



CONABIO

COMISIÓN NACIONAL PARA EL
CONOCIMIENTO Y USO DE LA BIODIVERSIDAD

Comisión Nacional para el Conocimiento y Uso de la Biodiversidad

SNIB-Ejemplares

Reporte de calidad de bases de datos financiadas

COMISIÓN NACIONAL PARA EL CONOCIMIENTO Y USO DE LA BIODIVERSIDAD

Liga Periférico-Insurgentes Sur No. 4903, Parques del Pedregal, Tlalpan 14010, México, D. F.

<https://www.gob.mx/conabio>

<http://www.biodiversidad.gob.mx>

<http://geoportal.conabio.gob.mx>

<http://www.snib.mx>

<http://enciclovida.mx>

URL persistente: <http://www.snib.mx/ejemplares/docs/CONABIO-SNIB-DescripcionReporteCalidadSNIB.pdf>

Forma de citar:

CONABIO. 2018. Sistema Nacional de Información sobre Biodiversidad (SNIB) - Ejemplares. Reporte de calidad de bases de datos financiadas, versión 2018-11. México. Disponible en <<http://www.snib.mx/ejemplares/docs/CONABIO-SNIB-DescripcionReporteCalidadSNIB.pdf>>

Licencia:

Este documento se publica y distribuye bajo una licencia [Creative Commons Atribución 4.0 Internacional](#) compatible con la licencia LibreusoMX de datos.gob.mx



Contenido

Contenido	3
SNIB-Ejemplares. Reporte de calidad de bases de datos financiadas	4
Introducción	4
Resumen de control de calidad	4
Metadatos	4
Posibles errores en los datos	5
Validación de datos que realiza la CONABIO	5
Validación Taxonómica	5
Validación Geográfica	6
Literatura citada	7
Anexo	9

SNIB-Ejemplares. Reporte de calidad de base de datos financiadas

Introducción

La CONABIO ha sido pionera en desarrollar desde sus inicios procesos de control de calidad en las bases de datos, principalmente para responder a la necesidad de contar con un Sistema Nacional de Información sobre Biodiversidad (SNIB) que permitiera integrar datos e información de diferentes fuentes incluso, con diferentes propósitos y enfoques de investigación. El reporte de control de calidad, representa un resumen sucinto de la calidad de información que tiene la base de datos del proyecto que ha financiado la CONABIO. Para dicho reporte se han considerado los principales aspectos del contenido de la base de datos estructurados conforme a las capas de información que se describen en el protocolo de control de calidad (CONABIO, 2018).

El reporte de control de calidad, está conformado por las siguientes secciones:

Resumen del control de calidad. Refiere a los datos de la autoría del proyecto desarrollado o [financiado](#) por la CONABIO en el que se generó la base de datos, así como a los datos relativos a su publicación:

	Descripción
Resumen del control de calidad	Título del Proyecto.
Responsable de la base de datos	Nombre completo del curador de la base de datos
Fecha de elaboración	día, mes y año en que se elaboró el reporte de calidad
Versión SNIB	Año – mes en que se publicó la base de datos en el Sistema Nacional de Información sobre Biodiversidad (SNIB)
Año de conclusión de la base de datos	Año en el que concluyó la base de datos
Modelo y diccionario de datos	Vínculo a los documentos del diagrama entidad-relación y al diccionario de datos del SNIB
Versión SNIB	Vínculo al documento que explica el proceso de revisión, validación, limpieza y estandarización de datos realizado por la CONABIO

Metadatos. Cada base de datos publicada cuenta con sus metadatos. En este apartado se vincula al reporte donde el usuario puede ver con detalle los metadatos, con el propósito de que lo tenga presente al hacer uso de los datos (ver [ejemplo](#)).

Posibles errores en los datos. Los posibles errores, se muestran en un cuadro con seis columnas:

Capa de información. Corresponde al primer nivel de integración de los posibles errores considerando el concepto integrador de los datos primarios: 'individuo - nombre científico – georreferencia – fecha'. Las capas de información consideradas son las siguientes

- Ejemplar. Datos de la curaduría del ejemplar.
- Taxonómica-biogeográfica. Datos nomenclaturales (nombre científico y ubicación taxonómica de la especie).
- Geográfica. Datos de la ubicación geoespacial del sitio de recolecta u observación del ejemplar.

