

Integridad Ecosistémica

La Integridad Ecosistémica (IE) describe la salud de un ecosistema, se refiere a que tan intacto, completo o funcional está respecto a su estado natural (Wildlife Conservation Society 2020). La integridad más alta la alcanzan las áreas no afectadas de manera significativa por actividades humanas, las cuales son fundamentales para la conservación de la biodiversidad, pues es en ecosistemas con alta integridad que las especies tienen menor riesgo de extinción (Di Marco et al. 2019). Además, estas áreas contribuyen en mayor medida, respecto a áreas degradadas, a los servicios ecológicos, como la absorción de CO₂, el suministro de agua y protección de riesgos causados por el cambio climático (Watson et al. 2018; Martin and Watson 2016). Por lo cual, es de gran importancia, medir la IE para preservar áreas con alta integridad y rehabilitar áreas de baja integridad.

Estimación de la Integridad Ecosistémica

Con el fin de estimar la IE para distintos años en todo el territorio mexicano, se ajustó un modelo llamado “extreme gradient boosting” (refuerzo del gradiente extremo), o XGBoost por sus siglas en inglés, el cual es un método de aprendizaje automático supervisado para clasificación o regresión, que mediante variables predictoras, ya sea continuas o categóricas, puede predecir el valor de interés, en este caso el Índice de IE (IIE). Éste método combina modelos débiles, es decir modelos con baja precisión, comúnmente árboles de decisión, para que en conjunto se obtenga una predicción mucho más exacta. El entrenamiento es iterativo, agregando en cada paso un nuevo árbol de decisión que predice el error de los árboles anteriores. Al final, se combinan las predicciones de los árboles en una predicción total, la cual suele ser más precisa que otros modelos, como bosque aleatorio o maquina de soporte vectorial.

Para predecir el valor del IIE, es necesario definir los valores del IIE que se tengan a un tiempo dado, para después entrenar el modelo con estos y con el conjunto de variables predictoras, de tal forma que se pueda crear la relación de los valores del IIE con las variables. Al hacer esto, tendremos un modelo que a partir de los valores de las variables predictoras de cualquier periodo, estima el valor del IIE. Es importante notar que una mejor modelación va a depender de la elección que se haga de los parámetros del modelo XGBoost y de lo relevante de las variables predictoras para el valor del IIE.

Índice de Integridad Ecosistémica

La IE es un concepto complejo, por lo que para su medición usamos una medida *proxy* que refleje el impacto humano en la naturaleza (Hansen et al. 2021). En este caso se definió un IIE de acuerdo al grado de transformación de la vegetación natural de los ecosistemas, respecto a su situación actual. Este concepto, también llamado *hemerobia*, se ha utilizado en trabajos previos [Equihua et al. (2024)](Instituto Nacional de Estadística Geografía e Informática 2021b) para la estimación de la IE.

Para su construcción se comparó la *vegetación primaria* (vegetación antes de los grandes cambios antropogénicos) (Instituto Nacional de Estadística Geografía e Informática 2003), con el mapa de uso de suelo y vegetación más reciente (Instituto Nacional de Estadística Geografía e Informática 2021a), el cual contiene datos de 2017. Se asignó un valor al índice de acuerdo a la severidad del cambio, tomando los valores descritos en [Equihua et al. (2024)](Instituto Nacional de Estadística Geografía e Informática 2021b) y de acuerdo a la siguiente tabla. Por ejemplo, si pasa de bosque de coníferas a agricultura de riego, se asigna un valor 13, que corresponde a una *degradación severa media superior*, o bien si discretizamos el índice en cuatro categorías (alta, media, baja y muy baja) corresponde a *integridad baja*. De esta forma el IIE es una medida que refleja el impacto de la intervención humana en la vegetación de los ecosistemas.

	IIE (19 categorías)	IIE (4 Categorías)	Descripción
0	Estasis	Integridad alta	No hay cambio de vegetación de estado primario
1	Pseudoestasis inferior	Integridad alta	No hay cambio de estado primario, si hay cambio de tipo de vegetación pero dentro de la misma ecovariante
2	Pseudoestasis media	Integridad alta	No hay cambio de estado primario, si hay cambio de tipo de vegetación a otra ecovariante pero el estrato dominante ascendiendo en porte (ej. matorral a bosque)
3	Pseudoestasis superior	Integridad alta	No hay cambio de estado primario, si hay cambio de tipo de vegetación a otra ecovariante pero manteniéndose el porte del estrato dominante (ej. boque de encino a selva baja)
4	Degradación muy somera	Integridad media	No hay cambio de estado primario, si hay cambio de tipo de vegetación a otra ecovariante pero el estrato dominante desceinde en porte (ej. matorral a pastizal o herbazal)
5	Degradación somera inferior	Integridad media	Cambio de bosques o selvas primarias a vegetación secundaria arbórea
6	Degradación somera superior	Integridad media	Cambio de cualquier tipo de vegetación primaria a bosque inducido o plantación forestal
7	Degradación moderada inferior	Integridad media	Cambio de bosques, selvas o matorrales primarios a vegetación secundaria arbustiva
8	Degradación moderada media inferior	Integridad media	Cambio de cualquier tipo de vegetación primaria a vegetación sabanoide
9	Degradación moderada media superior	Integridad baja	Cambio de bosques , selvsa, matorrales, herbazales o pastizales primarios a vegetación secundaria herbacea

