

INTRODUCCIÓN

La georreferenciación de localidades consiste en la asignación de coordenadas geográficas a la descripción textual de un sitio (Proctor, Blum, Chaplin, 2001; Wieczorek, 2001). El presente texto tiene como objetivo proporcionar herramientas técnicas sistematizadas para realizar la georreferenciación de las descripciones de las localidades en donde se hizo alguna colecta biológica.

La Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (Conabio) se ha dado la tarea de asignar coordenadas geográficas a las localidades contenidas en las bases de datos de las colecciones biológicas que carezcan de ellas y que sean de interés a esta institución.

La implementación de este proceso se realizó considerando el enorme volumen de registros en las colecciones que incluyen únicamente una descripción textual de la ubicación de una localidad. Las coordenadas geográficas de las colectas de especímenes constituyen la base de cualquier tipo de análisis espacial.

El área de georreferenciación de Conabio adoptó la metodología establecida por Wieczorek (2001), la cual se desarrolló dentro del proyecto '*Mammal Networked Information System*' (MaNIS) realizado por el museo de vertebrados de la Universidad de Berkeley. Esta metodología fue revisada de manera detallada con el propósito de adecuar los procedimientos, el material cartográfico y el nomenclatural específico a una serie de procesos sistematizados que permitieran establecer en Conabio la metodología más adecuada. El presente documento constituye propiamente un manual de procedimientos para la georreferenciación de localidades.

El proceso de georreferenciación está constituido en su primera etapa por la revisión y tratamiento de la información contenida en las bases de datos, la cual es depurada y organizada previamente a la asignación de coordenadas geográficas. Para la ubicación espacial de los sitios de colecta, cada una de las variables de la metodología establecida por Wieczorek (2001) fue considerada, diseñándose un formulario de

captura con el cual se obtienen las coordenadas de una descripción y además se incorporan los elementos necesarios para calcular la incertidumbre de la georreferenciación. La magnitud de la incertidumbre depende directamente de la precisión de los instrumentos y de los materiales utilizados en la descripción de una ubicación geográfica.

Las herramientas para georreferenciar que se describen técnicamente en este documento son: un formulario de *Microsoft Access*, los sistemas *Arcview* de *ESRI* y *Biótica* de la *Conabio*; también se especifica el manejo y utilización de la cartografía digital e impresa recopilada y procesada en nuestra institución.

Es factible que con los resultados que se vayan obteniendo de la georreferenciación de localidades, como la identificación de rasgos geográficos y la acumulación e identificación de descripciones o coordenadas referentes a un mismo sitio geográfico, se puedan ir creando catálogos que permitan la agilización temporal y formal de la realización de otras georreferenciaciones.

Por todo lo anterior, se tiene la ambiciosa tarea de integrar todos los procesos anteriormente descritos en un solo sistema que permita, a través de una serie de módulos, la determinación de las coordenadas geográficas de una localidad. Una herramienta de este tipo permitiría obtener de manera rápida y precisa la información básica para los estudios de análisis espacial como la distribución potencial de especies, patrones de distribución, estudios biogeográficos, etc.

Al final de este documento se presentan de manera detallada una serie de anexos que constituyen la base técnica de los procedimientos para la georreferenciación, los cuales estarán siendo actualizados de acuerdo a los ajustes que se vayan realizando en la metodología.

Descripción general de la georreferenciación de localidades

La descripción de la ubicación del sitio de colecta de un espécimen es parte esencial de la información asociada a cada colecta, pero su utilización en análisis espaciales a través de los sistemas de información geográfica (SIG) sólo es factible si esta información posee coordenadas.

Actualmente, el uso de los sistemas de posicionamiento global (GPS) ha permitido agregar de manera más frecuente y con –relativamente– mayor precisión el dato de las coordenadas de un sitio de colecta; sin embargo, el volumen de registros que aún permanece sin coordenadas es considerablemente elevado. En términos prácticos esto significa que una enorme cantidad de información permanece desaprovechada, de ahí la importancia de la georreferenciación de localidades.

