Selección de estudio de caso

Para poder seleccionar el caso de estudio analizo el proceso de geoanalítica del riesgo por desastre establecida por CENAPRED [80]. La metodología empleada consta de los siguientes pasos:

1. Identificación de los fenómenos naturales y antrópicos que pueden afectar una zona en estudio —geológicos, hidrometeorológicos, químicos, sanitario-ambientales y socio-organizativos.
2. Determinación del peligro asociado a los fenómenos identificados.
3. Identificación de los sistemas expuestos y su vulnerabilidad.
4. Evaluación de los niveles de riesgo asociado a cada tipo de fenómeno, tanto natural como antropogénico.
5. Integración de la información sobre los fenómenos naturales y antropogénicos, peligro, vulnerabilidad y riesgo, considerando los recursos técnicos y humanos. [80]

De acuerdo a lo anterior, tomo la base de datos geospaciales correspondientes a los desastres hidrometeorológicos publicados en el Sistema Nacional de Información del Agua (SINA) [81] de la Comisión Nacional del Agua (CONAGUA), así como la base de datos del Sistema de Consulta de Declaratoria del Centro Nacional de Prevención de Desastres (CENAPRED) [80]. La fuente de la que se obtiene la información para este sistema es de las declaratorias de desastre del Diario Oficial de la Federación (DOF) [82].

Analizando seis variables obtenidas de la declaratoria de desastre para 2020: Desastres hidrometeorológicos (presencia de sequías severas, inundaciones, ciclones tropicales, lluvias severas) y Desastres Geológicos (deslave/deslizamiento y sismos), su ubicación en el territorio mexicano y la frecuencia con la que suceden, se observa que los desastres ocurren de manera muy focalizada en algunas regiones del país. Por ejemplo, los desastres hidrometeorológicos como ciclones tropicales, lluvias severas e inundaciones afectaron mayormente al sureste del país; los desastres geológicos como los deslaves/deslizamientos y sismos afectaron de manera focalizada a los estados de Chiapas y Oaxaca; y las sequías severas afectaron principalmente al noroeste del país. Por otro lado, la presencia de incendios forestales se dio en casi todos los estados del país. [82]

De los distintos tipos de desastres analizados, las imágenes satelitales ofrecen una relevante oportunidad para estudiar los desastres hidrometeorológicos, ya que existen métodos que detectan agua superficial y permiten identificar los sistemas expuestos y su vulnerabilidad. Por lo anterior, si observamos las afectaciones por inundaciones, ciclones tropicales o lluvias severas, podemos observar que es en el Sureste del país, donde se presenta un mayor número de incidencias. Con la integración de información estadística y las observaciones de la Tierra, es posible evaluar la vulnerabilidad de los sistemas expuestos, a través de métodos que cuantifiquen los daños y pérdidas probables por efectos del fenómeno hidrometeorológico, sobre los sistemas expuestos [82].

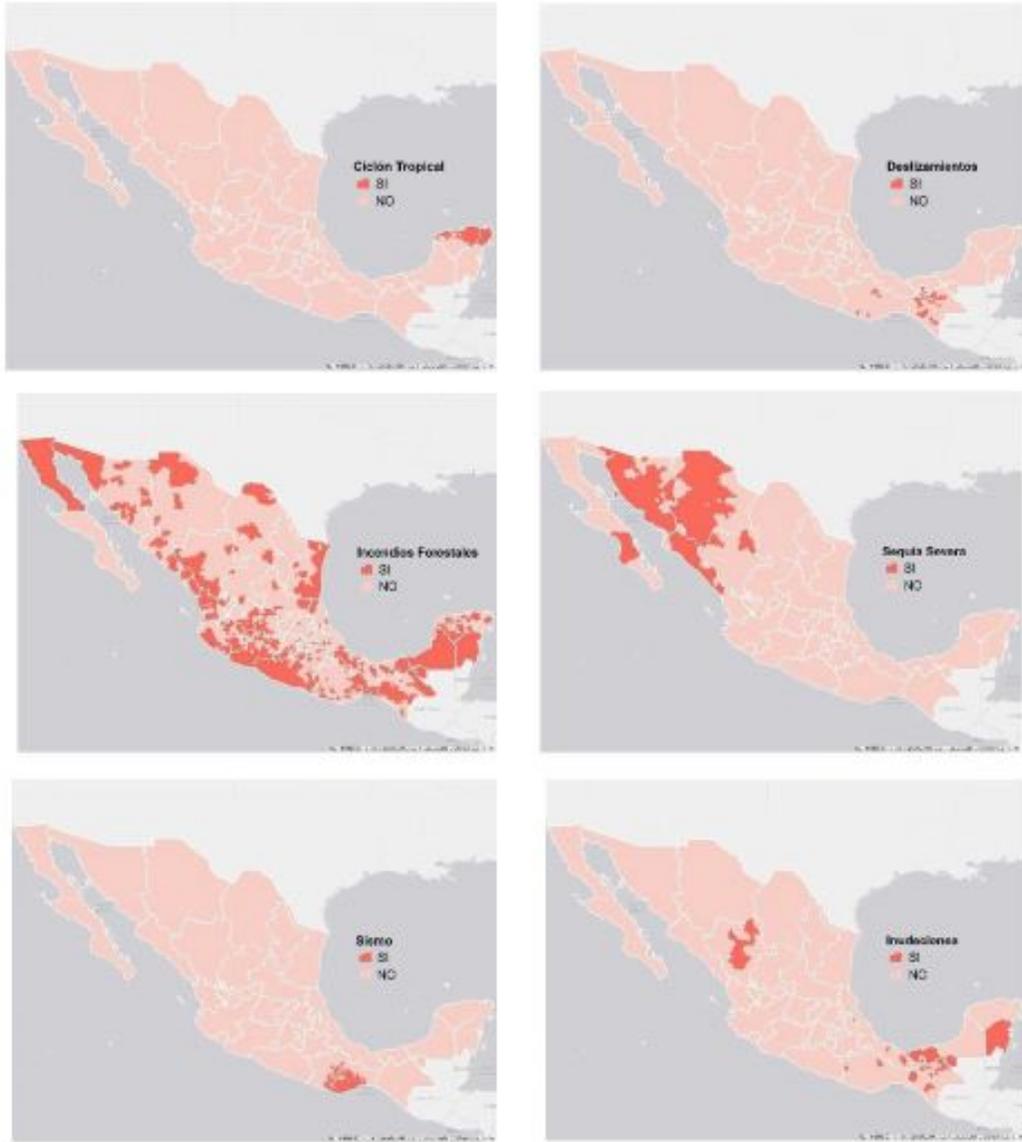


Figura 6: Declaratorias de Desastres 2020 [80]

Las inundaciones en la ciudad capital del estado de Tabasco suceden debido a que la planicie tabasqueña es atravesada por los dos ríos más caudalosos de México; el río Usumacinta y el río Grijalva. A causa de fuertes lluvias son sobrepasados y causan afectaciones en el territorio del estado. La Cuenca Hidrográfica Río Grijalva-Villahermosa se toma como área de estudio debido a que una cuenca es una unidad idónea para el estudio de los efectos de fenómenos hidrometeorológicos, dado que es, por definición, una región que es drenada hacia el mar por un solo sistema de drenaje natural. En la imagen, se muestra en color azul la región cubierta por la Cuenca. En color rojo la delimitación del Municipio Centro del Estado de Tabasco y de fondo la Geomediana 2020 (el promedio estadístico de la imagen satelital Landsat de 2020), con fines de alineación temporal con las inundaciones en dicho año.

Una vez seleccionada el área de interés, uniendo la información estadística relativa a la vulnerabilidad a nivel hogar, con la vulnerabilidad por presencia de riesgo de desastre hidrometeorológico, se mostrarán en la región de la cuenca, las zonas desfavorecidas.

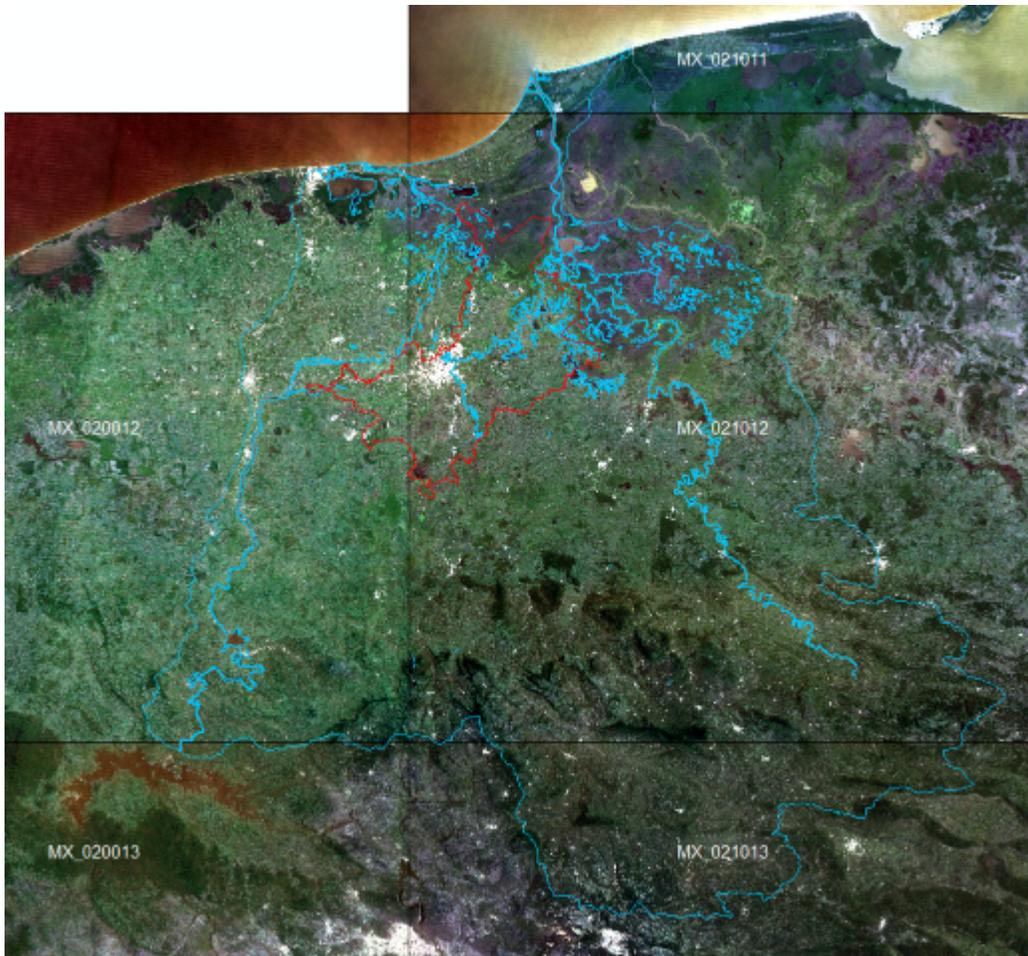


Figura 7: Cuenca Hidrográfica Río Grijalva-Villahermosa

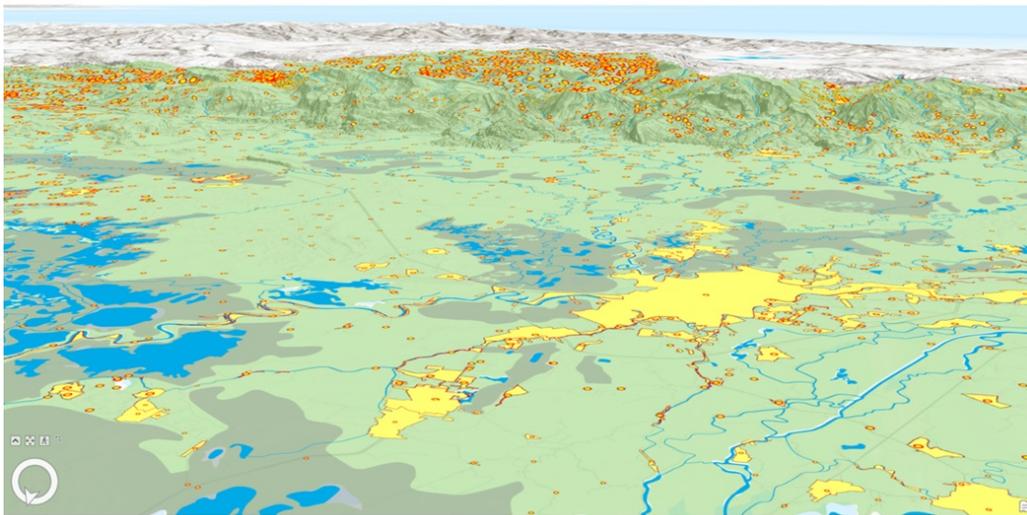


Figura 8: Localidades amanzanadas (amarillo, cuerpos de agua (azul), Terreno sujeto a inundación (gris).
Fuente: Elaboración propia con datos de Mapa Digital de México.

3.4. Observaciones de agua desde el espacio

Una vez identificadas las técnicas de percepción remota para mapear áreas desfavorecidas, analizo las técnicas utilizadas con sensores ópticos para mapear las zonas inundadas o propicias a ser inundadas en caso de un desastre.

Para localizar las áreas desfavorecidas ante desastres, existe una amplia literatura que analiza el potencial de las observaciones de la Tierra para reducir el riesgo de desastres. Específicamente en el caso de inundaciones con la llegada del radar de apertura sintética (SAR), la disponibilidad de imágenes satelitales de inundaciones ha proliferado y se sigue incrementando particularmente con el lanzamiento de satélites SAR de muy alta resolución (por ejemplo, TerraSAR-X, DLR (Del Centro Aeroespacial Alemán), Radarsat-2, CSA (Agencia Espacial Canadiense)) y también constelaciones (COSMO-SkyMed, ASI (Agencia Espacial Italiana), Sentinel-1, ESA (Agencia Espacial Europea) [83].