	IIE (19 categorías)	IIE (4 Categorías)	Descripción
10	Degradación moderada superior	Integridad baja	Cambio de cualquier tipo de vegetación primaria a palmar inducido o pastizal inducido o pastizal cultivado
11	Degradación severa inferior	Integridad baja	Cambio de cuerpo de agua a uso acuícola
12	Degradación severa media inferior	Integridad baja	Cambio de cualquier tipo de vegetación primaria a agricultura de temporal
13	Degradación severa media superior	Integridad baja	Cambio de cualquier tipo de vegetación primaria a agricultura de riego
14	Degradación severa superior	Integridad baja	Cambio de cualquier tipo de vegetación primaria a agricultura de humedad
15	Degradación muy severa inferior	Integridad muy baja	Cambio de cualquier tipo de vegetación primaria a cuerpo de agua
16	Degradación muy severa media inferior	Integridad muy baja	Cambio de cualquier tipo de vegetación primaria a área sin vegetación aparente o desprovista de vegetación
17	Degradación muy severa media superior	Integridad muy baja	Cambio de cualquier tipo de vegetación primaria a asentamientos humanos
18	Degradación muy severa superior	Integridad muy baja	Cambio de cualquier tipo de vegetación primaria a zona urbana

Variables

Las variables utilizadas para la predicción del IIE provienen de datos de percepción remota (satelitales), estos tienen como ventaja que cubren todo el territorio mexicano, están disponibles para distintos periodos, son de uso público, y se puede acceder a ellos casi en tiempo real, por lo que permiten una estimación del IIE actual e histórica.

Variables predictoras Con el fin de describir la estructura de la vegetación se incluyeron en el modelo datos de radar de apertura sintética de banda C, provenientes del satélite de la misión Sentinel-1. Este radar envía ondas a la Tierra que al rebotar de regreso permiten formar una imagen con información del terreno, incluso en condiciones de oscuridad, de lluvia o neblina. Las ondas de radar interactúan de manera distinta de acuerdo a la rugosidad, humedad y propiedades dieléctricas de la superficie, por lo que estos datos se han utilizado previamente para el estudio del suelo y la vegetación, por ejemplo, para la predicción de la altura del dosel (Nandy, Srinet, and Padalia 2021). Otra variable incluida para describir la vegetación, fue la productividad primaria bruta, que es la energía total fijada por fotosíntesis por las plantas.

También se consideró la cobertura terrestre, que identifica 17 clases, entre ellas cultivos y terrenos urbanos, los cuales están directamente relacionados a una integridad baja o muy baja. Por último, se consideró la distancia al borde del parche, ésta es la distancia de cada pixel al cultivo, terreno urbano o camino más cercano, suponiendo que entre mayor distancia exista a áreas con actividad humana mayor será la integridad.

Variables de control Para estimar el IIE, debemos considerar que México cuenta con distintos tipos de ecosistemas con diversas condiciones bioclimáticas, por lo que se incluyó como variable de control la clasificación propuestas por Holdridge, que utiliza la precipitación, la biotemperatura (temperatura promedio en la que se produce el crecimiento de las plantas) y la evapotranspiración (cantidad de agua que se evapora o transpira a través de las plantas) para clasificar el territorio en 31 zonas de vida. Además, se incluyó la elevación, pues de esta también depende el tipo de vegetación.

Fuente de datos	Definición	Variables	Resolución de origen (m)	Transformación a resolución de 250m	Fuente
Zona de vida de Holdridge	Agrupación en 28 zonas basándose en precipitación, biotemperatura y la evapotranspiración	Zona de vida de Holdridge	260	Interpolación con Nearest Neighbor	Portal de Geoinformación, CONABIO
Elevación (DEM)	Altura sobre nivel promedio del mar	Promedio	30	Promedio	DEM GLO-30, Copernicus
Productividad primaria bruta	Energía total fijada por fotosíntesis por las plantas	Mínimo Máximo Promedio	500	Mínimo Máximo Interpolación bilinear	Terra Gross Primary Productivity, NASA LP DAAC
Radar de apertura sintética en la banda C	Coeficiente de retrodispersión que depende de la geometría del terreno y sus características electromagnéticas	Desviación estándar Promedio de banda VH y VV	40	Promedio	Sentinel-1, Copernicus Sentinel data
		Desviación estándar de banda VH y VV Métricas de textura (entropía y promedio) de banda VH y VV			