La descripción de las localidades en donde son recolectados los organismos es realizada de maneras muy distintas, ya que involucra un elemento importante de subjetividad que determina su precisión, pues el colector describe un sitio de acuerdo a los medios que tiene a su disposición y a su propia percepción del lugar. Enseguida se presentan algunos ejemplos:

- Cuatlapán
- 1.3 mi N Acatlán
- 2.3 mi WSW Colima
- 1.9 km NE Puebla-Tlaxcala border on Mexico Hwy. 119
- 10 km N (by road) Canoa, La Malinche
- 1 mi W junction of Gomez Farías Rd. and Mexico Rte. 85
- CIUDAD VICTORIA, 8 MI S, 6 MI W OF; SIERRA MADRE ORIENTAL
- Near Veracruz

La primera etapa del proceso de georreferenciación de localidades consiste en la normalización de las bases de datos que contienen las descripciones. Por lo general, una base de datos siempre contiene información redundante que al ser eliminada,

permite estructurar adecuadamente toda la información y reducir de manera significativa su volumen; debido a esto la normalización es un paso esencial del proceso de georreferenciación.

Este proceso de normalización se realiza sobre una tabla que contenga los campos involucrados con la descripción de localidades, pues generalmente, las tablas poseen también campos relacionados con la taxonomía de los organismos, datos del colector, etc. Este procedimiento se realiza a través de una serie de agrupaciones en Access, y a través de la desagregación de las propias descripciones, en un proceso llamado Atomización.

La segunda etapa del proceso la constituye la asignación de coordenadas, la cual puede realizarse a través de un formulario de captura de Access, cartografía impresa o en un SIG con cartografía digital.

Se diseñó una base de datos de Access con formato de formulario que permite calcular las coordenadas y asignar la incertidumbre a ciertos tipos de descripciones de localidades; esta base tiene asociada toda la información nomenclatural y tiene programados los algoritmos que permiten calcular la incertidumbre de las coordenadas obtenidas.

Para la georreferenciación con mapas impresos se requiere básicamente de una carta topográfica (o cualquier carta temática con una base topográfica), un curvímeter para trazar las distancias sobre el mapa, regla y escuadra. Se ubica la localidad de referencia y se mide la distancia con el curvímeter en la dirección indicada; el trazo de la distancia puede realizarse sobre una carretera o en línea recta (dependiendo de lo que especifique el registro), la coordenada se obtiene utilizando las referencias geográficas de la carta.

Para realizar la georreferenciación con un SIG se requiere contar con la cartografía mínima: división política del país, un nomenclátor de las localidades y las carreteras y caminos rurales; es recomendable poseer cualquier información extra que permita

ampliar las posibilidades de búsqueda de elementos cartográficos. La manera de determinar una coordenada dependerá del SIG utilizado, pero básicamente se realiza haciendo una consulta en el nomenclátor de localidades o de cualquier otro rasgo y trazando la distancia indicada, el sistema proporcionará las coordenadas. En Conabio se utiliza el *software* de ArcView y el de Biótica.

El cálculo de la incertidumbre constituye un elemento esencial en el proceso de georreferenciación, pues asocia a cada par de coordenadas un parámetro suficiente que permite su utilización en análisis científicos. En Conabio se adoptó un método desarrollado por Wiecezorek (2001) el cual describe la manera de cuantificar las distintas fuentes de incertidumbre involucradas en la descripción de un sitio de colecta. Se desarrolló una base de datos de tipo formulario con la que se cuantifica el valor total de la incertidumbre, de manera que cada registro georreferenciado posee un par de coordenadas geográficas y una medida de longitud que determina su precisión.

El proceso de georreferenciación concluye con la validación general de las coordenadas obtenidas y la realización de un reporte.

Clasificación de los tipos de descripciones de localidades

Todas las descripciones de localidades se pueden clasificar en 9 tipos básicos, cada uno de los cuales recibe un tratamiento específico en el proceso de georreferenciación y en la asignación de la incertidumbre.