Para el mapeo de inundaciones, SAR tiene la ventaja de penetrar la capa de nubes y no se ve afectado en gran medida por condiciones meteorológicas adversas, las cuales a menudo persisten durante las inundaciones de gran impacto [84]. Lo anterior ha llevado a contar con una mayor confiabilidad en el mapeo de inundaciones y progreso acelerado en la predicción de inundaciones.

Recientemente, se han publicado varios artículos que revisan los méto-

dos comúnmente utilizados para mapear inundaciones de sensores satelitales SAR y ópticos [83]. Schumann [83] proporciona una lista de los sensores que se usan típicamente en el mapeo de inundaciones y analiza las principales ventajas de ambos tipos de sensores (SAR y óptico). Algunos sensores como Landsat tienen la ventaja de contar con una serie histórica que permite analizar cambios a lo largo del tiempo, otros tienen la ventaja de proveer más de una imagen diaria como Modis/Terra Aqua, y otros, proveen de avances tecnológicos en la resolución o la capacidad de mapear inundaciones debajo de la vegetación.

Las inundaciones se encuentran entre los desastres de mayor costo anual en pérdidas aseguradas y no aseguradas [83]. Dado que los eventos de alto impacto a menudo cubren escalas espaciales que van más allá de las operaciones tradicionales de monitoreo regional, la teledetección, en particular desde satélites, presenta un enfoque atractivo.

Los métodos para detectar inundaciones a través de imágenes ópticas están bien establecidos [85]. La teledetección permite obtener información de objetos situados en la superficie terrestre. A través del sensor de los satélites se recibe la radiación de los objetos [86].

Varios métodos aprovechan las ventajas espectrales de los índices que explotan la diferencia de reflectancia entre lo visible y las partes infrarrojas del espectro como el Índice de vegetación (NDVI, por sus siglas en inglés) [87] y la diferencia normalizada Índice de Agua (NDWI, por sus siglas en inglés) [88].

Métodos de clasificación estadística aplicados a los datos de Landsat han demostrado resultados muy precisos, incluidos métodos de máxima verosimilitud supervisados [89] y árboles de decisión [90]. En este trabajo, utilizo el producto llamado Observaciones de Agua desde el Espacio (WOFS, por sus siglas en inglés) [91]. Las razones para emplear este producto son:

- permite analizar agua en la superficie a grandes escalas, es decir, se puede calcular para todo el territorio nacional y no es necesario reducir el análisis a una región,
- se utilizan los sensores ópticos de los satélites Landsat los cuales están disponibles gratuitamente y se cuenta con un acervo histórico de más de 30 años, lo que permite hacer comparaciones en el tiempo,
- los resultados obtenidos del árbol de clasificación permiten tener intervalos de confianza a nivel pixel para realizar una valoración sobre la precisión de los resultados.

El producto Wofs, cuenta con tres bandas. En la banda 1 *Wofs* se tienen los datos a nivel pixel de los valores calculados de agua, expresados en porcentaje (de cero a 100 por ciento). En la banda 2 *WET*, se encuentran los valores de las lecturas que corresponden a agua. Por último en la banda 3 *TOTAL* se escribe el dato de la cantidad de observaciones disponible para ese pixel. De manera general, el cálculo de Wofs se expresa de la siguiente manera:

$$WOFS = (WET/TOTAL) * 100 \quad (3)$$

Un ejemplo ilustrativo sería si, para un lago, tenemos un total de 24 observaciones (escenas de imágenes Landsat) en la banda 3, de esas 24 escenas observamos un total de 24 con presencia de agua (al ser un lago esperaríamos que tenga agua todo el año). Al aplicar la fórmula antes mencionada, la banda 1 Wofs expresa un 100 por ciento de presencia de agua. En la siguiente sección utilizaremos el Wofs como uno de los indicadores más de vulnerabilidad.

3.4.1. El código

Uno de los objetivos de esta investigación es el uso de datos no tradicionales como son las imágenes de satélite en las ciencias económicas. El uso de estos datos, también llamados *Big Data* requiere del uso de paquetes computacionales que permitan implementar técnicas de inteligencia artificial para el análisis.

Es por ello que en esta investigación se utilizan herramientas de software abierto para analizar las imágenes de satélite. En específico, utilizo la herramienta Google Sandbox [92]. Dicha interfaz de programación abierta y gratuita que conecta a los usuarios con los conjuntos de datos de Google Earth Engine.

La herramienta utiliza código abierto y permite ejecutar aplicaciones del software Python, el lenguaje de programación más común entre los científicos de datos. Esta herramienta, a través del entorno de cuadernos de Colab, permite procesar los datos satelitales sin necesidad de descargarlos lo cual puede limitar su uso en computadoras personales.

El algoritmo utilizado en esta investigación toma el código abierto puesto a disposición de usuarios a través del Open Earth Alliance [92] y es modificado para generar los datos para la zona de estudio. El código empleado para generar el algoritmo de agua desde el espacio puede ser consultado en el Apéndice C así como en la siguiente [liga](https://colab.research.google.com/drive/1Em121ELqqlEZ4fTrNDr0peV7K-a?usp=sharing): <https://colab.research.google.com/drive/1Em121ELqqlEZ4fTrNDr0peV7K-a?usp=sharing>.

3.5. Clasificación

La aproximación interdisciplinaria es el centro de la presente investigación. He analizado la integración de la información estadística que es generada y analizada a través de la ciencia de probabilidad y muestreo. Por otra parte, he reflexionado sobre el uso de imágenes raster que incluyen desde la geodesia para conocer la exactitud posicional, hasta la teledetección y el análisis geoespacial a través de sistemas de información geográfico. En esta sección, introducimos la ciencia computacional como una herramienta fundamental para el análisis y el mapeo de las zonas desfavorecidas. Barabasi 2012 [93], identifica que en la era actual existe una explosión de datos. Desde las redes sociales hasta la biología celular, la información generada en nuestros días es una oportunidad única para documentar las entrañas de los sistemas complejos. Al tener datos masivos, podemos conocer con mayor detalle el comportamiento de los agentes y las interacciones entre ellos. Para poder hacer uso eficiente de los datos masivos, necesitamos a las ciencias computacionales.

El objetivo de esta sección es especificar el método para clasificar áreas desfavorecidas o áreas no desfavorecidas de acuerdo con los dominios de estudio definidos en la sección 3.1.2. Se presenta un algoritmo que recibe como un input una celda de la retícula y regresa como output la clasificación.

La metodología se conforma por los siguientes pasos:

1. Recolección, limpieza e integración de bases de datos
2. Selección de variables o features relevantes
3. Construcción de las categorías de vulnerabilidad
4. Construcción de un modelo predictivo a nivel celda
5. Clasificación de la imagen
6. Validación
7. Mapeo de áreas desfavorecidas

En la Figura 9 se muestra el diagrama de flujo, que define el inicio y el fin del método empleado para generar información sobre las áreas desfavorecidas de manera general. Cada figura geométrica dentro del diagrama de flujo representa un proceso de la metodología a seguir que deberá ser desarrollado con detalle. El círculo marca el inicio y el fin del proceso metodológico. El paralelogramo indica que la computadora debe de obtener datos de entrada

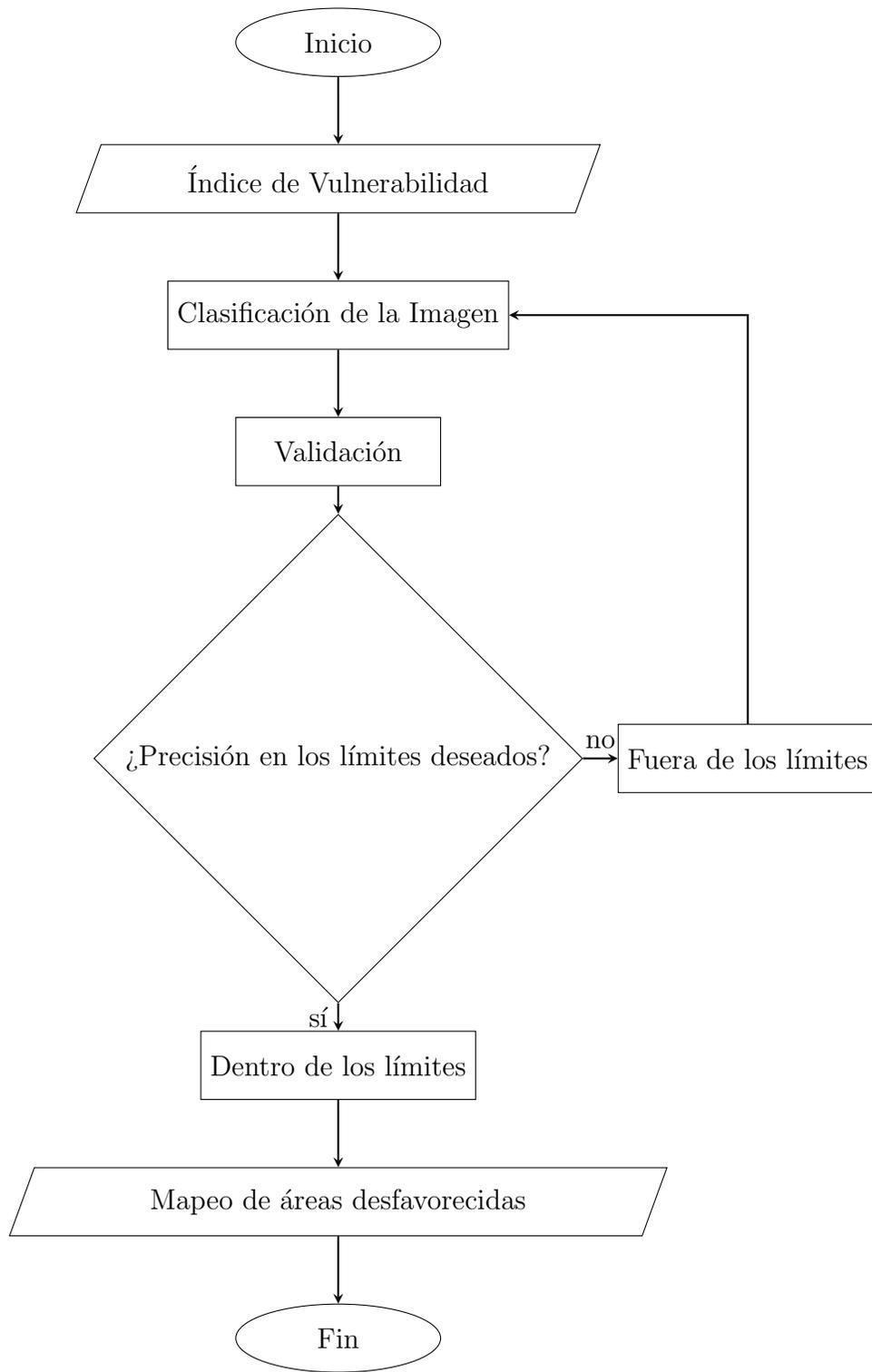


Figura 9: Diagrama de Flujo

o salida. El rectángulo denota un proceso de operaciones aritméticas o manipulación de datos. El rombo señala que una decisión debe tomarse. Y las flechas, guían el flujo del proceso.

El índice que empleo para medir vulnerabilidad es el índice de rezago social [94] que especifica, de acuerdo con la literatura sobre índices de vulnerabilidad física [95] [96] [97] [98] [99] y vulnerabilidad social [100] [101] [102] las variables a los pasos a seguir son:

1. Definición de las variables a considerar con base en la teoría
2. Recolección, limpieza e integración de bases de datos
3. Cómputo estadístico
4. Selección de las variables dominantes para cada dominio
5. Ponderación del índice de acuerdo a las variables dominantes para cada componente
6. Combinar las variables a través de la suma ponderada
7. Normalización del índice
8. Clasificación de categorías de vulnerabilidad alta, media, baja
9. Mapeo a nivel celda del índice de vulnerabilidad