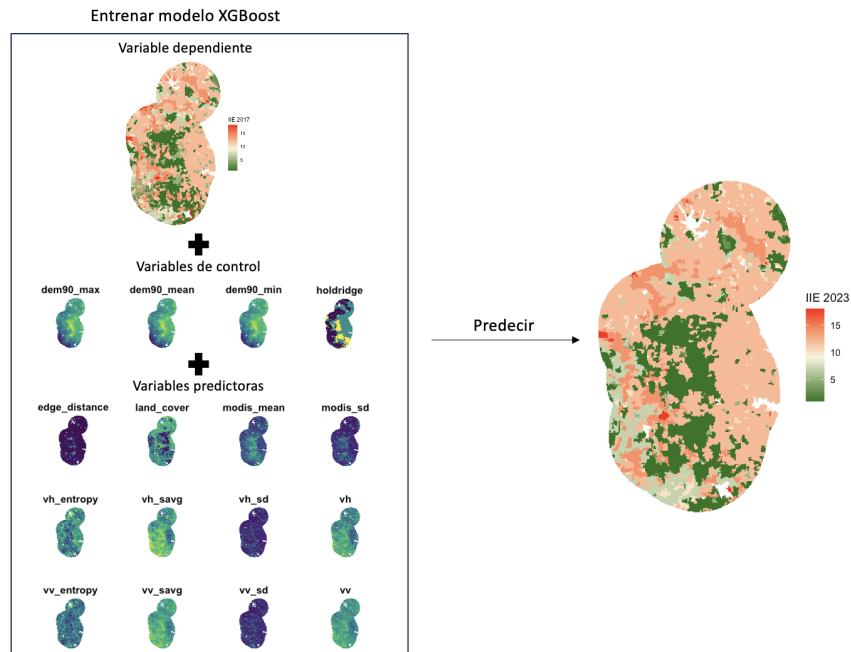
Fuente de datos	Definición	Variables	Resolución de origen (m)	Transformación a resolución de 250m	Fuente
Cobertura del suelo	Estimación basada en imágenes satelitales del tipo de uso de suelo, que incluye cultivos, asentamientos urbanos, bosques, etc.	Uso de suelo	500	Interpolación con Nearest Neighbor	MODIS Land Cover Type, NASA LP DAAC
Distancia al borde	Distancia, en metros, de cada pixel al borde del parche	Distancia al borde	250	-	MODIS Land Cover Type, NASA LP DAAC. Red Nacional de Caminos, INEGI

Modelaje

Para entrenar el modelo XGBoost se tomó el IIE de 2017 debido a que el mapa de uso de suelo y vegetación (Instituto Nacional de Estadística Geografía e Informática 2021a), con el que éste se genera, es de 2017. Se recolectaron los datos para estimar las variables predictoras y de control para este año. Si la fuente de datos está ligada a variaciones estacionales, por ejemplo, la productividad primaria bruta, se tomaron los datos de un sólo mes al año (marzo), con el fin de que estos tengan menor variabilidad. Los mapas de cada una de las variables, así como el del IIE de 2017, fueron proyectados a la misma extensión, sistema de coordenadas (EPSG) y resolución, obteniendo mapas con píxeles de 250m x 250m.

Con el fin de tomar en cuenta la autocorrelación espacial, se aplicó el algoritmo SLIC (Simple Linear Iterative Clustering), que crea agrupaciones de píxeles, llamadas *superpíxeles*, con características similares de acuerdo a las variables deseadas. En este caso, se utilizaron aquellas que representan la estructura de la vegetación, es decir datos de radar y de fotosíntesis, en específico el promedio de las bandas VV y VH, así como la productividad primaria bruta media y su desviación estándar. Se transformó cada uno de los mapas de variables y el del IIE a estos *superpíxeles*.

Para relacionar las variables con el IIE, se entrenó un modelo XGBoost. Con este modelo se obtiene la probabilidad de que cada *superpixel* pertenezca a cada clase del IIE, asignando el valor con la mayor probabilidad. Se puede predecir el IIE para el periodo y territorio deseado utilizando el modelo entrenado con datos del 2017, mediante las variables.



Validación

La validación del índice de IE se realizó mediante datos externos, que a diferencia de las variables predictoras, provenientes de imágenes satelitales, no tienen amplia cobertura. La validación fue con dos tipos de información: el conocimiento de expertos en el territorio y datos de fauna.

Validación de expertos Se acudió al personal que labora dentro de las áreas naturales protegidas (AP), quienes cuentan con conocimiento del estado del territorio conservado. Se elaboró un cuestionario, donde cada uno de los expertos marcó sobre un mapa el estado en el que considera se encuentra el territorio del que él o ella tenga conocimiento, esto se realizó eligiendo una de cuatro categorías en celdas de 1 km x 1 km de una rejilla sobre el mapa. Las categorías son:

- IE Alta: Celda sin indicios de impacto humano aparente, dominio de bosques o selvas maduros (+50 años), habitan la mayoría de las especies de mamíferos silvestres de gran tamaño (+10 kg), y aves terrestres (+1 kg) y/o dominan especies en alguna categoría de riesgo/endémicas. Sin indicios de fauna exótica (vacas, perros, gatos, ovejas y cabras).
- IE Media: Celda con impacto humano moderado, vegetación secundaria dominante, con alta posibilidad de recuperación y restauración de la cobertura vegetal. Menor número de especies de mamíferos de gran tamaño (+10kg) y de aves terrestres (+ 1 kg) al esperado en condición prístina, o menor incidencia de cada especie. Sin agricultura o con escasa presencia de milpas, sin plantaciones forestales. Incidencia moderada de especies de flora y fauna exótica.
- IE Baja: Celda con impacto humano evidente, sin capacidad de recuperación de la cobertura vegetal original, presencia de monocultivos, el suelo no está severamente erosionado (ej. acahuales). Cercanía zonas urbanas o hay escasos asentamientos humanos (dispersos). Dominancia de especies de flora y fauna exótica.
- IE Muy Baja: Celda dominada por impactos humanos, dominancia de monocultivos, de asentamientos humanos (rurales o urbano), presencia de minería, dominancia de ganadería intensiva, suelo transformado con respecto al original o erosionado. Dominancia de especies de flora y fauna exótica.