Tipo	Descripción	Ejemplos
1. Dudosa	El propio registro describe que es dudoso el nombre de la localidad o rasgo de referencia	'Isla Boca Brava?', 'presumably central Chile?'
2. No puede ser ubicada	Los datos de la descripción se perdieron o contienen información que no corresponde con una descripción, o los nombres presentan homónimos que	'localidad no registrada', 'Bob Jones', 'obtenida en laboratorio', 'cumbre', 'San José, México'.

	no pueden ser diferenciados, o no se encuentra la localidad o el rasgo en ninguna fuente disponible	
3. Comprobadamente inexacta	La descripción tiene elementos inconsistentes	'municipio de Poza Rica, Puebla'
4. Coordenadas	La descripción ya posee coordenadas	'42.4535 -84.8429', 'UTM 553160 4077280',
5. Localidad o rasgo	La descripción sólo posee el nombre de un elemento geográfico (una localidad, un cerro, una cueva, isla, etc.)	'Aguascalientes', 'Alice springs', 'unión de la carretera 70 con la 45'
6. Distancia	La descripción consiste en una distancia asociada a un elemento geográfico	'A 5 km de Calgary'
7. Distancia a través de una ruta	Se describe una localidad con una distancia recorrida sobre una carretera, camino, río, etc.	'3.5 km S de Toluca hwy 20'
8. Distancias en direcciones ortogonales	Se utilizan dos distancias trazadas en direcciones ortogonales a partir de un elemento geográfico	'6 km N y 4 km W de Durango'
9. Distancia en una dirección	La descripción contiene una distancia definida en una dirección	'50 km NE de San Luis Potosí'

Tabla 1. Tipos de descripciones de localidades comúnmente encontradas en las colecciones biológicas (Wieczorek, Guo, Hijmans; 2003).

Dentro de estos tipos se encuentran muchas descripciones que presentan características que necesitan ser precisadas con criterios especiales, así que han sido documentados varios de ellos para estandarizar la forma en que se les asignan coordenadas.

La incertidumbre en la georreferenciación

El proceso de asignar una coordenada a una descripción de una localidad implica un margen de error geográfico determinado por el nivel de precisión con el cual es

realizada dicha descripción, este depende directamente de los instrumentos utilizados para obtenerla y de la propia percepción del colector, es decir, ¿cómo se definió una distancia desde un punto?, ¿se utilizó un mapa para ubicar un sitio?, ¿se especificó una distancia inter-cardinal o solamente una cardinal (SW o S)?, ¿se realizó el recorrido sobre una carretera?, etc.

Los criterios y el proceso para determinar la incertidumbre de una coordenada obtenida está ampliamente descrito en el artículo *'The point-radius method for georeferencing locality descriptions and calculating associated uncertainty'* (Wieczorek, Qinghua, Hijmans; 2003 - Anexo 01), a continuación sólo presentaremos los elementos que definen la incertidumbre y las adecuaciones a la metodología desarrolladas en Conabio.

La imprecisión implícita en la descripción de una localidad puede estar determinada por varios factores que interactúan entre sí, el objetivo de su determinación es obtener un valor único que considere todos los factores y que pueda ser utilizado fácilmente; a este valor único se le denomina **INCERTIDUMBRE**, y se expresa como una medida de longitud. El método de Punto-radio define una descripción de localidad con dos elementos: el punto de coordenadas y su incertidumbre, representada con una medida de longitud. Esta distancia define una circunferencia que delimita el área en donde, con mayor probabilidad, se ubica el sitio de colecta.

La incertidumbre está compuesta básicamente por un factor de distancia y otro de dirección que interactúan entre sí; en el primero se considera:

- a) La extensión de la localidad de referencia
- b) El desconocimiento del *datum*
- c) La imprecisión en la determinación de la distancia
- d) La imprecisión en la medición de las coordenadas
- e) Escala del mapa utilizado

Dentro del factor de dirección, se considera:

- f) la imprecisión con que se definió la dirección de referencia

El primer elemento, la *extensión de la localidad*, implica que una localidad no se circunscribe a una coordenada x-y, sino que se distribuye sobre una superficie determinada; por lo tanto, al establecer una distancia de referencia hacia un punto debemos considerar esta extensión dentro del cálculo de la incertidumbre.