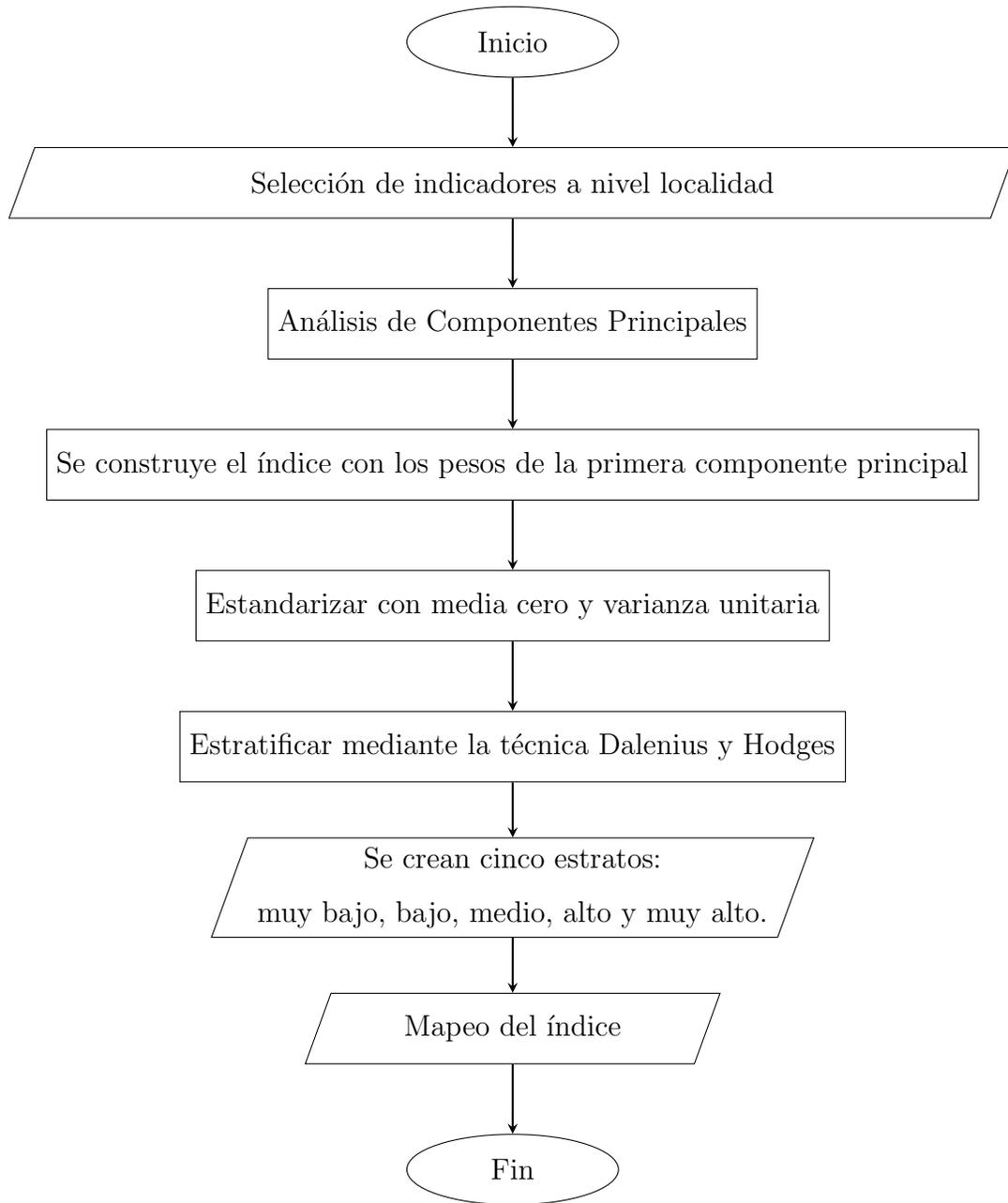


Figura 10: Índice de Rezago Social

3.6. Método del trabajo de campo

En esta sección establezco el método de trabajo de campo que permita comprobar y ratificar la información obtenida. El objetivo es conocer a profundidad, a través del conocimiento de expertos en campo, las respuestas a las preguntas de investigación y la teoría empírica que orienta las respuestas. A través de técnicas de medición cualitativas busco relacionar las teorías analizadas en capítulos anteriores con un análisis empírico realizado a través de un proceso de entrevistas en campo.

La entrevista es un procedimiento para el levantamiento de información en una población. Las entrevistas son técnicas de investigación social idóneas para obtener información valiosa. De acuerdo con Aravena y Cols (2006), se define la entrevista como una estrategia de investigación basada en las declaraciones verbales de una población concreta, a la que se realiza una consulta para conocer determinadas circunstancias políticas, sociales o económicas, o el estado de opinión sobre un tema en particular [103].

Para elaborar la estrategia del proceso de entrevista se realizó una indagación exploratoria sobre lo que se quiere lograr en este trabajo de campo que es validar los resultados del análisis cuantitativo de esta investigación. Se prepararon las preguntas y posibles respuestas en función de las variables a estudiar, en este caso la relación entre inundaciones y el rezago social para el caso de la Cuenca Río Grijalva-Villahermosa. Se analizó el tiempo, espacio y población a encuestar para definir el instrumento a utilizar, su formato y aplicación que se definen en los siguientes pasos. Siguiendo la teoría sobre la

metodología de encuestas en campo, todo esto fue consultado con un comité de expertos, en este caso se trató del comité tutor de esta investigación [103].

3.6.1. Criterios de selección de los entrevistados

La encuesta debe realizarse a una muestra de la población para obtener un resultado válido. Por lo anterior, en el primer semestre de 2022, entrevisté a un grupo de personas involucradas en la atención de la marginación y vulnerabilidad ante inundaciones en Tabasco. Es por esto que los entrevistados provienen de distintos ámbitos como son el poder ejecutivo y legislativo. Del mismo modo, entrevisté a personas del ámbito académico y la sociedad civil. Además, integro actores que se han desempeñado en administraciones públicas actuales y anteriores buscando un balance de sectores, género y tiempo en los entrevistados.

3.6.2. Cuestionario

Para el diseño del cuestionario se recurrió a la operacionalización de variables de la investigación a partir de las preguntas y las hipótesis planteadas en esta investigación, con una validación previa del comité experto. El formato del cuestionario puede ser de preguntas abiertas, cerradas o ambas. En este caso, se formuló a través de preguntas abiertas. Formulo preguntas sencillas, iniciando por las más generales a las más particulares.



Figura 11: Perfil de los entrevistados

La teoría de la investigación social cualitativa indica que el guión de entrevistas no debe ser tan extenso que haga redundante la información que obtenemos. Además, el lenguaje debe ser sencillo y favorecer el proceso comunicativo [104]. En el apéndice A, presento la guía utilizada para la entrevista con las preguntas.

La pandemia por COVID-19, ha propiciado el trabajo remoto y lo ha normalizado en nuestras vidas. Cada vez más, las entrevistas cara a cara se llevan a cabo a través de medios virtuales. Es por lo que se ha seleccionado llevar a cabo el trabajo de campo vía virtual a través de entrevistas cara a cara realizadas a través de una plataforma de videollamada. Así, la administración del cuestionario se realizó virtualmente.

4. Resultados

La carta internacional activada en 2020 para la Cuenca Río Grijalva-Villahermosa, está compuesta por agencias y operadores de sistemas espaciales de todo el mundo que trabajan juntos para proporcionar imágenes satelitales con fines de seguimiento a desastres [105]. El 6 de noviembre de 2020, fue activada la carta debido al paso del Huracán Eta que afectó los estados de Tabasco y Chiapas.

El objetivo de esta sección es integrar información estadística y raster para el análisis de vulnerabilidad ante inundaciones en la Cuenca Río Grijalva-Villahermosa. Para ello, utilizo las imágenes satelitales para mapear zonas inundadas y la información estadística que nos permiten identificar la población afectada. En concreto, empleo como insumos, las imágenes Landsat y la retícula de población con los datos del Censo de Población y Vivienda 2020.

La retícula es generada en concordancia con la información ráster a utilizar. Por ello, uso una retícula de 30 metros por 30 metros que coincide con las imágenes de los satélites Landsat. Las variables seleccionadas son la población y las viviendas totales del Censo de Población y Vivienda 2020. Para asignar las variables a cada celda de la retícula aplico el método definido en la sección 4 que consiste en una distribución proporcional en área y distribuye equitativamente los datos en las celdas.

En cuanto a la información geoespacial, utilizo la variable Wofs [91] pro-



Figura 12: Macuspana, Tabasco el 9 de noviembre de 2020. Fuente: atalayar.com

veniente del algoritmo citado en la sección 4.4 como un predictor de área desfavorecida. El algoritmo Wofs (Water Observation from Space) nos indica el porcentaje de tiempo en el que el píxel tuvo agua. En la Figura 13, se puede observar el Wofs para el año 2020 donde el color rojo denota que no hubo agua, mientras que el color azul intenso denota un píxel que tuvo agua todo el año. Es decir, los ríos y los lagos se marcan en azul intenso, ya que tuvieron agua el 100 por ciento del tiempo.

Para comprobar que existe una correlación positiva entre inundaciones y áreas desfavorecidas, se debe validar que un valor positivo de la variable Wofs en una celda con población mayor a cero implica que esa celda es un área desfavorecida. Los pasos que se siguieron para la integración de la información son:

1. Adquirir datos de población y vivienda a nivel celda para la región de

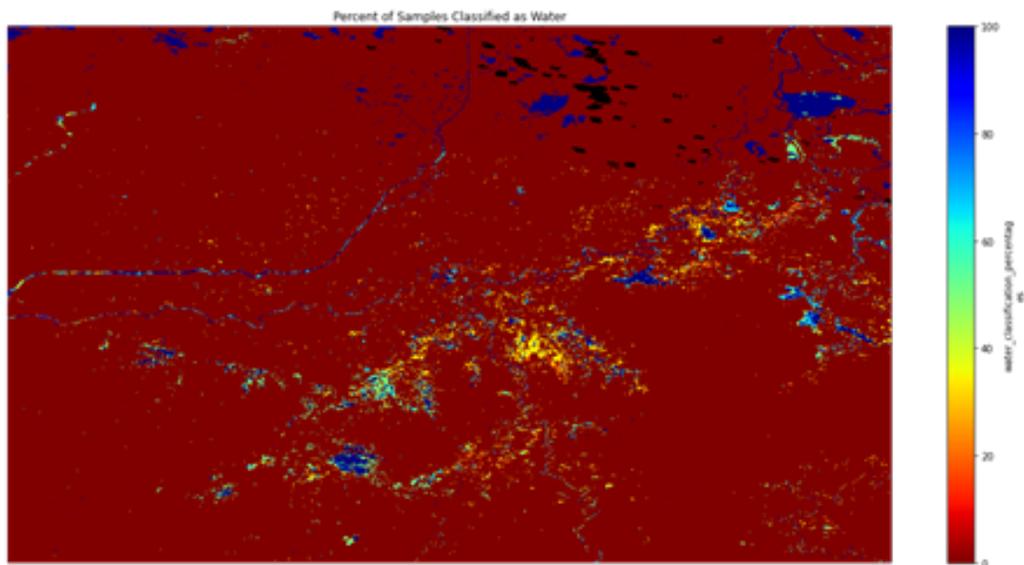


Figura 13: Resultado de la Clasificación WofS.

interés.

2. Incluir los datos de las imágenes para realizar la clasificación de cada celda como su porcentaje de Agua.
3. Integrar los datos e identificar el número de hogares en áreas desfavorecidas.

Los resultados de la integración de los datos a nivel celda con la variable Wofs nos arroja que, de las más de 3 millones de personas que habitaban la Cuenca en el 2020, 52,630 fueron afectadas por inundaciones. Este dato corresponde a todas las personas ubicadas en una vivienda cuya retícula tuvo un valor en la variable de Wofs mayor a cero. Esto equivale a un total de 17,799 viviendas afectadas.



Figura 14: Elaboración propia con datos del Censos de Población y Vivienda 2020.

La variable de interés son las áreas desfavorecidas que definimos en la sección 3.3.1 como un polígono que presenta un conjunto de factores de riesgos. En esta sección cuantificamos el área desfavorecida de la cuenca que corresponde a 2.1 kilómetros cuadrados afectados.

4.1. Validación cuantitativa

Hasta ahora hemos categorizado las áreas desfavorecidas como aquellas que cuentan con población y un valor de Wofs mayor a cero, lo que implica que en dichas áreas existen hogares que se han visto afectados por inundaciones.

Para validar que dichos hogares son realmente vulnerables debemos pro-

bar la consistencia entre la información proveniente del Wofs y la información sociodemográfica que es utilizada comúnmente para definir a un hogar como vulnerable.

El Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) es el organismo encargado de generar información objetiva sobre la medición de la pobreza en México y sobre la situación de la política social. El indicador de pobreza generado por el CONEVAL, únicamente tiene representatividad a nivel estatal y municipal. Estos datos no son de utilidad debido a que el indicador Wofs se estimó a nivel celda, necesitamos un indicador con menor nivel de granularidad.

El índice de rezago social que publica dicho organismo, presenta una alternativa para trabajar con datos con mayor nivel de desagregación. Se estima a través de la técnica estadística de componentes principales que, de acuerdo con CONEVAL [106], combina información de indicadores de carencias en un índice, que sintetiza numéricamente diferentes dimensiones de la pobreza. Este índice se tiene a nivel estatal, municipal y localidad. El Índice de Rezago Social, se estratifica en cinco categorías, de tal forma que dentro de cada categoría las unidades fueran lo más homogéneas posibles y entre los estratos lo más distintos posibles. De esta forma se crearon cinco estratos que son: muy bajo, bajo, medio, alto y muy alto rezago social, y que muestran de un menor a una mayor carencia en los indicadores que conforman el índice, respectivamente [106].

Para emplear el índice de rezago social, es necesario trabajar con los datos a nivel localidad, por lo que agrupamos la variable Wofs como el promedio de las celdas de cada localidad. Se utiliza el índice de rezago social como la verdad fundamental, comúnmente llamado “ground truth” en técnicas de ciencia de datos, para validar los resultados.

4.1.1. Índice de Moran

El índice de Moran calcula la relación de una variable en cierta localidad i versus otra variable en las localidades vecinas de i [107]. Utilizo el índice de Moran como una medida que nos determina la autocorrelación espacial para medir la influencia que tiene la variable Wofs en la ocurrencia del Índice de Rezago Social en las vecindades más próximas. En este caso, R son las localidades, x el Wofs, y el índice de rezago social y w son las coordenadas geográficas de cada localidad.

Se observan los vecinos de cada una de las localidades, por ejemplo, en el siguiente mapa de la figura 15, se señala el punto rojo como la localidad i y las localidades vecinas son coloreadas en verde oscuro.