Previo a responder el cuestionario, se les explicó a los expertos el concepto de IE y cada una de las categorías. Una vez obtenida la respuesta de los expertos, se eligieron las celdas que tuvieron consenso. Se considera que existió consenso si más de la mitad (la mitad más uno, siempre y cuando haya respondido más de un experto) coincidieron en la misma respuesta para una celda (excluyendo a los expertos que no asignaron respuesta a esa celda). Si sólo un experto respondió en alguna AP, ésta será considerada como si hubiera consenso.

Se calculó en qué proporción coincide las respuestas con consenso de los expertos y el valor del IIE de 2023. Otra manera de comparar estos valores, es fijándonos únicamente en el orden que asignaron los expertos, por ejemplo, si éstos opinaron que el área B de la AP tiene mayor integridad que un área A y menor integridad que un área C, es decir, $IE_{Expertos}(A) < IE_{Expertos}(B) < IE_{Expertos}(C)$, entonces el IIE 2023 debería mantener este mismo orden, para lo cual, comparamos el valor del IIE 2023 de un pixel en B, contra la mediana del IIE 2023 en el área A y B. Si está bien ordenado debe cumplir $mediana(IIE_{2023})(A) < IIE_{2023}(pixel) \leq mediana(IIE_{2023})(C)$.

Para ello seguimos los siguientes pasos por cada AP:

1. Dividir la AP en áreas que tengan el mismo valor de la respuesta de expertos con consenso, por ejemplo A, B, C.
2. Ordenar estas áreas de menor degradación a mayor:

$$IE_{Expertos}(A) < IE_{Expertos}(B) < IE_{Expertos}(C)$$

3. Calcular la mediana del IIE 2023 por área:

$$mediana(IIE_{2023})(A), mediana(IIE_{2023})(B) \text{ y } mediana(IIE_{2023})(C).$$

4. Para cada pixel i perteneciente al área B, si su valor de IIE 2023 es menor a la mediana del IIE 2023 del área C y mayor a la del área A, decimos que está ordenado de manera correcta:

$$o_i = \begin{cases} 1 & mediana(IIE_{2023})(A) < IIE_{2023}(i) \leq mediana(IIE_{2023})(C) \\ 0 & eoc \end{cases}$$

Validación mediante datos de Fauna La fauna tiene diferentes características que los hacen más o menos vulnerables al mismo tiempo pueden ser indicadores de la salud del ecosistema, debido al impacto que su biomasa causa sobre la remoción de nutrientes y uso de recursos, o al papel que juega en el ecosistema. Es por ello que datos de fauna fueron utilizados para la validación del IIE, estos provienen de dos fuentes: registros del SNIB (Sistema Nacional de Información sobre Biodiversidad de México) y de fototruampas de SiPeCaM (Sitios Permanentes de Calibración y Monitoreo de la Biodiversidad).

De acuerdo a la vulnerabilidad a la cacería y exposición al humano, se seleccionó un conjunto de registros provenientes del SNIB de mamíferos y aves de talla grande para la validación de clases altas de IIE, los cuales incluyeron: *Canis latrans*, *Canis lupus*, *Caracara plancus*, *Cathartes aura*, *Coragyps atratus*, *Dicotyles tajacu*, *Leopardus pardalis*, *Lontra longicaudis*, *Lynx rufus*, *Meleagris gallopavo*, *Odocoileus virginianus*, *Puma concolor*.

Por su parte, la incidencia de ganadería se utilizó para la evaluación de IE baja y muy baja.

Por otro lado, se cuenta con observaciones de especies derivado del monitoreo con fototruampas de SiPeCaM. Los datos fueron recabados siguiendo un diseño de muestreo, donde se colocaron las cámaras en unidades que cuentan con las mismas características bioclimáticas, con la diferencia que algunas zonas son íntegras y otras degradadas. Por lo que se pudo llevar a cabo pruebas estadísticas para saber si cierta especie fue observada con mayor frecuencia en zonas íntegras o degradadas, validando el IIE. Con estos datos se tiene una representatividad de en al menos 3 de los 7 ecosistemas del país.

Resultados

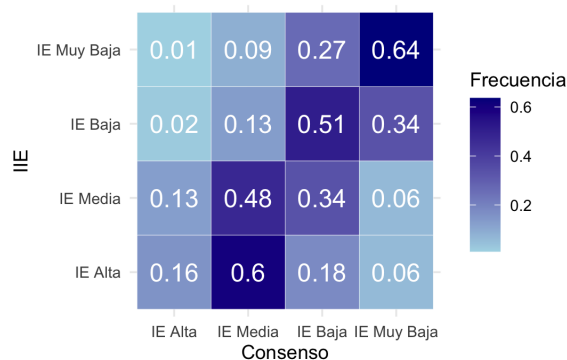
Validación

Validación de expertos Se realizó el cuestionario a 41 expertos, abarcando 19 AP, con un promedio de 2.2 (DE=1.1) expertos por cada AP.

El 57.6% de las celdas donde hubo dos o más respuestas tuvieron consenso. El mapa de consenso, conformado por las celdas que tuvieron consenso y por aquellas que pertenecen a una AP donde sólo participó un experto, es el 69.7% del área total que abarcó el cuestionario y corresponde al 0.33% del total del territorio que abarca el IIE.