El desconocimiento del *datum* de una carta genera incertidumbre cuando es utilizada para determinar una coordenada; Wieczorek (2001) muestra la estimación que se ha realizado para Norteamérica basada en NAD27, NAD83 y WGS84 expresada como una magnitud lineal que puede ser incorporada al cálculo de la incertidumbre. También la referencia de un mismo punto de coordenadas cambia dependiendo del datum que haya sido utilizado.

En la estimación de una distancia de referencia rara vez se especifica el método instrumental con la que fue obtenida; de manera que de acuerdo al nivel de precisión con que fue registrada (dígitos significantes), se le determina una proporción de incertidumbre. Si una distancia se registra sin fracciones decimales se toma como incertidumbre una unidad de medida; si existen decimales, éstos se convierten a notación fraccional y la incertidumbre se determina por la división de la unidad (1) entre el denominador.

La precisión con que se registra una coordenada también influye en la magnitud del error, mientras más dígitos significantes sean registrados, menor será su incertidumbre; esta incertidumbre depende de la escala de la fuente de donde se obtiene la coordenada (anexo 2).

Otro elemento que influye en el factor de distancia es la escala del mapa que es utilizado para determinar un par de coordenadas, mientras más pequeña sea la escala, el error será más grande. Si no se conoce la precisión de un mapa de acuerdo a su

escala, se sugiere tomar 1 mm de incertidumbre que corresponde al mínimo error gráfico detectable.

Por último, la magnitud de la incertidumbre asociada a la dirección está determinada por la propia precisión con que fue reportada en un registro. En el método se asumió (conservadoramente) que se debe asociar una incertidumbre más grande a una dirección cardinal (N, S, E, O), una incertidumbre intermedia a las intercardinales (NE, SE, SO, NO), y una mínima a las medias (inter-intercardinales: NNE, SSE, etc.) (anexo 01).

El formulario para el cálculo de la incertidumbre

En Conabio se diseñó una base de datos en *Microsoft Access* (con un formato de formulario) que permite introducir todos los parámetros antes mencionados de cada descripción de localidad, el formulario define las coordenadas correspondientes y calcula la incertidumbre total. También permite la captura manual de las coordenadas (cuando se obtienen las coordenadas con *Arcview* o *Biótica*), calculando solamente la incertidumbre.

En esta sección se muestra la estructura de la base de datos del formulario, su utilización se explicará en el capítulo correspondiente a la asignación de coordenadas e incertidumbre.

La base de datos tiene la siguiente estructura: (ver figura 1)

- a) **Identificador:** (id). valor único de identificación de un registro.
- b) **Escala del mapa:** Escala del material utilizado para definir una coordenada o ubicar un sitio; también se utiliza para especificar si se posee la escala que se haya utilizado para definir coordenadas ya existentes en el registro original.
- c) **Datum:** La incertidumbre correspondiente por desconocimiento del datum del mapa utilizado para asignar coordenadas. Se despliega un mapa de México

con los valores de incertidumbre, se puede seleccionar el estado en el que se localiza el registro.