El valor del índice tiene un signo negativo con un valor de -0.163 . Este valor indicaría que en localidades con valores altos del Índice de Rezago Social (más marginados) se tienen valores bajos en la variable Wofs y en sus localidades vecinas (lo cual es contraintuitivo a lo que se busca). O dicho de otra forma, las localidades donde se tienen mayores valores de Wofs, es decir,

$$I_t = \frac{R \sum_{i=1}^R \sum_{j=1}^R x_i y_j w_{ij}}{R_b \sum_{i=1}^R x_i^2}$$

Figura 15: Índice de Moran.

las más inundadas son aquéllas que presentan menores rezagos. Esto no va en línea con la hipótesis sobre el uso de las áreas inundadas para predecir áreas desfavorecidas.

Este resultado puede deberse a que si observamos la proporción de ceros para las localidades rurales con diferentes grados de rezago social observamos que el problema es que hay muchos ceros en las variables Wofs en las localidades rurales, y esto, al llevarlo al cálculo del índice podría, estar generando el sesgo o el signo contrario que observamos, y por ello, no está generando un índice de concordancia con respecto al Índice de Rezago Social, acorde a lo esperado.

Por lo anterior, el índice de Moran no es concluyente ya que no nos permite observar la relación entre las áreas inundadas y los altos índices de rezago social.

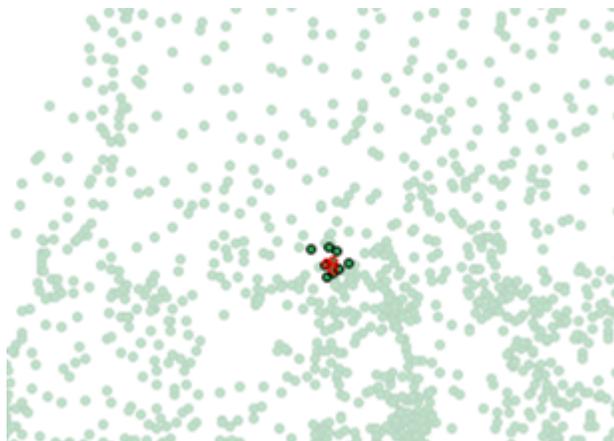


Figura 16: El Índice de Moran determina la autocorrelación espacial.

4.1.2. Datos con mayor desagregación

Los resultados del índice de Moran calculados para la variable Wofs y el índice de rezago social presentados en la sección anterior podrían contener ruido proveniente de la agrupación de los datos. Esto es, debido a que el índice de rezago social lo calculamos a nivel localidad, al agrupar la variable Wofs a dicho nivel de agregación podríamos estar perdiendo variabilidad en los datos.

Para calcular el índice con datos a nivel celda, se necesita generar una variable que permita identificar las características socioeconómicas a dicho nivel. Después de un riguroso análisis sobre las variables a nivel manzana con las que se cuenta en el Censo de Población y Vivienda 2020, se elige trabajar con el Índice de Nivel Socioeconómico (NSE) [108] que utiliza seis variables censales con el fin de modelar el ingreso corriente. Las variables que toma en cuenta son:

1. Nivel educativo del jefe de hogar
2. Número de baños completos en la vivienda
3. Número de autos en el hogar
4. Tenencia de conexión a internet en el hogar
5. Número de integrantes en el hogar mayores de 14 años que trabajan
6. Número de dormitorios en la vivienda

De acuerdo con el problema planteado, esperaríamos que a mayor vulnerabilidad, medida como un menor nivel socioeconómico, se observe un mayor nivel de la variable Wofs. Es decir, esperamos que los hogares vulnerables tengan mayor propensión a ser inundados. Para comprobar dicha hipótesis corremos de nuevo el índice de Moran ahora con la variable NSE a nivel de celda.

El valor del índice de Moran de nueva cuenta es muy cercano a cero (-.031), lo cual implicaría que hay muy poca relación negativa espacial. Por lo anterior, los resultados no son concluyentes.

4.1.3. Validación con modelos nulos

El adelanto en la tecnología computacional, ha detonado avances significativos en la modelación y la simulación en las ciencias. Diversos estudios

crean un vínculo entre la teoría y los modelos con supuestos utilizados para describir escenarios fiables. Por ejemplo, véase Lugo y Martínez 2022 [109].

En la modelación matemática, para asegurar que el modelo se comporta como deseamos, se utiliza la verificación para contar con un nivel de precisión consistente con la intención de la aplicación del modelo [110]. A partir de entonces, con la validación del modelo, buscamos tener la suficiente confianza en el método y sus resultados que nos permitan considerar al modelo válido. Así, el proceso de validación forma parte del proceso de desarrollo del modelo.

Existen diversas técnicas que pueden utilizarse en el proceso de validación [111]. En todas ellas se busca crear confianza en la correspondencia entre la aplicación del modelo y la teoría. Desde este momento, el proceso de evaluación permite confiar en las explicaciones dadas sobre los comportamientos del modelo y la comprensión teórica. Por otra parte, si el modelo no satisface los criterios de evaluación, puede ser necesario reformular el modelo teórico en el que se basa el modelo [112]. No existe un algoritmo o procedimiento para seleccionar qué técnicas de validación utilizar. La mayoría dependen del modelo, del sistema y del problema que se quiere analizar. En lo sucesivo, el investigador debe utilizar la más apropiada para su análisis [110].

En este caso, utilizo la validación con modelos nulos para comparar el comportamiento de salida del modelo para varios conjuntos de datos. Los modelos nulos se utilizan para validar si existe, al azar, algún patrón que esté afectando nuestros resultados. En este sentido, y en seguimiento a los

resultados de las secciones anteriores, queremos validar que no existe un patrón al azar relacionado con la variable WOFS. Si la correlación resulta positiva, entenderíamos que los datos no son los mejores para el estudio. Para corroborarlo empleamos una distribución imaginada.

Para ello, tomamos dos variables aleatorias. La primera con una distribución Normal estándar, con media cero y varianza uno. Y la segunda, una distribución Uniforme estándar con de igual manera media uno y varianza cero.

1. *Normal (0,1)*
2. *Uniforme (0,1)*

El resultado de la validación arroja que el índice de Moran, es prácticamente cero entre WOFS y la distribución Normal (0.005). De igual manera, el resultado con la distribución Uniforme es cero (-.001).

Con estos resultados concluyo que no existe asociación entre la variable WOFS y las variables aleatorias generadas mediante las distribuciones Normal y Uniforme.

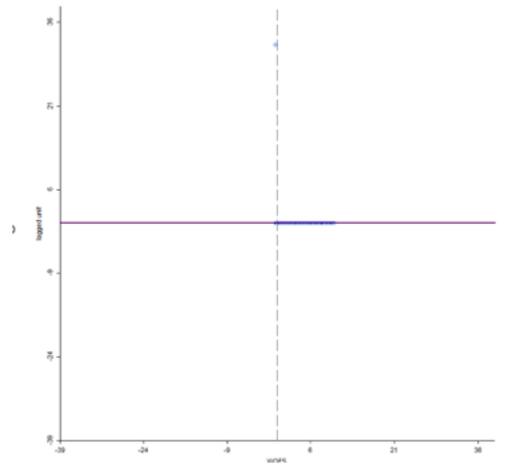
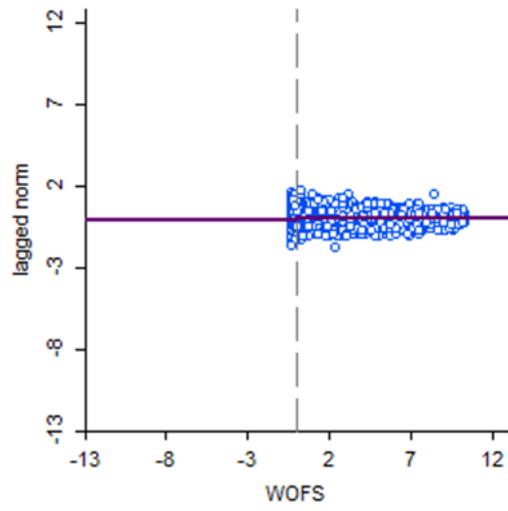


Figura 17: Resultados Modelos Nulos

4.2. Efectos temporales en el ingreso

La sección 5.1 nos indica que no hay una correlación de las inundaciones con los hogares vulnerables. Dichos resultados coinciden con diversas investigaciones que han mostrado que los efectos económicos de las inundaciones son temporales. El riesgo de desastre por inundación ocasiona disminuciones del ingreso por trabajo [113] y de la actividad económica medida, por ejemplo, través de las luces nocturnas [114].

Sperling [113], demuestra que las inundaciones en específico, para Tabasco, México en 2007, provocaron una disminución inmediata de los ingresos laborales de la población afectada. Este resultado se debe principalmente a un aumento en el número de trabajadores, predominantemente hombres, que continúan trabajando después de la inundación pero que ahora, ganan menos del salario mínimo. En este artículo Sperling observa que el sector agrícola experimentó la caída de ingresos más severa.

La literatura muestra que los efectos son temporales también, para otros tipos de desastres. Rodríguez-Oreggia [115] estudia el efecto de los huracanes en el mercado laboral en México, encontrando un aumento en los ingresos de trabajadores poco especializados debido a la mayor actividad en el sector de la construcción. Por su parte, Martínez [116] analiza el efecto del terremoto de Chile de 2010 en el mercado laboral, encontrando choques que desaparecen después de siete años de análisis. De tal forma, la capacidad destructiva de terremotos y huracanes generan choques temporales en la demanda (aumento) y la oferta por trabajo (reducción), incrementando el salario del sector

construcción. Estos efectos de desastres destructivos como los terremotos y los huracanes son contrarios a las inundaciones, donde Sperling [113] muestra una disminución inmediata en los ingresos por trabajo, particularmente en el sector primario.

Las variables e indicadores de vulnerabilidad o pobreza se centran generalmente en bienes duraderos como las condiciones de la vivienda (material de techos, pisos) y el equipamiento (televisor, electrodomésticos) que se explican principalmente por el ingreso permanente de los hogares. Es decir, no esperaríamos que varíen de manera significativa ante cambios temporales en el ingreso, por lo que es posible que las variables utilizadas para medir vulnerabilidad y pobreza no capturen los choques económicos generados por las inundaciones.

Adicionalmente, de acuerdo con Townsend [117], los hogares se hacen de mecanismos de aseguramiento formales o informales, lo cual puede ayudar a reducir la volatilidad del ingreso ante choques temporales. Dado que las inundaciones son recurrentes en la zona de estudio y tienen una conducta estacional, es de esperarse que los hogares se aseguran ante posibles pérdidas de ingreso y activos, y ajustan su conducta laboral y de consumo de forma estacional, reduciendo el impacto en el ingreso permanente y por ende en las variables que construye el índice de rezago social o el índice de nivel socio económico.

Por lo tanto, futuras investigaciones deberían concentrarse en analizar

variables que capturen cambios en el ingreso temporal que afectan el ahorro y el consumo en el corto plazo.

4.3. Validación en campo

« La realidad es compleja, multivariada y difícil de comprender, presentándose ante los ojos del investigador o del estudioso como múltiples realidades [118].»

Las entrevistas realizadas han permitido indagar las razones de los resultados obtenidos. Estas entrevistas han resultado idóneas para conocer el contexto local en cuanto a la medición de la pobreza y la vulnerabilidad ante inundaciones. A través de una serie de preguntas estructuradas interpreto la relación entre las áreas desfavorecidas ante inundación y la condición de rezago social en el estado de Tabasco. La libre expresión de las entrevistadas y los entrevistados proporciona relaciones complejas que deben entenderse en su propio contexto.

Los principales hallazgos del trabajo de campo pueden resumirse en los siguientes cinco puntos:

1. **El fenómeno de inundaciones en Tabasco es un problema histórico y recurrente.** El estado de Tabasco se caracteriza por la abundancia de agua, debido a que su sistema de ríos y lagunas desembocan en las costas del Golfo de México. El estado está incluido en dos regiones hidrológicas: Coatzacoalcos y Grijalva-Usumacinta. La segunda es

nuestro caso de estudio. El estado de Tabasco reúne un conjunto de características que lo hacen susceptible a sufrir inundaciones de manera frecuente [119]. En el trabajo de campo los entrevistados mencionaron recordar inundaciones como las de los años 1999,2007,2010,2011 y 2020. Mencionan un conjunto de causas tanto naturales como sociales, así como un mal manejo de presas y asentamientos en zonas de alta vulnerabilidad. Aunado a lo anterior existe una deforestación ocasionada por la agricultura, ganadería e industria, y la ausencia de una planeación territorial holística. Los entrevistados coinciden en que los fenómenos hidrometeorológicos de Tabasco son recurrentes y suceden en promedio cada 6.5 años. Además, caracterizan el fenómeno como uno complejo de resolver por tener múltiples causas e impactos.

- 2. La gran mayoría de la población tiene un grado de exposición alto a las inundaciones.** Los resultados obtenidos en el Capítulo 5 de esta investigación a través del análisis cuantitativo son confirmados en las entrevistas en campo. Las inundaciones, afectan a aproximadamente 40 por ciento del territorio del Estado. Es por eso, que se considera que el grado de exposición a las inundaciones es independiente del nivel de rezago social de la población.

De acuerdo Bollin y Hidajat [120] el riesgo por desastre está conformado por la amenaza, la exposición, la vulnerabilidad y las capacidades. Utilizado este marco conceptual podríamos inferir que los resultados obtenidos en esta investigación muestran que la exposición al riesgo es bastante homogénea para los distintos niveles de índice de rezago social. Sin embargo, los niveles altos de dicho índice aumentan el com-

ponente de vulnerabilidad del riesgo por desastre. Por lo tanto, como resultado de las entrevistas podemos definir la vulnerabilidad como un componente que, junto a información de las amenazas y la exposición de las personas, sus bienes y la infraestructura, permite calcular el riesgo como probabilidad de un daño no aceptable o crítico, en este caso las inundaciones.