Del total de celdas con consenso, el 30.25% coinciden con el valor del IIE. De las que no coinciden, en el 91.0% los expertos asignaron mayor degradación de la que se predice con el IIE. Al respecto, uno de los resultados que sobresalen es que, a pesar del alto impacto humano en el centro del país, el 57.6% del IIE 2023 tiene valores de Integridad Alta en el territorio que abarca el mapa de consenso, en contraste, sólo el 11.4% presentó Integridad Alta de acuerdo a los expertos.

En la siguiente imagen podemos observar la distribución de las respuestas con consenso de cada categoría del IIE. Como podemos ver en la categoría IE Alta es donde existe una mayor discrepancia con las respuestas de los expertos, pues estos consideraron que solo el 16% de las celdas con IIE 2023 con valor IE Alta tienen realmente este valor, más bien, creen que la mayoría de estas celdas (60%) tienen un valor de IE Media. Esto podría explicarse por la subjetividad del cuestionario, pues un experto puede dudar en que cualquier zona tenga IE Alta y prefiere reservarse diciendo que tiene IE Media. El 29.3% de expertos no consideraron que existan zonas con Integridad Alta dentro de la AP en la que laboran, lo cual es poco probable, ya que son zonas protegidas.

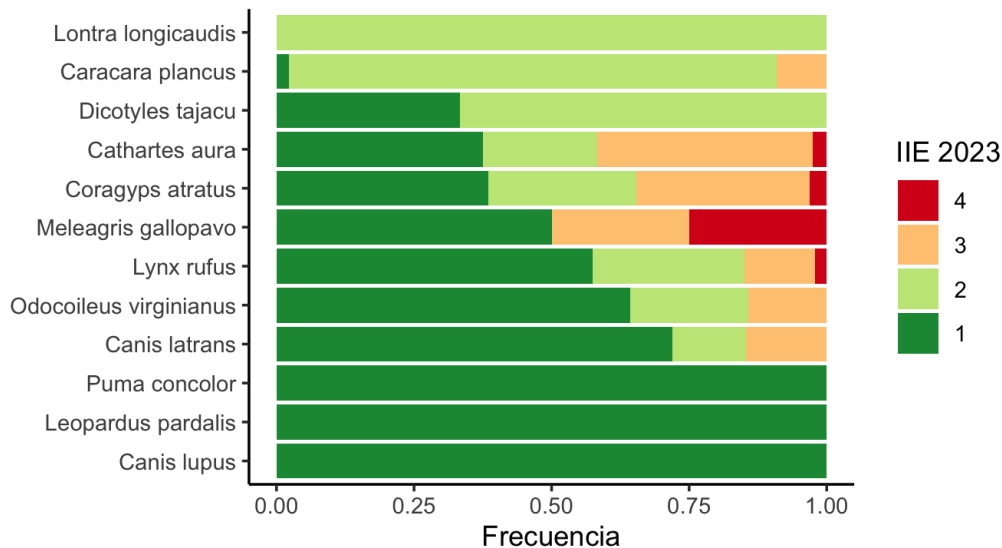


Cuando dividimos por AP, se tuvo una coincidencia promedio de 36.0% (DE=26.9%). Las AP con mayor coincidencia entre el consenso de expertos y el IIE 2023, fueron la Laguna de Zempoala (100%), Cumbres de Ajusco (68.9%), Insurgentes Miguel Hidalgo y Costilla (66.7%), y El Veladero (55.6%). Y las AP con menor coincidencia fueron Grutas de Cacahuamilpa, Desierto de los Leones y El Chico, donde no hubo coincidencias y Mariposa Monarca con 3.5%.

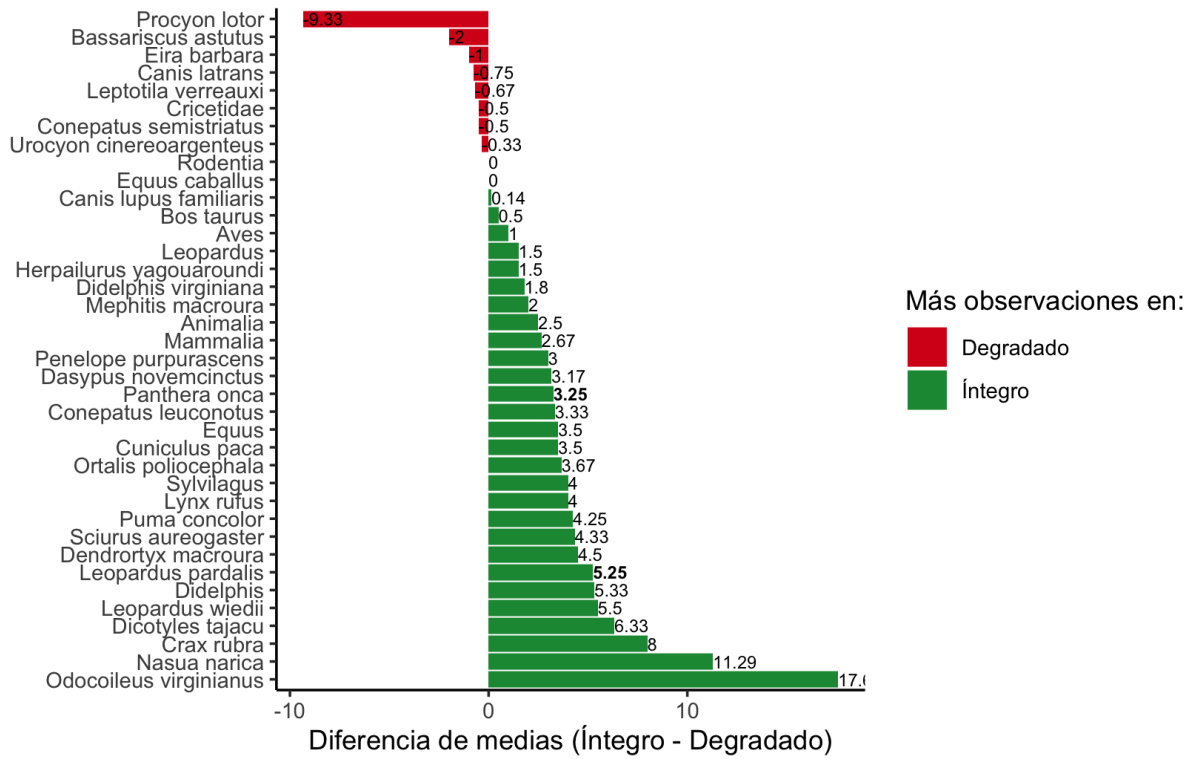
En cuanto a la comparación del orden de las categorías, el 39.36% del total coinciden en el orden del IIE 2023 y las respuestas con consenso de expertos. Considerando por AP, en promedio tuvieron una coincidencia en el orden del 53.4% (DE=23.8%), siendo las de mayor coincidencia Barranca de Metztitlán (92.4%), La Montaña Malinche o Matlalcuéyatl (88.7%), El Veladero (76.3%) y Lagunas de Zempoala (75.2%). Las de menor coincidencia fueron Sierra Gorda de Guanajuato (12.9%), Sierra de Huautla (18.7%), Mariposa Monarca (22.7%) e Iztaccíhuatl-Popocatepetl (32.6%).