- d) **Coordenadas de referencia:** Con el botón BD se despliega un menú en donde se puede seleccionar el rasgo geográfico de referencia, el cual se puede consultar a partir del termino, división política estatal y municipal, las coordenadas se registran en los cuadros de latitud y longitud; también se pueden teclear directamente si ya se posee el dato o si no aparece en el menú.
- e) **Extensión de la localidad de referencia:** Al seleccionar un rasgo geográfico con el botón BD, aparece la extensión máxima calculada en kilómetros para las poblaciones, en los rasgos que no llevan asociada dicha extensión ésta deberá obtenerse con ayuda de la cartografía.
- f) **Precisión de la coordenada:** Precisión determinada por el número de decimales que posee la coordenada, la cual esta en función de la escala de la fuente de donde se obtiene.
- g) **Incertidumbre de la coordenada:** En este campo se despliega la incertidumbre generada por la precisión de la coordenada y la extensión de la localidad.
- h) **Rosa de los vientos:** Para asignar la dirección en la que está ubicada la localidad según la descripción.
- i) **Unidades de distancia:** Unidades de distancia especificadas en la descripción (mi, m, km, etc).
- j) **Distancia:** La distancia especificada en la descripción de una localidad .
- k) **Distancias ortogonales:** Se pueden especificar medidas ortogonales en la descripción de un sitio; éstas se definen sobre los vértices de un ángulo recto a partir de un punto de referencia, el sitio que definen está determinado por la intersección de los vectores paralelos a dichas ortogonales.
- l) **Incertidumbre de distancia:** La incertidumbre generada por la precisión de la distancia.
- m) **Calcula:** Realiza la operación del cálculo de la coordenada y del error máximo.

- n) **Coordenada calculada e Incertidumbre:** Coordenada obtenida a partir de los parámetros de referencia e incertidumbre calculados que determina la incertidumbre de dicha coordenada.
- o) **Coordenada manual:** Esta opción se activa para poder introducir manualmente una coordenada de un sitio obtenida con el trazo de una distancia a través de una carretera en una carta o en una cobertura digital; es decir, cuando no se calcula la coordenada automáticamente a través del formulario. Al activar la opción dar clic en calcula aparece el recuadro en amarillo (o’).
- p) **Entidad destino:** Menú para seleccionar la entidad en la que se ubicó finalmente un registro; se utiliza para aquellos registros en donde la localidad de referencia y las coordenadas sitio se localizan en estados diferentes.
- q) **Fuente:** Se pueden seleccionar las fuentes cartográficas y nomenclaturales utilizadas en la asignación de las coordenadas de cada registro.
- r) **Clave observaciones:** Se seleccionan las observaciones que sean necesarias para especificar la georreferenciación de un registro determinado.
- s) **Validación:** Este botón genera el punto de coordenadas obtenido y lo despliega sobre el mapa de México, sirve para validar la coordenada con respecto al mapa de división política estatal y nacional.
- t) **Sucesión de puntos:** permite obtener coordenadas a partir de la medición de distancias y direcciones diferentes, que en la descripción indican ser consecutivas, p.e: Michoacán, 15 km W sobre la carretera 15 y 3 km S de la autopista.
- u) **Configuración:** Área geográfica en donde quedan incluidas las coordenadas.
- v) **Conversión de coordenadas:** Se puede hacer la conversión de coordenadas geográficas a coordenadas UTM; de coordenadas en grados, minutos y segundos a coordenadas en grados decimal y viceversa.
- w) **Equivalencias:** transformación de los diferentes tipos de unidades.

Toda la información que se registra en el formulario se almacena en una tabla llamada ‘precisión’, la cual se ubica en la sección de tablas de la base de datos en la que está contenido el propio formulario. Para utilizar el formulario en el proceso de asignación de coordenadas es necesario copiar todas las tablas de los nomenclátors asociados y el

mismo formulario a la base de datos que contenga la tabla con los registros por georreferenciar.

The screenshot shows a Microsoft Access form titled "Georreferenciación". The form is divided into several sections with various input fields, buttons, and a compass rose.