- 3. La población en pobreza es más vulnerable ante el riesgo de inundaciones.** Ahora bien, a pesar de que una gran parte del Estado está expuesto al riesgo por inundaciones, el grado de vulnerabilidad es distinto de un hogar a otro dependiendo de los mecanismos con los que se cuenta para hacer frente a las inundaciones. En las entrevistas, fue repetida la anécdota de que la gente sube sus electrodomésticos al segundo piso cuando vienen las lluvias. Es cierto que la escala de los datos limita el análisis de vulnerabilidad. Si queremos medir el ingreso a nivel municipal, tenemos problemas debido a que no existe disponibilidad de datos a esa escala. Las variables censales son las que cuentan con una menor escala y con una periodicidad regular. El problema de dichos datos es que se generan cada diez años. El trabajo que aquí se presenta muestra una investigación que muestra el uso de datos no tradicionales como son las imágenes de satélite que permiten hacer análisis en los años intercensales. Con los datos aquí presentados se pueden mapear las áreas desfavorecidas ante inundaciones, y esto contribuye a la literatura que visualiza las diferencias que existen en el territorio. Aquí hemos encontrado, y el trabajo de campo lo confirma, que tanto hogares marginados como no marginados son vulnerables ante inundaciones

y han internalizado los costos por inundación.

4. **Se requieren mecanismos de cambio estructural para atender el fenómeno.** La necesidad de contar con mecanismos financieros de protección ante el riesgo de desastre por inundación fue mencionada en las entrevistas semiestructuradas realizadas. Los hacedores de política pública mencionaron la falta de mecanismos como los que existían en el pasado como el Fideicomiso Fondo de Desastres Naturales (Fonden) que permitía a los Estados atender la situación post desastre. Además, se mencionó la necesidad de contar con mecanismos financieros de aseguramiento que permitan proteger los bienes. En el caso de los hospitales y los bienes del gobierno, se requiere asegurar los materiales para poder garantizar la operación en el futuro. De igual forma, se requiere una inversión en vivienda con materiales de calidad adecuada que evite la construcción a ras de piso. Así como buscar la localización de la población fuera de zonas de riesgo.
5. **La información de alerta temprana y los atlas de riesgos deben difundirse y acercarse a la ciudadanía.** Por último, existe la necesidad de contar con estaciones meteorológicas a las que se les provea de mantenimiento constante que permitan dar avisos oportunos a la población y contar con sistemas de alerta temprana que provean del conocimiento necesario y la difusión a través de redes sociales. De igual forma, los atlas de riesgo no son difundidos y poco se conocen. A través de las entrevistas se concluye que la información oportuna, accesible y de fácil interpretación es imperante para la toma de decisiones ante el

riesgo por desastre. Los tomadores de decisiones que fueron entrevistados confirman la necesidad planteada en esta investigación de contar con datos oportunos listos para el análisis que sean difundidos a la población en lenguaje ciudadano. Los métodos utilizados actualmente para medir pobreza requieren de una actualización conforme a las fuentes de datos para poder proveer información con mayor oportunidad. Adicionalmente, los efectos del cambio climático deberán incluirse en los análisis de pobreza a fin de contar con un enfoque integral de vulnerabilidad social que permita generar políticas públicas que provean redes de protección a la población.

5. Discusión

A continuación discuto las contribuciones de esta investigación. En primera instancia menciono las aportaciones a la teoría económica a través del uso de fuentes de datos no tradicionales para mapear de manera replicable y con datos gratuitos las zonas desfavorecidas. En segundo lugar, planteo la utilidad de la investigación para contribuir a la evaluación de daños y pérdidas de los efectos del cambio climático en el contexto global actual a través de herramientas sistematizadas de fácil acceso a datos. Por último, considero relevante las aportaciones relativas a la reproducibilidad de la ciencia.

5.1. Teoría Económica

- ***Desarrollo Económico***: Aporta elementos a la medición de la pobreza a través del reconocimiento de los agentes y los hogares como pertenecientes a un entorno. Es decir, la medición de la pobreza tradicional se enfoca en las variables a nivel del hogar. En esta investigación se identifica a los hogares como unidades independientes, que existen en un entorno y que se encuentran interconectados con otros. La propuesta aquí planteada identifica la pobreza como un fenómeno complejo que debe ser estudiado de forma integral.
- ***Economía Regional y Urbana***: El análisis aquí presentado contiene una visión perteneciente al campo de la economía regional y urbana ya que analiza el fenómeno desde una perspectiva integral y toma a la región (las zonas desfavorecidas versus zonas no desfavorecidas) como

objeto de análisis económico. La aportación a dicha rama de la economía se basa en la identificación de los patrones de localización de la pobreza en ciudades del territorio mexicano.

- ***Geografía Económica***: La contribución a la geografía económica de esta investigación se basa en el estudio del fenómeno de la pobreza desde una perspectiva espacial y el análisis de la localización y características de los hogares en pobreza desde la dimensión espacio-temporal. La innovación en esta rama de la economía es la incorporación de la vulnerabilidad ante desastres como una variable dependiente del espacio y el tiempo que es constante y que coloca a los hogares en situación de catástrofe con regularidad.
- ***Métodos Cuantitativos***: La metodología aquí propuesta utiliza métodos de aprendizaje de máquina en la identificación de polígonos marginados. Esto contribuye al uso del cómputo en los métodos cuantitativos de la ciencia económica. El método no solo contribuye en la parte metodológica sino que también muestra el uso y la aplicabilidad de datos heterogéneos en el análisis económico cuantitativo.

Adicionalmente, hay distintas aportaciones de esta investigación a la ***Economía Aplicada*** a través de la generación de información para política pública. El mapeo de áreas desfavorecidas contribuye a la generación de políticas públicas en beneficio de las ciudades y aporta elementos para cumplir con los objetivos planteados en compromisos internacionales, como son los Objetivos de Desarrollo Sostenible de la Agenda 2030 de las Naciones Unidas. Principalmente, se contribuye al Objetivo 1: Poner fin a la pobreza

en todas sus formas en todo el mundo y al Objetivo 11: Lograr que las ciudades sean más inclusivas, seguras, resilientes y sostenibles.

5.2. Herramientas para la evaluación de daños y pérdidas ante los efectos del cambio climático

Una aportación clara e inmediata de esta investigación se da en el marco de los retos de cambio climático y la resiliencia ante desastres. Como se mencionó en la sección 2.2, el Panel Intergubernamental para el Cambio Climático ha alertado que, de no contener el calentamiento global, las pérdidas y los daños resultantes no serán prevenibles y las consecuencias serán inevitables.

Herramientas sistemáticas de generación de información son necesarias para el diseño y la evaluación de un desarrollo sostenible. Ante la creciente crisis climática, existe la necesidad de contar con herramientas sistematizadas para evaluar **los daños y pérdidas** relacionados con los desastres. De acuerdo con la Convención para el Cambio Climático de las Naciones Unidas, la pérdida y el daño generalmente pueden entenderse como los impactos negativos del cambio climático que ocurren a pesar de, o en ausencia de, acciones de mitigación y adaptación. Es por ello que, existe una necesidad latente de generar estimaciones ex ante y ex post de los daños y pérdidas

resultantes de los efectos del cambio climático. En este sentido, esta investigación ofrece métodos de cuantificación espacial de los posibles daños y pérdidas a través del uso de información estadística y geoespacial. Las herramientas metodológicas y analíticas que aquí se presentan, son de utilidad para medir los impactos devastadores del cambio climático.

Existen dos desafíos principales para la implementación y la operacionalización del uso de imágenes de satélite en los procesos de producción de información ante desastres: los desafíos institucionales y los desafíos tecnológicos. Por un lado, institucionalmente los hacedores de política pública no cuentan con experiencia y conocimiento sobre el uso de imágenes de satélite. Esto representa un gran obstáculo para la integración de nuevos sistemas de trabajo, técnicas y metodologías diferentes a las asociadas con procesos más tradicionales. Se necesita un gran esfuerzo en el desarrollo de capacidades para integrar completamente el flujo de trabajo con tecnología y nuevas metodologías. Por otro lado, el desafío tecnológico se refiere a la necesidad de personalizar la infraestructura institucional, de modo que permita gestionar, procesar y analizar los grandes volúmenes de datos; tarea especialmente difícil de diseñar e implementar sin un flujo de trabajo de producción previamente detallado como punto de partida. Aquí hemos presentado una metodología replicable y sencilla para mapear zonas desfavorecidas que no requiere de grandes poderes computacionales, lo que hace accesible la herramienta.

Para poder medir y evaluar los efectos de los desastres en la población, la Comisión Económica para América Latina y el Caribe CEPAL [121], ha

elaborado metodologías para la evaluación del impacto socioeconómico de los desastres. Dichas metodologías buscan estandarizar la medición del impacto de los desastres para contar con elementos para dar respuesta con información del impacto [122]. Las evaluaciones establecidas por la CEPAL y aplicadas en México a través del Centro Nacional de Prevención de Desastres, presentan información de forma retrospectiva, a partir de información dispersa elaborada con criterios dispares [123]. El método que presenta este trabajo de investigación contribuye a generar información uniformada de medición y evaluación de los efectos económicos y sociales de los desastres. La información aquí presentada, muestra las ventajas de dichos métodos para registrar sistemáticamente el impacto socioeconómico de los desastres para disponer de datos en un periodo amplio de tiempo que contribuya a establecer los fondos de contingencia. Asimismo, contribuye a generar mapas de amenazas, estimar vulnerabilidades y aporta a los esfuerzos de mitigación y prevención.

Este trabajo muestra los beneficios potenciales de cerrar la brecha entre el análisis espacial en ciencias sociales y la teledetección, caracterizando las áreas desfavorecidas fuertemente influenciadas por su medio ambiente. En los últimos años se ha presentado un aumento en las innovaciones de mapeo que utilizan *big data*, imágenes de satélite y aprendizaje automático. Si bien, su potencial para focalizar las políticas se ha demostrado de manera limitada, aún falta lograr una adopción generalizada: la comunicación y el compromiso son clave para promover la colaboración y, por lo tanto, el progreso hacia la sostenibilidad. Este trabajo colabora a la difusión del uso de métodos cuantitativos para la detección de zonas desfavorecidas así como la

evaluación de daños y pérdidas ante desastres. En esta investigación, el análisis se centró en el caso de estudio del fenómeno de inundaciones, sin embargo el análisis podría extenderse en futuras investigaciones a otros desastres hidrometeorológicos como ciclones tropicales y lluvias severas. Lo mismo para desastres geológicos como los deslaves y deslizamientos o incendios forestales.

Una reflexión interesante que debe realizarse al analizar las herramientas sistemáticas de generación de datos requeridas para la atención de desastres y el mapeo de áreas desfavorecidas es la relativa a la resolución y la desagregación de la información. Cada día, requerimos datos con una escala geográfica menor que nos permita contar con resolución suficiente para focalizar la política. Para ello, necesitamos datos con el mayor nivel de desagregación en el territorio, pero también para conocer por grupos de edad, sexo y distintas condiciones de la población. Los métodos de generación de información que se apoyan en tecnologías cada vez más avanzadas, están apuntando hacia nuevas metodologías de generación de datos que permitan dichas desagregaciones. La estandarización y la adopción de metodologías que utilizan datos no tradicionales son incipientes.

La información necesaria para evaluar los daños y pérdidas del cambio climático es necesaria para el diseño de instrumentos de política pública que permitan mitigar los efectos. Por ello, se busca que la información sea sostenible, regular, de calidad y de acceso gratuito. Esta investigación proporciona información que cuenta con estas características al utilizar datos abiertos de calidad.

5.3. Reproducibilidad de la Ciencia

Finalmente, la reproducibilidad en la ciencia es relevante en la generación de trabajos de investigación para actualizar las metodologías tradicionales ante la producción de datos masivos y el avance de la tecnología [124]. El método presentado en este trabajo de tesis, permite la validación por parte de otros investigadores ya que cuenta con un código que es público y trabaja con software abierto y datos gratuitos. Existen herramientas y guías como las Directrices de Promoción de la Transparencia y la Apertura (TOP, por sus siglas en inglés) [125] que establecen estándares que se pueden usar ampliamente en todas las disciplinas científicas como un mecanismo para el consenso sobre los estándares de integridad científica y para equipar mejor a los científicos para operar en un entorno de investigación que cambia rápidamente. De la misma manera, el marco MDAR [124], provee de una lista de verificación de criterios para impulsar la mejora de la investigación y, en última instancia, una mayor confianza en la ciencia. El beneficio de contar con herramientas para la fácil reproducibilidad del método, permite que esta investigación sea aplicada a otros contextos, así como a otros datos.

6. Conclusiones

El avance tecnológico ha puesto a disposición nuevas fuentes de información para la toma de decisiones. En esta investigación, utilizo fuentes de información no tradicionales, en este caso, observaciones de la Tierra para la identificación de áreas desfavorecidas. La metodología planteada, emplea variables estadísticas para definir vulnerabilidad y clasifica imágenes de satélite a través de métodos de aprendizaje de máquina. En el análisis, identifiqué espacialmente las áreas desfavorecidas a través de un enfoque multidisciplinario, contribuyendo a la teoría económica al identificar la pobreza como un fenómeno complejo desde una perspectiva espacio-temporal.

A través de un método riguroso, la investigación contribuye a la economía empírica ya que muestro el uso de datos no tradicionales, en específico, datos masivos como son las imágenes de satélite, para generar análisis empírico sobre estimaciones de pobreza.