Si los expertos están en lo correcto parecería que el modelo de IIE tiende a sobrestimar la Integridad de los Ecosistemas en la categoría de IE Alta, pues los expertos consideran que en estas zonas la IE es menor, en su mayoría IE Media.

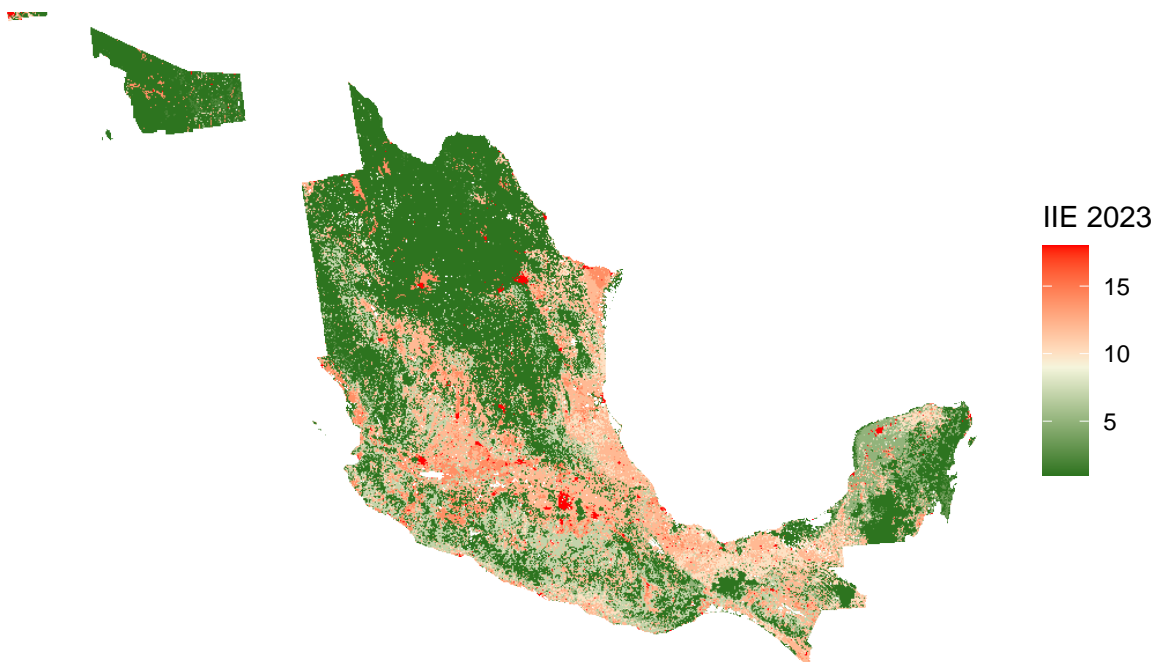
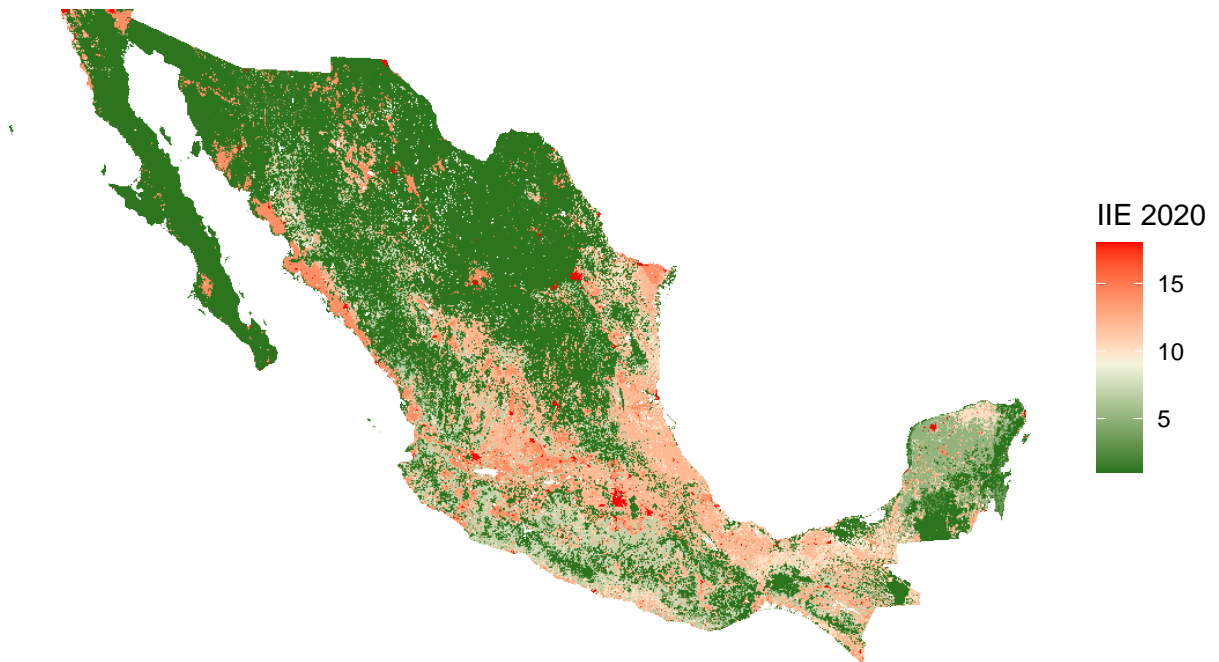
Validación mediante datos de Fauna Los resultados de los datos obtenidos del SNIB de observaciones de mamíferos y aves de talla grande se pueden ver en la siguiente gráfica, la cual muestra que estas especies se encuentran en mayor proporción en áreas Integra de acuerdo al IIE 2023, ya sea en áreas con IE Alta o IE Media, a excepción del *Meleagris gallopavo*.