- Top Section:** Includes a menu bar with "Equivalencias", "G M S <-> DD", "Geo <-> UTM", and a "Configuración" button (labeled 'u'). Below this is a "Sucesión de puntos" button (labeled 't') and a "Mapas" section (labeled 's') showing three map icons.
- Left Section:** Contains an "Id:" field (labeled 'a') with the value "130" (labeled 'v'). Below it is a "escala del mapa:" section (labeled 'i') with fields for "1: 50000", "0.05 Km", and "1:50 000". A "Datum:" field (labeled 'c') shows "0 Km" and a "Mapa" button (labeled 'e'). Below that is an "Extensión de la localidad:" field (labeled 'i') with the value "0.11 Km".
- Center Section:** Features a compass rose (labeled 'b') with directions N, NNE, NE, ENE, E, ESE, SE, SSE, S, SSO, SO, OSO, O, ONO, and NO. A "Dirección:" field (labeled 'i') is above it. Below the compass is a "Grados:" field (labeled 'd') with the value "180°".
- Right Section:** Includes a "Unidades:" dropdown (labeled 'i') set to "mi", a "Distancia" field (labeled 'j') with the value "4.2", and a "Conversión a Kilómetros" button. Below this is a "Distancias Ortogonales" section with two rows of input fields and dropdowns (labeled 'k').
- Bottom Left Section:** Contains a "Coordenada del punto de referencia en Grados decimales" section (labeled 'f') with fields for "Longitud:" (-101.89472), "Latitud:" (24.33639), and "Precisión de la coordenada:" (0.0000001). Below this is an "Incertidumbre:" field (labeled 'g') with the value "0.000015 Km".
- Bottom Center Section:** Includes a "Coordenada del sitio" section (labeled 'o') with fields for "Longitud:" (-101.955836) and "Latitud:" (24.33639). Below this is a "Coordenada UTM del sitio" section (labeled 'n') with fields for "X:" (200226.953), "Y:" (2711105.67214), and "zona:" (14).
- Bottom Right Section:** Contains a "Distancia en Km" field (labeled 'n') with the value "6.76", an "Incertidumbre:" field (labeled 'o') with the value "5.3 Km", and a "Coordenada Manual" checkbox (labeled 'm'). Below this is a "País:" field (México) and an "Estado:" field (Zacatecas) (labeled 'p').
- Footer Section:** Includes a "Fuente" section (labeled 'q') with text about INEGI data, an "Observaciones:" section (labeled 'r') with a text area, and a "Registro:" section at the bottom showing "1330" of "1331".

Figura 1. Formulario de *Microsoft Access* para determinar la coordenada de una localidad según la descripción del sitio y el cálculo de la incertidumbre.

ORGANIZACIÓN DE LA INFORMACIÓN

Descripción y organización de la información original

Las tablas de las bases de datos originales que se georreferencian se presentan en distintos formatos por lo que la información que contienen relativa a las descripciones de localidades se puede distribuir en diferentes campos (Ver Figura 2); la primera revisión de la información original, por lo tanto, tendría como objetivo identificar todos los campos relacionados con la descripción de los sitios de colecta. Se recomienda crear una nueva tabla con todos estos campos para iniciar con ella el proceso de normalización. El término Normalización que utilizaremos en el texto se refiere al proceso de reestructuración de la información que contiene una tabla, con el objetivo de reducir al máximo la información redundante, y de esta manera, realizar de manera más eficiente el proceso de asignación de coordenadas.

La información mínima necesaria que se debe localizar en la base de datos para poder realizar la georreferenciación es la siguiente:

- País
- Entidad federativa (división política)
- Descripción del sitio de colecta: rasgo principal de referencia, distancia, dirección o rasgo que indique una dirección (por lo general, todos estos datos se presentan en un solo campo)

Sin embargo, en ocasiones esta información puede resultar insuficiente, por varias razones: a) porque se presentan localidades homónimas en un mismo municipio, b) no se encuentra el rasgo geográfico al cual se refiere la *descripción de localidad*, c) porque las localidades han cambiado de nombre, d) porque existieron en el momento de colecta y ya no existen ahora, y e) hay incertidumbre en la distancia o la dirección recorrida. Para disponer de más elementos que permitan resolver estas situaciones, se recomienda seleccionar también (si es que existen) los campos de la tabla original que posean: el nombre de la especie, la altitud, el año de colecta, las rutas de los

colectores, el nombre del colector, tipos de vegetación, etc. (ver Papayero y Llorente, 1999). Cualquier información adicional que se considere conveniente para determinar la ubicación de un sitio, debe incluirse en el proceso de organización de la información original.