Tipos de error. Se consideran principalmente los tipos de error que tienen relación con la ausencia de información (error de omisión), redundancia de información (repetición de datos) y, los errores de congruencia (relación incorrecta o incongruente entre el conjunto de datos de un registro).

Descripción del error. Se describe el tipo de error encontrado en los datos.

No. de registros con error. Se indica el número de registros que presentan el error. Al dar clic sobre el dato numérico, se muestra un archivo en formato .csv con los registros erróneos.

Total de registros. Se indica el número total de registros, por capa de información.

Porcentaje de registros con error. Se muestra el porcentaje de registros en la base de datos con el error.

Validación de datos que realiza la CONABIO. Corresponde a un resumen del proceso de validación que realiza la CONABIO para la publicación de la información en el Sistema Nacional de Información sobre Biodiversidad (SNIB), disponible en <<http://www.snib.mx/documents/docsnib.html>>.

Validación taxonómica. El resultado de la validación taxonómica, se muestra en un cuadro con cinco columnas:

Descripción. Se describe el error encontrado en los nombres científicos contenidos en la base de datos al compararlos con los catálogos de autoridades taxonómicas (CONABIO, 2014)

Número de nombres. Se indica el número de registros (nombres científicos) que presentan el error. Al dar clic sobre el dato numérico, se muestra un archivo en formato .csv con los registros erróneos.

Total de nombres en la base de datos. Se indica el número total de registros (nombres científicos) contenidos en la base de datos.

Número de ejemplares. Se indica el número total de registros de ejemplares que tienen el error.

Total de ejemplares en la base de datos. Se indica el número total de registros de ejemplares (especímenes) contenidos en la base de datos.

Validación geográfica. El resultado de la validación geográfica, se muestra en un cuadro con cinco columnas:

Descripción. Se describe el error encontrado en la información geoespacial contenida en la base de datos.

Número de sitios. Se indica el número de sitios (coordenada x, y) que presentan el error. Al dar clic sobre el dato numérico, se muestra un archivo en formato .csv con los registros erróneos.

Total de sitios en la base de datos. Se indica el número total de registros (sitios) contenidos en la base de datos.

Número de ejemplares. Se indica el número total de registros de ejemplares que tienen el error.

Total de ejemplares en la base de datos. Se indica el número total de registros de ejemplares (especímenes) contenidos en la base de datos.

Literatura citada

CONABIO (comp.) 2014. *Catálogo de autoridades taxonómicas de especies con distribución en México. Bases de datos del Sistema Nacional de Información sobre Biodiversidad*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, en <<http://www.biodiversidad.gob.mx/especies/CAT.html>>.

CONABIO. 2018. *Protocolo de control de calidad. Vol. I: Datos primarios del Sistema Nacional de Información sobre Biodiversidad (SNIB), características y reglas*. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México.

Anexo

EJEMPLO DE REPORTE DE CALIDAD DE BASES DE DATOS DEL PROYECTO “ANFIBIOS Y REPTILES DE LAS REGIONES 30 (PARQUE NATURAL SIERRA MADERAS DEL CARMEN); 31 (SIERRA EL BURRO-RÍO SAN RODRIGO); 36 (SIERRA LA ENCANTADA); Y 37 (SIERRA EL PINO), COAHUILA”

Resumen del control de calidad aplicado a la base de datos del proyecto:

Anfibios y reptiles de las regiones 30 (Parque Natural Sierra Maderas del Carmen); 31 (Sierra El Burro-Río San Rodrigo); 36 (Sierra La Encantada); y 37 (Sierra El Pino), Coahuila

Responsable del proyecto: Dr. Julio Alberto Lemos Espinal

Fecha de elaboración: 23 de junio de 2018

Versión SNIB: 2017-12



Ver metadatos de la base de datos:

Capa de información	Tipo de error	Descripción error	No de registros con error	Total de registros	Porcentaje de registros con error
Ejemplar (curatorial)	Omisión	Ejemplar sin fecha de recolecta u observación	<u>3,701</u>	50,903	7.27%
		Ejemplar sin dato del recolector	<u>6,985</u>		13.72%
		Ejemplar reportado sin dato del recolector u observador	<u>60</u>		0.12%
		Ejemplar sin dato de la colección	<u>64</u>		0.13%
		Ejemplar recolectado sin número de catálogo	<u>5</u>		0.01%
		Ejemplar reportado sin número de catálogo	<u>64</u>		0.13%
		Ejemplar recolectado sin número de recolecta	<u>22,285</u>		43.78%
		Ejemplar observado sin número de observación	<u>15,424</u>		30.3%
		Ejemplar reportado sin número de recolecta	<u>64</u>		0.13%
		Ejemplar sin dato en tipo nomenclatural	<u>9</u>		0.02%
		Ejemplar sin fecha de determinación	<u>18,390</u>		36.13%
		Ejemplar recolectado sin determinador	<u>12,130</u>		23.83%
		Ejemplar reportado sin determinador	<u>60</u>		0.12%
		Ejemplar sin tipo de vegetación	<u>9,961</u>		19.57%
		Ejemplar sin altitud o profundidad	<u>41,041</u>		80.63%
		Ejemplar sin dato de tipo de preparación	<u>84</u>		0.17%
		Ejemplar sin dato de hábitat	<u>22,963</u>		45.11%
		Ejemplar sin dato de método de recolecta	<u>50,903</u>		100%
		Ejemplar sin dato de edad	<u>50,336</u>		98.89%
		Ejemplar sin dato de sexo	<u>50,264</u>		98.74%
Ejemplar sin dato de hora de recolecta u observación	<u>50,903</u>	100%			

Capa de información	Tipo de error	Descripción error	No de registros con error	Total de registros	Porcentaje de registros con error
	Redundancia	Ejemplares con el mismo número de recolecta y recolector	<u>2</u>	274	0%
	Congruencia	Ejemplar observado con dato de número de catálogo	<u>15,424</u>		30.3%
		Ejemplar observado con dato de tipo de preparación	<u>245</u>		0.48%
		Ejemplar observado con diferente fecha de observación y de determinación	<u>6,505</u>		12.78%
		Ejemplar con ambiente marino y con valor de profundidad positivo	<u>1</u>		0%
		Ejemplar con recolectores u observadores con intervalo de recolectas u observaciones mayor o igual a 40 años	<u>19,414</u>		38.14%
		Ejemplar con grupo de recolecta con intervalo de recolectas mayor o igual a 40 años	<u>4,868</u>		9.56%
		Ejemplar con grupo de determinación con intervalo de determinaciones mayor o igual a 40 años	<u>1,929</u>		3.79%
Taxonómica - biogeográfica	Omisión	Taxón sin dato en el sistema de clasificación, catálogo de autoridad taxonómico, diccionario o listado taxonómico	<u>2</u>	274	0.73%
	Congruencia	Infraespecie típica con autoridad diferente al de la especie	<u>1</u>		0.36%
Geográfica	Omisión	Coordenada geográfica sin datum	<u>1,420</u>	1,515	93.73%
	Congruencia	Misma coordenada geográfica con diferentes estados o regiones administrativas equivalentes	<u>1</u>		0.07%
		Misma coordenada geográfica con diferentes municipios o regiones administrativas equivalentes	<u>1</u>		0.07%

Validación de datos que realizó la CONABIO¹

Validación taxonómica				
Descripción	Número de nombres	Total de nombres en la base de datos	Número de ejemplares	Total de ejemplares en la base de datos
Nombres de especies no encontradas en los Catálogos de Autoridades Taxonómicas (CAT) de la CONABIO o en fuentes complementarias	0	257	0	50,903
Nombres de especies sin la relación al válido o correcto a partir de los Catálogos de Autoridades Taxonómicas (CAT) de la CONABIO o en fuentes complementarias	0		0	

Control de calidad de la base de datos del proyecto JF001

Validación geográfica				
Descripción	Número de sitios	Total de sitios en la base de datos	Número de ejemplares	Total de ejemplares en la base de datos
Información geográfica (país-estado-municipio-localidad-coordenada) no válida	<u>2</u>	1,695	<u>65</u>	50,903

¹Para mayor información del proceso de validación que realiza la CONABIO, puede consultar el documento **Versión SNIB** disponible en <http://www.snib.mx/ejemplares/docs/>

Agradecemos que cualquier duda o comentario relacionado con este documento, sea enviado a la Subcoordinación de Inventarios Bióticos (SIB) mediante el correo electrónico sib@conabio.gob.mx.