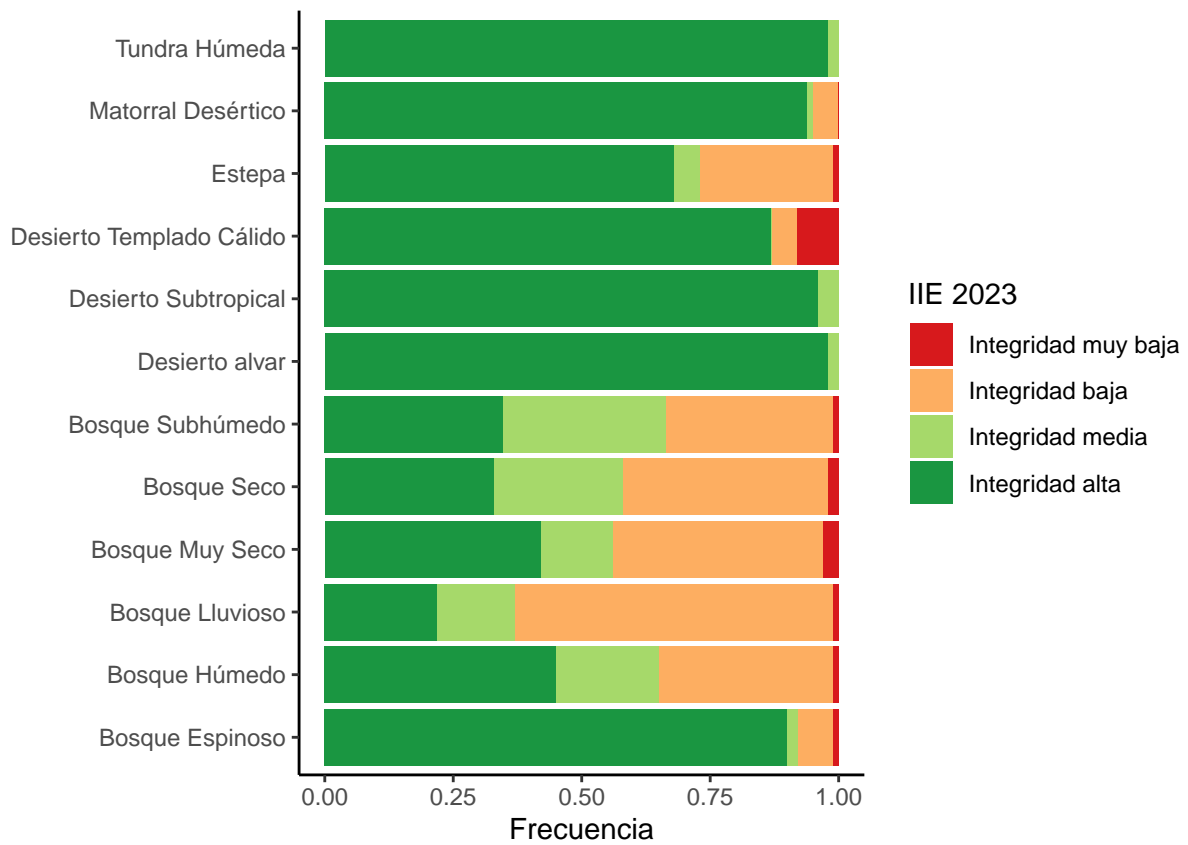


Los datos de SiPeCaM, procesados hasta el momento, muestran que las especies *Leopardus pardalis* y *Panthera onca* se encuentran con mayor frecuencia, de manera estadísticamente significativa, en zonas donde el IIE 2023 tiene un valor de Integridad Alta, lo cual es válido para estas especies. Por otra parte, aunque no existe significancia estadística hay evidencia de que el *Procyon lotor* se observó en áreas donde IIE 2023 muestra degradación, lo cual se esperaría de esta especie.