La dificultad de medir pobreza a nivel hogar para poder generar índices de vulnerabilidad, es reflejo de la motivación de esta investigación. Las estimaciones de pobreza que producen los países, pocas veces permite la focalización de política pública con el detalle requerido debido a que las estimaciones provienen de censos y encuestas que, por un lado, cuentan con representatividad estatal y limita el análisis a nivel municipio o localidad. Por el otro, debido a la complejidad y costos de dichos eventos estadísticos, la información está disponible cada cinco o diez años lo que limita su uso

ante emergencias o desastres. Por ello, es necesario contar con variables a un mayor nivel de granularidad. Una primera conclusión de esta investigación es que necesitamos datos a nivel celda que permitan estimar los efectos de los desastres y su relación con vulnerabilidad. En ese sentido, estudios posteriores deberán utilizar microdatos, es decir, datos a nivel unidad de estudio, para observar el efecto en los ingresos y poder así mapear las zonas desfavorecidas.

Últimamente, ha crecido el interés por parte de las oficinas de estadística y organizaciones como la CEPAL en las metodologías llamadas “estimación para áreas pequeñas” [126], que permite combinar la información censal y de las encuestas para generar estimaciones con un nivel de desagregación mayor. En esta investigación, contribuimos al estudio de técnicas estadísticas para áreas pequeñas con el uso de un insumo adicional que son las Observaciones de la Tierra. Colombia, a través de su Departamento Administrativo Nacional de Estadística, el cual publica estimaciones de pobreza provenientes de análisis de imágenes de satélite utilizando el método de redes neuronales. Este trabajo contribuye al uso de dichas técnicas en México con el objetivo de contar con datos de pobreza con mayor desagregación geográfica y mayor oportunidad, es decir que se produzcan con mayor frecuencia.

Adicionalmente, como caso de estudio en este trabajo, he analizado el riesgo ante desastres por inundaciones. La conclusión principal es que el efecto de las inundaciones en Tabasco, debe estudiarse como un efecto temporal que afecta el ingreso por trabajo y tiene efectos poco significativos en los ingresos permanentes. Esto lo he concluido a partir del análisis geoestadístico

que he llevado a cabo utilizando el índice de rezago social correlacionado al algoritmo que cataloga la presencia de agua en las imágenes de satélite.

Esta investigación prueba que las fuentes de datos no tradicionales y técnicas computacionales como aprendizaje de máquina, pueden ser empleadas para el análisis económico. En específico, se integra información estadística y de observaciones de la Tierra para analizar la correlación entre inundaciones y vulnerabilidad. El principal hallazgo de la investigación consiste en identificar que el riesgo ante desastres por inundaciones debe analizarse como un efecto temporal que afecta el ingreso por trabajo y tiene efectos poco significativos en los ingresos permanentes.

A partir de dicho análisis, ofrezco recomendaciones de política pública. Primero, las políticas orientadas a subsanar las afectaciones por inundaciones deben definirse para hacer frente al daño potencial. Es decir, las políticas públicas deben implementarse para atender eventos que aún no han ocurrido. Esto, en relación a que la definición de vulnerabilidad nos indica que es un estado dinámico y por ello, la identificación de vulnerabilidades apoya la definición de políticas que permitan atender a la población en momentos críticos.

Adicionalmente, el trabajo empírico reveló que, como se planteó al inicio de esta investigación, es necesario contar con métodos no tradicionales de mapeo de pobreza. Los métodos de recolección de información a nivel nacional son poco útiles para comprender las realidades locales. Por ejemplo, el acceso a agua potable, que en algunas regiones del país es un buen indicador

para clasificar como un hogar pobre, en el contexto del amplio acceso al agua de Tabasco, pierde relevancia. Esto fue mencionado en las entrevistas por los expertos quienes señalan que la línea de pobreza de Tabasco subestima a la población en dicha condición. La necesidad de contar con mediciones locales con mayor nivel de granularidad y más oportunas se hace evidente en el trabajo de campo, dando relevancia a los métodos de identificación de población vulnerable a través de fuentes de datos no tradicionales. Para avanzar en este camino, trabajos de investigación futuros podrán continuar con el análisis, para aportar a la generación de datos y metodologías no tradicionales para un desarrollo sustentable.

Apéndice A Guía de Entrevista de Campo

Hola, buenos días/buenas tardes Estimado/a., le agradezco su tiempo para esta entrevista y su disposición para responder a mis preguntas. Sus respuestas van a ser, sin duda, muy valiosas para el objetivo de esta investigación.

Como le comentaba en el correo de invitación a esta entrevista, actualmente me encuentro en el quinto semestre del doctorado en Economía de la Facultad de Economía de la UNAM y realizo una investigación referente al uso de imágenes de la Tierra para la detección de vulnerabilidad. En particular, estoy interesada en estudiar el vínculo que existe entre la vulnerabilidad social y el riesgo ante desastres por inundación en la Cuenca Grijalva Villahermosa.

Los objetivos generales de la investigación son estudiar la relevancia del uso de datos no tradicionales en la identificación de áreas desfavorecidas y plantear soluciones para el manejo de datos no tradicionales en la producción de información espacial sobre vulnerabilidad. Esto, con el objetivo específico de acercar las técnicas de inteligencia artificial a las ciencias sociales, en específico a la economía en la generación de información estadística y geográfica. Además de ejemplificar el uso de metodologías de aprendizaje de máquina y otras técnicas innovadoras para la localización de áreas desfavorecidas.

En este sentido, en esta entrevista busca conocer desde su experiencia y perspectiva local la relación que existe entre las áreas inundadas y la pobreza

existente en el Estado. Esta información, será de utilidad para reconocer y visualizar el uso que se le puede otorgar a la información estadística y la información de percepción remota para localizar a la población vulnerable ante inundaciones.

Antes de empezar con la entrevista, quiero pedir su autorización para grabarla. Las grabaciones sólo serán utilizadas para este propósito y no se publicarán en ninguna parte sin su consentimiento.

Muchas gracias, ahora sí, comencemos con la entrevista:

1. ¿Considera que la población en pobreza es más vulnerable a desastres?
2. ¿Me puede comentar sobre el fenómeno de inundaciones en el estado de Tabasco?
3. En su experiencia ¿considera que existe una relación entre la pobreza y la vulnerabilidad ante inundaciones en el estado de Tabasco?
4. Desde su perspectiva y experiencia, ¿cuáles son los retos que existían y existen para atender a la población vulnerable?
5. ¿Qué piensa sobre los métodos utilizados en la localización de la población vulnerable? ¿Son efectivos?
6. Desde su perspectiva y experiencia, ¿Es posible identificar áreas desfavorecidas a través de fuentes de datos alternativos como son las imágenes de satélite clasificadas a través de técnicas de percepción remota?
7. Por último, ¿algo que le gustaría agregar?, una experiencia o una reflexión sobre el tema.

Le agradezco el tiempo y la disposición para esta entrevista. Sin duda, la información recabada aquí será de mucha utilidad.

Apéndice B Indicadores del Índice de Rezago Social

Para la construcción del índice de rezago social correspondiente al año 2020, se consideraron los siguientes indicadores de la base en información del Censo de Población y Vivienda 2020:

1. Educación

- Porcentaje de la población de 15 años y más analfabeta.
- Porcentaje de la población de 6 a 14 años que no asiste a la escuela.
- Porcentaje de la población de 15 años o más con educación básica incompleta.

2. Acceso a servicios de salud

- Porcentaje de la población sin derechohabiencia a servicios de salud.

3. Calidad y espacios en la vivienda

- Porcentaje de las viviendas particulares habitadas con piso de tierra.

4. Servicios básicos en la vivienda

- Porcentaje de las viviendas particulares habitadas que no disponen de excusado o sanitario.
- Porcentaje de las viviendas particulares habitadas que no disponen de agua entubada de la red pública.
- Porcentaje de las viviendas particulares habitadas que no disponen de drenaje.
- Porcentaje de las viviendas particulares habitadas que no disponen de energía eléctrica.

5. Activos en el hogar

- Porcentaje de las viviendas particulares habitadas que no disponen de lavadora.
- Porcentaje de las viviendas particulares habitadas que no disponen de refrigerador.

Apéndice C Código

El algoritmo utilizado en esta investigación toma el código abierto puesto a disposición de usuarios a través del Open Earth Alliance [92] y es modificado para generar los datos para la zona de estudio. El código empleado para generar el algoritmo de agua desde el espacio puede ser consultado en la siguiente [liga](https://colab.research.google.com/drive/1Em12lELqqlEZ4fTrNDR0pVoRt-eV7K-a?usp=sharing): <https://colab.research.google.com/drive/1Em12lELqqlEZ4fTrNDR0pVoRt-eV7K-a?usp=sharing>.

```
1 1. Title and author
2 title: PMG_Water_TimeSeries_L8_estructura
3 author: Paloma Merodio Gomez @palomamerodio tomado y
      modificado de Open Earth Alliance Sandbox [ver referencia
      en sección 3.4.1 ].
4
5 2. Import libraries
6
7 2.1 Import and connect to OpenDataCube colab
8 !wget -nc https://raw.githubusercontent.com/ceos-seo/odc-
      colab/master/odc_colab.py
9 from odc_colab import odc_colab_init
10 odc_colab_init(install_odc_gee=True)
11
12 2.2 Import populate data base
13 downloads an existing index
14 from odc_colab import populate_db
15
16 #populate data base
17 #populates the new ODC environment with index
18 populate_db()
```

```
19
20 # 2.3 Import dependencies and connect to data
21
22 import warnings
23
24 from odc_gee import earthengine
25
26 # Import Data Cube API
27 import utils.data_cube_utilities.data_access_api as dc_api
28
29 # Import Data Cube display map function
30 from utils.data_cube_utilities.dc_display_map import
    display_map
31
32 # Import Data Cube clean mask function
33 from utils.data_cube_utilities.clean_mask import
    landsat_qa_clean_mask
34
35 #Import Data Cube water classifier
36 from utils.data_cube_utilities.dc_water_classifier import
    wofs_classify
37
38 #Import the median mosaic function
39 from utils.data_cube_utilities.dc_mosaic import
    create_median_mosaic
40
41 # Import the plotting utility
42 from utils.data_cube_utilities.dc_rgb import rgb
43
44 #Import Data Cube export function
```

```

45 from utils.data_cube_utilities.import_export import
    export_slice_to_geotiff
46 from utils.data_cube_utilities.import_export import
    export_xarray_to_geotiff
47
48 # Import Utilities
49 import datetime
50 import numpy as np
51 import xarray as xr
52 import pandas as pd
53
54 import matplotlib.pyplot as plt
55
56 # Import color-scheme and set "no data" to black color
57 from matplotlib.cm import jet_r
58
59 # set config and authorize access by Earht Engine
60
61 # Suppress Warning Messages
62 warnings.filterwarnings('ignore')
63
64 # Load Data Cube Configuration
65 dc = earthengine.Datacube(app='Water_Observations_from_Space
    ')
66
67 #connect to Data Cube API
68 api = dc_api.DataAccessApi()
69
70 # 3. Class definition
71 #3.1 Class "ROI" Region of Interest

```

```

72
73 class ROI:
74
75     def __init__(self, latitude, longitude, platform, time_extents,
76                 product):
77         self.roi_latitude= latitude
78         self.roi_longitude= longitude
79         self.roi_data = dc.load(latitude=latitude, longitude=
80                                 longitude, platform=platform, time=time_extents,
81                                     product=product, measurements=['red
82                                     ', 'green', 'blue', 'nir', 'swir1', 'swir2', 'pixel_qa'])
83         self.roi_mask = landsat_qa_clean_mask(self.roi_data,
84                                                 platform=platform)
85
86     def wofs_percentages (self):
87
88         # Apply WOFs water classification. Only considering non-
89         # cloudy pixels
90         dataset=self.roi_data
91         mask=self.roi_mask
92
93         ts_water_classification = wofs_classify(dataset,
94                                                 clean_mask=mask.values)
95
96         # Ignore "no data" values
97         ts_water_classification = ts_water_classification.where(
98             ts_water_classification != -9999).astype(np.float16)
99
100        # Calculate the time series water classification product

```

```

94     roi_analysis = (ts_water_classification.mean(dim=['time
      ']) * 100).wofs.rename('water_classification_percentages')
95
96     # Return percentages of classification of water
97     return roi_analysis
98
99     def median_mosaic (self):
100         cleaned_dataset = self.roi_data.where(self.roi_mask)
101         roi_analysis = create_median_mosaic(cleaned_dataset, self
      .roi_mask)
102         return roi_analysis
103
104     def display_ROI(self):
105         return display_map(self.roi_latitude,self.roi_longitude)
106
107     #3.2 Class "TSA_result"   Time Series Analysis Result
108
109     class TSA_result:
110         def __init__(self,my_results):
111             self.resultsArray=my_results
112
113         def display_wofs (self,Title):
114             jet_r.set_bad('black', 1)
115             res = self.resultsArray()
116             res.plot(cmap=jet_r, figsize=(12,10))
117             plt.axis('off')
118             plt.title(Title)
119             plt.show()
120
121     def display_median (self,Title):