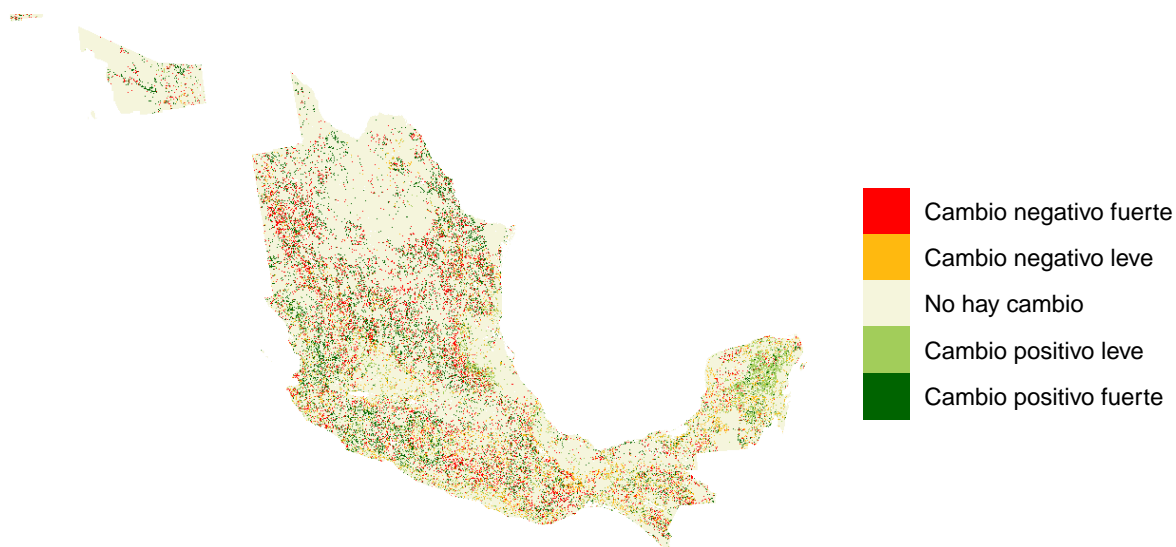


IIE





Cambio en el IIE



Referencias

- Di Marco, Moreno, Simon Ferrier, Tom D Harwood, Andrew J Hoskins, and James EM Watson. 2019. "Wilderness Areas Halve the Extinction Risk of Terrestrial Biodiversity." *Nature* 573 (7775): 582–85.
- Equihua, Miguel, Octavio Pérez-Maqueo, Julián Equihua, Pedro Maeda, Michael Schmidt, Melanie Kolb, Nashieli Garcia Alaniz, et al. 2024. "A Big Data and Machine Learning Approach for Monitoring the Condition of Ecosystems."
- Hansen, Andrew J, Benjamin P Noble, Jaris Veneros, Alyson East, Scott J Goetz, Christina Supples, James EM Watson, et al. 2021. "Toward Monitoring Forest Ecosystem Integrity Within the Post-2020 Global Biodiversity Framework." *Conservation Letters* 14 (4): e12822.
- Instituto Nacional de Estadística Geografía e Informática. 2003. "Conjunto de Datos Vectoriales de La Carta de Vegetación Primaria. Escala 1: 1000000 (Niveles I y II)." INEGI.
- . 2021a. "Conjunto de Datos Vectoriales de La Carta de Uso Del Suelo y Vegetación. Escala 1:250 000. Serie VII. Conjunto Nacional." INEGI.
- . 2021b. "Cuentas de Condición de Los Ecosistemas Terrestres de México. Cuentas de Los Ecosistemas de México. Resultados Del Proyecto Natural Capital Accounting and Valuation of Ecosystem Services (NCAVES)." INEGI.
- Martin, Tara G, and James EM Watson. 2016. "Intact Ecosystems Provide Best Defence Against Climate Change." *Nature Climate Change* 6 (2): 122–24.
- Nandy, Subrata, Ritika Srinet, and Hitendra Padalia. 2021. "Mapping Forest Height and Aboveground Biomass by Integrating ICESat-2, Sentinel-1 and Sentinel-2 Data Using Random Forest Algorithm in Northwest Himalayan Foothills of India." *Geophysical Research Letters* 48 (14): e2021GL093799.
- Watson, James EM, Tom Evans, Oscar Venter, Brooke Williams, Ayesha Tulloch, Claire Stewart, Ian Thompson, et al. 2018. "The Exceptional Value of Intact Forest Ecosystems." *Nature Ecology & Evolution* 2

(4): 599–610.

Wildlife Conservation Society. 2020. “WCS - FAQ: Ecosystem Integrity in the Post-2020 Global Biodiversity Framework.” https://cdn.wcs.org/2020/02/21/6iq1kmonow_Integrity_FAQ_2.2_21_February_2020_.pdf; WCS.