```

```

122     res = self.resultsArray()
123     median_rgb =res[['swir2', 'nir', 'green']].to_array()
124     median_rgb.plot.imshow(vmin=0, vmax=5000, figsize=(10,10)
125 )
126     plt.axis('off')
127     plt.title(Title)
128     plt.show()
129
130 def export_wofs(self,fileName):
131     file_name = "output/" + fileName + ".tif"
132     res = self.resultsArray()
133     dataset_to_export = xr.Dataset(coords=res.coords,attrs=
134 res.attrs)
135     dataset_to_export['wofs_pct'] = (res/100).astype(np.
136 float32)
137     export_xarray_to_geotiff(dataset_to_export,file_name)
138     !ls -lah output/
139
140 def export_median(self,fileName):
141     file_name = "output/" + fileName + ".tif"
142     res = self.resultsArray()
143     export_xarray_to_geotiff(res,file_name)
144     !ls -lah output/
145
146 # 4. Function definition
147
148 def latitude_limits (box_center, box_size_deg):
149     half_side = box_size_deg/2
150     return (box_center [0]-half_side,box_center [0]+half_side)
151

```

```

149 def longitude_limits (box_center, box_size_deg):
150     half_side = box_size_deg/2
151     return (box_center[1]-half_side,box_center[1]+half_side)
152
153 # 5. Variable generation and setting
154
155 # Define the Product and Platform
156 # This data is indexed from Google Earth Engine data sources
157
158 product = "ls8_google"
159 platform = "LANDSAT_8"
160
161 # 5.1 Villahermosa
162
163 # Select a Latitude-Longitude point for the center of the
164     analysis region
165 # Select the size of the box (in degrees) surrounding the
166     center point
167
168 # Villahermosa
169 # next: set the center of the region
170 lat_long = (17.99, -92.95)
171 # next: set the size of the square region to study (in
172     degrees)
173 box_size_deg = 0.5
174
175 # Select a time range
176 # The inputs require a format (Min,Max) using this date
177     format (YYYY-MM-DD)

```

```

174 # The Landsat-8 allowable time range is: 2013-04-07 to
      current
175 # next: set time extent to study
176 ej1_time_extents = ('2020-01-01', '2020-12-31')
177
178 # Calculate the latitude and longitude bounds of the analysis
      box:
179 ej1_latitude = latitude_limits(lat_long,box_size_deg)
180 ej1_longitude = longitude_limits(lat_long,box_size_deg)
181
182 Creation of Villahermosa Region Of Interest
183 roi_villahermosa = ROI(ej1_latitude,ej1_longitude,platform,
      ej1_time_extents,product)
184
185 #view ROI
186 roi_villahermosa.display_ROI()
187
188 # 7.1 Algorithms of WOFs and Median applied to ROI
189
190 # generate WOFs in villahermosa
191 roi_villahermosa_wofs = roi_villahermosa.wofs_percentages
192 tsa_villahermosa_wofs = TSA_result(roi_villahermosa_wofs)
193
194 # generate Median in Villahermosa
195 roi_villahermosa_median = roi_villahermosa.median_mosaic
196 tsa_villahermosa_median = TSA_result(roi_villahermosa_median)
197
198 # 8. Printing or showing output
199
200 # display WOFs in villahermosa

```

```
201 tsa_villahermosa_wofs.display_wofs("Villahermosa , WOFs")
202
203 #9. Export results
204 tsa_villahermosa_wofs.export_wofs("pmg_wofs_villahermosa")
205 tsa_villahermosa_median.export_median("
    pmg_median_villahermosa")
```

Referencias

- [1] J. E. Blumenstock, “Fighting poverty with data,” *Science*, vol. 353, no. 6301, pp. 753–754, 2016.
- [2] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon, “Combining satellite imagery and machine learning to predict poverty,” *Science*, vol. 353, no. 6301, pp. 790–794, 2016.
- [3] K. Shi, Z. Chang, Z. Chen, J. Wu, and B. Yu, “Identifying and evaluating poverty using multisource remote sensing and point of interest (poi) data: A case study of chongqing, china,” *Journal of Cleaner Production*, vol. 255, p. 120245, 2020.
- [4] M. Kuffer, K. Pfeffer, and R. Sliuzas, “Slums from space—15 years of slum mapping using remote sensing,” *Remote Sensing*, vol. 8, no. 6, p. 455, 2016.
- [5] A. Kirman, *Complex economics: individual and collective rationality*. Routledge, 2010.

- [6] J. Foster, J. Greer, and E. Thorbecke, “A class of decomposable poverty measures,” *Econometrica: journal of the econometric society*, pp. 761–766, 1984.
- [7] D. L. Newhouse, P. Suarez Becerra, and D. T. T. Doan, “Sri lanka: poverty and welfare: recent progress and remaining challenges,” 2016.
- [8] O. OECD, “Framework for statistics on the distribution of household income, consumption and wealth,” 2013.
- [9] M. Ravallion and S. Chen, “Measuring pro-poor growth,” *Economics letters*, vol. 78, no. 1, pp. 93–99, 2003.
- [10] A. Smith, “The wealth of nations [1776],” 1937.
- [11] S. Amartya, “Desarrollo y libertad, editorial planeta,” 2006.
- [12] J. D. Domínguez and A. M. M. Caraballo, “Medición de la pobreza: una revisión de los principales indicadores,” *Revista de Métodos Cuantitativos para la Economía y la Empresa*, vol. 2, pp. 27–66, 2006.
- [13] S. Alkire and M. E. Santos, “Acute multidimensional poverty: A new index for developing countries,” 2011.
- [14] A. B. Atkinson *et al.*, “On the measurement of inequality,” *Journal of economic theory*, vol. 2, no. 3, pp. 244–263, 1970.
- [15] A. García Andrés, N. A. Fuentes, and O. Montes García, “Desigualdad y polarización del ingreso en méxico: 1980-2008,” *Política y cultura*, no. 37, pp. 285–310, 2012.

- [16] C. CONEVAL, “Metodología para la medición multidimensional de la pobreza en México,” 2016.
- [17] C. CONEVAL, “10 años de medición de pobreza en México, avances y retos en política social (comunicado de prensa),” 2019.
- [18] H. Nuissl and D. Heinrichs, “Slums: perspectives on the definition, the appraisal and the management of an urban phenomenon,” *DIE ERDE—Journal of the Geographical Society of Berlin*, vol. 144, no. 2, pp. 105–116, 2013.
- [19] E. L. Birch *et al.*, “Alan Mayne 2017: Slums: The history of a global injustice. London: Reaktion Books Ltd,” *International Journal of Urban and Regional Research*, vol. 42, no. 6, pp. 1158–1159, 2018.
- [20] A. Patel, N. Koizumi, and A. Crooks, “Measuring slum severity in Mumbai and Kolkata: A household-based approach,” *Habitat International*, vol. 41, pp. 300–306, 2014.
- [21] t. . Naciones Unidas.
- [22] G. A. M. Botello and G. E. G. Robles, “Documento final de la reunión temática sobre espacios públicos de Habitat III,” *“PATRIMONIO”: Economía Cultural y Educación para la Paz (MEC-EDUPAZ)*, vol. 2, no. 14, pp. 137–159.
- [23] L. G. D. A. HUMANOS and C. PRIMERO, “Ley general de asentamientos humanos,” *Extraído de <http://www.diputados.gob.mx/LeyesBiblio/ref/lgahotdu.htm> Última reforma publicada en el DOF del*, vol. 3, 2020.

- [24] SEDATU, “Reglas de operación del programa para regularizar asentamientos humanos irregulares.”
- [25] INEGI, “Censo nacional de gobiernos municipales y demarcaciones territoriales de la ciudad de México 2019.”
- [26] D. R. Thomson, M. Kuffer, G. Boo, B. Hati, T. Grippa, H. Elsey, C. Linard, R. Mahabir, C. Kyobutungi, J. Maviti, *et al.*, “Need for an integrated deprived area “slum” mapping system (ideamaps) in low- and middle-income countries (lmics),” *Social Sciences*, vol. 9, no. 5, p. 80, 2020.
- [27] D. Kohli, R. Sliuzas, N. Kerle, and A. Stein, “An ontology of slums for image-based classification,” *Computers, Environment and Urban Systems*, vol. 36, no. 2, pp. 154–163, 2012.
- [28] R. Lilford, C. Kyobutungi, R. Ndugwa, J. Sartori, S. I. Watson, R. Sliuzas, M. Kuffer, T. Hofer, J. P. de Albuquerque, and A. Ezech, “Because space matters: conceptual framework to help distinguish slum from non-slum urban areas,” *BMJ global health*, vol. 4, no. 2, p. e001267, 2019.
- [29] D. Verma, A. Jana, and K. Ramamritham, “Transfer learning approach to map urban slums using high and medium resolution satellite imagery,” *Habitat International*, vol. 88, p. 101981, 2019.
- [30] M. R. Ibrahim, H. Titheridge, T. Cheng, and J. Haworth, “predicts-lums: A new model for identifying and predicting informal settlements

- and slums in cities from street intersections using machine learning,” *Computers, Environment and Urban Systems*, vol. 76, pp. 31–56, 2019.
- [31] Á. Abascal, N. Rothwell, A. Shonowo, D. R. Thomson, P. Elias, H. Elsey, G. Yeboah, and M. Kuffer, “” domains of deprivation framework” for mapping slums, informal settlements, and other deprived areas in lmics to improve urban planning and policy: A scoping review,” 2021.
- [32] J. Birkmann, “Measuring vulnerability to promote disaster-resilient societies: Conceptual frameworks and definitions,” *Measuring vulnerability to natural hazards: Towards disaster resilient societies*, vol. 1, pp. 9–54, 2006.
- [33] N. Ruiz Rivera, “La definición y medición de la vulnerabilidad social. un enfoque normativo,” *Investigaciones geográficas*, no. 77, pp. 63–74, 2012.
- [34] J. Alwang, P. B. Siegel, S. L. Jorgensen, *et al.*, “Vulnerability: a view from different disciplines,” tech. rep., Citeseer, 2001.
- [35] C. E. Natenzon and A. B. Parkinson, “El derecho como instrumento de adaptación al cambio climático. revisión de sentencias relativas a desastres por inundaciones urbanas,” *AREA, Agenda de Reflexión en Arquitectura, Diseño y Urbanismo*, no. 26, p. 17, 2020.
- [36] B. L. Turner, R. E. Kasperson, P. A. Matson, J. J. McCarthy, R. W. Corell, L. Christensen, N. Eckley, J. X. Kasperson, A. Luers, M. L. Martello, *et al.*, “A framework for vulnerability analysis in sustainabi-

- lity science,” *Proceedings of the national academy of sciences*, vol. 100, no. 14, pp. 8074–8079, 2003.
- [37] E. L. Quarantelli, *What is a disaster?: perspectives on the question*. Psychology Press, 1998.
- [38] D. Ríos and C. Natenzon, “Una revisión sobre catástrofes, riesgo y ciencias sociales,” *Riesgos, catástrofes y vulnerabilidades. Buenos Aires: Ediciones Imago Mundi*, 2015.
- [39] C. Gilbert, “Studying disaster: a review of the main conceptual tools,” *International Journal of Mass Emergencies and Disasters*, vol. 13, no. 3, pp. 231–240, 1995.
- [40] A. Oliver-Smith, “What is a disaster? anthropological perspectives on a persistent question,” *The angry earth: Disaster in anthropological perspective*, pp. 18–34, 1999.
- [41] U. N. Conference of the Parties, “Decisions taken at the sharm el-sheikh climate change conference,” 2022.
- [42] C. on Biological Diversity, “Decisión adoptada por la conferencia de las partes de la convención de la diversidad biológica,” 19 de diciembre de 2022.
- [43] N. Unidas, “Centro regional de información,” 2022.
- [44] UNDRR, “Declaración de punta del este. viii plataforma regional para la reducción del riesgo de desastres en las américas y el caribe,” 2023.

- [45] P. Krugman, “The self-organizing economy, mitsui lectures in economics,” 1996.
- [46] D. Colander and R. Kupers, *Complexity and the art of public policy: Solving society’s problems from the bottom up*. Princeton University Press, 2014.
- [47] M. Batty, *The new science of cities*. MIT press, 2013.
- [48] I. Turok and J. Borel-Saladin, “The theory and reality of urban slums: Pathways-out-of-poverty or cul-de-sacs?,” *Urban Studies*, vol. 55, no. 4, pp. 767–789, 2018.
- [49] A. Datta, “How many satellites orbit earth and why space traffic management is crucial,” Aug 2020.
- [50] R. Engstrom, D. Pavelesku, T. Tanaka, and A. Wambile, “Mapping poverty and slums using multiple methodologies in accra, ghana,” in *2019 Joint Urban Remote Sensing Event (JURSE)*, pp. 1–4, IEEE, 2019.
- [51] J. C. Duque, J. E. Patino, L. A. Ruiz, and J. E. Pardo-Pascual, “Measuring intra-urban poverty using land cover and texture metrics derived from remote sensing data,” *Landscape and Urban Planning*, vol. 135, pp. 11–21, 2015.
- [52] M. Xie, N. Jean, M. Burke, D. Lobell, and S. Ermon, “Transfer learning from deep features for remote sensing and poverty mapping,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, 2016.

- [53] D. Roy, D. Bernal, and M. Lees, “An exploratory factor analysis model for slum severity index in mexico city,” *Urban studies*, vol. 57, no. 4, pp. 789–805, 2020.
- [54] S. Mason and C. Fraser, “Image sources for informal settlement management,” *The Photogrammetric Record*, vol. 16, no. 92, pp. 313–330, 1998.
- [55] U. J. Dorji, A. Plangprasopchok, N. Surasvadi, and C. Siripanpornchana, “A machine learning approach to estimate median income levels of sub-districts in thailand using satellite and geospatial data,” in *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery*, pp. 11–14, 2019.
- [56] C. A. Baquero Oleas, “Evaluación de ecosistemas de páramo inundables, mediante variables espectrales y técnicas de clasificación orientada a objetos; en la reserva de producción de fauna chimborazo,” 2020.
- [57] I. R. Buitrago Arévalo, “Clasificación y evaluación de una imagen satelital utilizando los métodos de redes neuronales artificiales y segmentación espacial,”
- [58] R. Sliuzas, N. Kerle, and M. Kuffer, “Object-oriented mapping of urban poverty and deprivation,” 2008.
- [59] I. Baud, M. Kuffer, K. Pfeffer, R. Sliuzas, and S. Karuppappan, “Understanding heterogeneity in metropolitan india: The added value of remote sensing data for analyzing sub-standard residential areas,” *In-*

ternational Journal of Applied Earth Observation and Geoinformation, vol. 12, no. 5, pp. 359–374, 2010.

- [60] P. Merodio, “Configuración espacial de las zonas marginadas en la ciudad de ... - nexos,” Sep 2020.
- [61] I. N. de Estadística y Geografía, “Censo de población y vivienda 2020. síntesis metodológica y conceptual,” 2020.
- [62] N. CEPAL, “Boletín de actividades de censos de población y habitación, números 1 a 9,” 1981.
- [63] M. KUFFER, “Global urban data gaps: Machine learning, earth observation and deprived urban areas,” 2020.
- [64] J. Wang, M. Kuffer, D. Roy, and K. Pfeffer, “Deprivation pockets through the lens of convolutional neural networks,” *Remote sensing of environment*, vol. 234, p. 111448, 2019.
- [65] M. Kuffer, D. R. Thomson, G. Boo, R. Mahabir, T. Grippa, S. Vanhuysse, R. Engstrom, R. Ndugwa, J. Makau, E. Darin, *et al.*, “The role of earth observation in an integrated deprived area mapping “system” for low-to-middle income countries,” *Remote Sensing*, vol. 12, no. 6, p. 982, 2020.
- [66] I. N. de Estadística y Geografía, “Producción y publicación de la geomediana nacional a partir de imágenes del cubo de datos geoespaciales,” 2020.

- [67] D. Martin and I. Bracken, “Techniques for modelling population-related raster databases,” *Environment and Planning A*, vol. 23, no. 7, pp. 1069–1075, 1991.
- [68] W. Tobler, U. Deichmann, J. Gottsegen, and K. Maloy, “World population in a grid of spherical quadrilaterals,” *International Journal of Population Geography*, vol. 3, no. 3, pp. 203–225, 1997.
- [69] W. Tobler, U. Deichmann, J. Gottsegen, and K. Maloy, “The global demography project (95-6),” 1995.
- [70] M. F. Goodchild and N. S.-N. Lam, *Areal interpolation: a variant of the traditional spatial problem*. Department of Geography, University of Western Ontario London, ON, Canada, 1980.
- [71] J. Mennis and T. Hultgren, “Intelligent dasymetric mapping and its application to areal interpolation,” *Cartography and Geographic Information Science*, vol. 33, no. 3, pp. 179–194, 2006.
- [72] E. Doxsey-Whitfield, K. MacManus, S. B. Adamo, L. Pistolessi, J. Squires, O. Borkovska, and S. R. Baptista, “Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4,” *Papers in Applied Geography*, vol. 1, no. 3, pp. 226–234, 2015.
- [73] J. E. Cohen and C. Small, “Hypsographic demography: the distribution of human population by altitude,” *Proceedings of the National Academy of Sciences*, vol. 95, no. 24, pp. 14009–14014, 1998.

- [74] D. R. Thomson, D. A. Rhoda, A. J. Tatem, and M. C. Castro, “Gridded population survey sampling: a systematic scoping review of the field and strategic research agenda,” *International journal of health geographics*, vol. 19, no. 1, pp. 1–16, 2020.
- [75] S. Leyk, A. E. Gaughan, S. B. Adamo, A. d. Sherbinin, D. Balk, S. Freire, A. Rose, F. R. Stevens, B. Blankespoor, C. Frye, *et al.*, “The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use,” *Earth System Science Data*, vol. 11, no. 3, pp. 1385–1409, 2019.
- [76] S. Leyk, M. Ruther, B. P. Buttenfield, N. N. Nagle, and A. K. Stum, “Modeling residential developed land in rural areas: A size-restricted approach using parcel data,” *Applied Geography*, vol. 47, pp. 33–45, 2014.
- [77] F. R. Stevens, A. E. Gaughan, C. Linard, and A. J. Tatem, “Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data,” *PloS one*, vol. 10, no. 2, p. e0107042, 2015.
- [78] J. Mennis, “Generating surface models of population using dasymetric mapping,” *The Professional Geographer*, vol. 55, no. 1, pp. 31–42, 2003.
- [79] D. R. Thomson, A. E. Gaughan, F. R. Stevens, G. Yetman, P. Elias, and R. Chen, “Evaluating the accuracy of gridded population estimates in slums: A case study in nigeria and kenya,” 2021.
- [80] CENAPRED, “Sistema de consulta de declaratorias.”

- [81] CONAGUA, “Sistema nacional de información del agua.”
- [82] P. Merodio and A. Ramirez, “Afectación municipal por múltiples desastres: un índice para México,” 2021.
- [83] G. J. Schumann, G. R. Brakenridge, A. J. Kettner, R. Kashif, and E. Niebuhr, “Assisting flood disaster response with earth observation data and products: A critical assessment,” *Remote Sensing*, vol. 10, no. 8, p. 1230, 2018.
- [84] G. J.-P. Schumann and D. K. Moller, “Microwave remote sensing of flood inundation,” *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 83, pp. 84–95, 2015.
- [85] L. C. Smith, “Satellite remote sensing of river inundation area, stage, and discharge: A review,” *Hydrological processes*, vol. 11, no. 10, pp. 1427–1439, 1997.
- [86] U. de Murcia, “Fotointerpretación y teledetección.”
- [87] C. J. Tucker, “Red and photographic infrared linear combinations for monitoring vegetation,” *Remote sensing of Environment*, vol. 8, no. 2, pp. 127–150, 1979.
- [88] A. Thakkar, V. Desai, A. Patel, and M. Potdar, “Land use/land cover classification using remote sensing data and derived indices in a heterogeneous landscape of a Khan-Kali watershed, Gujarat,” *Asian Journal of Geoinformatics*, vol. 14, no. 4, 2015.

- [89] P. S. Frazier, K. J. Page, *et al.*, “Water body detection and delineation with landsat tm data,” *Photogrammetric engineering and remote sensing*, vol. 66, no. 12, pp. 1461–1468, 2000.
- [90] M. G. Tulbure and M. Broich, “Spatiotemporal dynamic of surface water bodies using landsat time-series data from 1999 to 2011,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 79, pp. 44–52, 2013.
- [91] N. Mueller, A. Lewis, D. Roberts, S. Ring, R. Melrose, J. Sixsmith, L. Lymburner, A. McIntyre, P. Tan, S. Curnow, *et al.*, “Water observations from space: Mapping surface water from 25 years of landsat imagery across australia,” *Remote Sensing of Environment*, vol. 174, pp. 341–352, 2016.
- [92] B. Killough, S. Rizvi, and A. Lubawy, “Advancements in the open data cube and the use of analysis ready data in the cloud,” in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 1793–1795, IEEE, 2021.
- [93] A.-L. Barabási, “The network takeover,” *Nature Physics*, vol. 8, no. 1, pp. 14–16, 2012.
- [94] CONEVAL, “Índice de rezago social 2020,” Apr 2021.
- [95] S. Khajehei, A. Ahmadalipour, W. Shao, and H. Moradkhani, “A place-based assessment of flash flood hazard and vulnerability in the contiguous united states,” *Scientific reports*, vol. 10, no. 1, pp. 1–12, 2020.

- [96] B. E. Flanagan, E. W. Gregory, E. J. Hallisey, J. L. Heitgerd, and B. Lewis, “A social vulnerability index for disaster management,” *Journal of homeland security and emergency management*, vol. 8, no. 1, 2011.
- [97] A. Fekete, “Validation of a social vulnerability index in context to river-floods in germany,” *Natural Hazards and Earth System Sciences*, vol. 9, no. 2, pp. 393–403, 2009.
- [98] M. Esperón-Rodríguez and V. L. Barradas, “Comparing environmental vulnerability in the montane cloud forest of eastern mexico: A vulnerability index,” *Ecological Indicators*, vol. 52, pp. 300–310, 2015.
- [99] C. Borja-Vega and A. de la Fuente, *Municipal vulnerability to climate change and climate related events in Mexico*. The World Bank, 2013.
- [100] M. Ravallion, *Issues in measuring and modeling poverty*. The World Bank, 1999.
- [101] M. C. Schmittlein, R. C. Deutsch, W. W. Piegorsch, and S. L. Cutter, “A sensitivity analysis of the social vulnerability index,” *Risk Analysis: An International Journal*, vol. 28, no. 4, pp. 1099–1114, 2008.
- [102] E. Skoufias, B. Davis, and S. De la Vega, “Targeting the poor in mexico: An evaluation of the selection of households for progresas,” tech. rep., 2001.
- [103] M. Aravena, E. Kimelman, B. Micheli, R. Torrealba, and J. Zúñiga, “Investigación educativa i,” 2006.

- [104] S. Toledo, “Diferencias entre cuestionario y encuesta,” *Universidad virtual de salud Manuel Fajardo*, 2012.
- [105] J.-L. Bessis, J. Bequignon, and A. Mahmood, “The international charter “space and major disasters” initiative,” *Acta Astronautica*, vol. 54, no. 3, pp. 183–190, 2004.
- [106] C. N. de Evaluación de la Política de Desarrollo Social, “Los mapas de pobreza en México,” 2007.
- [107] D. Oxoli, S. Sabri, A. Rajabifard, and M. A. Brovelli, “A classification technique for local multivariate clusters and outliers of spatial association,” *Transactions in GIS*, vol. 24, no. 5, pp. 1227–1247, 2020.
- [108] A. N. S. AMAI, “Nota metodológica,” *Mexico: Asociacion Mexicana de Agencias de Inteligencia Mercado y Opinion*, 2018.
- [109] I. Lugo and G. Martínez-Mekler, “Theoretical study of the effect of ports in the formation of city systems,” *Journal of Shipping and Trade*, vol. 7, no. 1, p. 16, 2022.
- [110] R. G. Sargent, “Simulation model validation,” in *Simulation and model-based methodologies: an integrative view*, pp. 537–555, Springer, 1984.
- [111] W. L. Oberkampf and C. J. Roy, *Verification and validation in scientific computing*. Cambridge University Press, 2010.
- [112] C. A. Aumann, “A methodology for developing simulation models of complex systems,” *Ecological Modelling*, vol. 202, no. 3-4, pp. 385–396, 2007.

- [113] L. L. Sperling, “Flooded jobs: income development after the 2007 tabasco flood,” *Climate and Development*, pp. 1–26, 2021.
- [114] A. Kocornik-Mina, T. K. McDermott, G. Michaels, and F. Rauch, “Flooded cities,” *American Economic Journal: Applied Economics*, vol. 12, no. 2, pp. 35–66, 2020.
- [115] E. Rodríguez-Oreggia, “Hurricanes and labor market outcomes: Evidence for Mexico,” *Global environmental change*, vol. 23, no. 1, pp. 351–359, 2013.
- [116] M. J. Martínez, M. J. Martínez, and R. Romero-Jarén, “How resilient is the labour market against natural disaster? evaluating the effects from the 2010 earthquake in Chile,” *Natural Hazards*, vol. 104, no. 2, pp. 1481–1533, 2020.
- [117] R. M. Townsend, *Financial systems in developing Economies: growth, inequality and policy Evaluation in Thailand*. Oxford University Press, 2011.
- [118] M. García Ferrando, J. Ibáñez, and F. Alvira, “El análisis de la realidad social. métodos y técnicas de investigación (the analysis of social reality. research methods and techniques),” 1998.
- [119] F. I. Arreguín-Cortés, H. Rubio-Gutiérrez, R. Domínguez-Mora, and F. d. Luna-Cruz, “Análisis de las inundaciones en la planicie tabasqueña en el periodo 1995-2010,” *Tecnología y ciencias del agua*, vol. 5, no. 3, pp. 05–32, 2014.

- [120] C. Bollin, R. Hidajat, and J. Birkmann, “Community-based risk index: Pilot implementation in indonesia,” *Measuring vulnerability to natural hazards: Towards disaster resilient societies*, vol. 271, p. 89, 2006.
- [121] P. CEPAL, “Manual para la evaluación del impacto socioeconómico y ambiental de los desastres,” *Santiago de Chile*, 2003.
- [122] D. Bitrán, “Evaluación retrospectiva del impacto socioeconómico de los desastres y consideraciones metodológicas para llevarla a cabo,” 2006.
- [123] D. Bitrán, *Metodología para la evaluación del impacto socioeconómico de los desastres*. CEPAL, 2009.
- [124] M. Macleod, A. M. Collings, C. Graf, V. Kiermer, D. Mellor, S. Swaminathan, D. Sweet, and V. Vinson, “The mdar (materials design analysis reporting) framework for transparent reporting in the life sciences,” 2021.
- [125] A. Kretser, D. Murphy, S. Bertuzzi, T. Abraham, D. B. Allison, K. J. Boor, J. Dwyer, A. Grantham, L. J. Harris, R. Hollander, *et al.*, “Scientific integrity principles and best practices: recommendations from a scientific integrity consortium,” *Science and Engineering Ethics*, vol. 25, pp. 327–355, 2019.
- [126] C. Elbers, J. O. Lanjouw, and P. Lanjouw, “Micro-level estimation of poverty and inequality,” *Econometrica*, vol. 71, no. 1, pp. 355–364, 2003.