a)

manis_211102 : Tabla				
LocalityID	CollectionCod	HigherGeog	Estado	SpecLocality
19	BIBM	Mexico, Baja California		no specific locality recorded
20	BIBM	Mexico, Baja California Norte		no specific locality recorded
12210	CAS	North America, Mexico		no specific locality recorded
12211	CAS	North America, Mexico		3 miles west Rio Frio
12212	CAS	North America, Mexico		Santa Cruz del Valle
12568	CAS	North America, Mexico, Baja California		Cholla Bay
12569	CAS	North America, Mexico, Baja California		Gulf of California
12570	CAS	North America, Mexico, Baja California Norte		1 mile south of Huerfano
12571	CAS	North America, Mexico, Baja California Norte		1-1/2 mile south Socorro Ranch
12572	CAS	North America, Mexico, Baja California Norte		1/2 mile south of Huerfano
12573	CAS	North America, Mexico, Baja California Norte		10 miles south El Marmol
12574	CAS	North America, Mexico, Baja California Norte		14 miles NE Punta Prieta
12575	CAS	North America, Mexico, Baja California Norte		14 miles NNE Punta Prieta
12576	CAS	North America, Mexico, Baja California Norte		14 miles north Laguna Hanson, Sierra Juarez
12577	CAS	North America, Mexico, Baja California Norte		14 miles north Laguna Hanson, Sierra Suarez
12578	CAS	North America, Mexico, Baja California Norte		15 miles NE Rancho Chapala
12579	CAS	North America, Mexico, Baja California Norte		18 miles NW Bahia de los Angeles
12580	CAS	North America, Mexico, Baja California Norte		2 miles ENE Socorro Ranch
12581	CAS	North America, Mexico, Baja California Norte		2 miles ENE Socorro Ranch

b)

hu-aves.txt - Bloc de notas				
Archivo	Edición	Formato	Ayuda	
orden, familia, genero, especie, pais, estado, sitio, colector				
Charadriiformes, Laridae, Larus, occidentalis, Mexico, Baja California, 18 mi S Point Loma, Robert D. Moore				
Pelecyaniformes, Phaethontidae, Phaethon, aethereus, Mexico, Nayarit, 20 mi off Mexican Coast, Harold H. Bailey				
Pelecaniformes, Sulidae, Sula, nebouxii, Mexico, Nayarit, 20 mi off Mexican Coast, Harold H. Bailey				
Passeriformes, Corvidae, Aphelocoma, californica, Mexico, Baja California Sur, "Agua Caliente, Cape region", Chester C. Lamb				
Podicipediformes, Podicipedidae, Tachybaptus, dominicus, Mexico, Baja California Sur, "Agua Caliente, Cape region", Chester C. Lamb				
Podicipediformes, Podicipedidae, Tachybaptus, dominicus, Mexico, Tamaulipas, Alta Mira Est., F. B. Armstrong Jr.				
Ciconiiformes, Cathartidae, Cathartes, aura, Mexico, Tamaulipas, Altamira, E Tamaulipas, F. B. Armstrong Jr.				
Charadriiformes, Jacanidae, Jacana, spinosa, Mexico, Tamaulipas, Altamira, not recorded				
Pelecaniformes, Sulidae, Sula, nebouxii, Mexico, Nayarit, "Bay of Banderas, coast of Tepic", V. W. Owen				
Pelecaniformes, Sulidae, Sula, nebouxii, Mexico, Nayarit, "Bay of Banderas, coast of Tepic", V. W. Owen				
Passeriformes, Icteridae, Agelaius, phoeniceus, Mexico, Veracruz, Camargo, not recorded				
Galliformes, Cracidae, Ortalis, vetula, Mexico, Tamaulipas, Camargo, T. H. Jackson				
Passeriformes, Icteridae, Icterus, bullockii, Mexico, Chihuahua, Camargo, T. H. Jackson				
Falconiformes, Accipitridae, Buteo, albicaudatus, Mexico, Tamaulipas, Camargo, not recorded				
Ciconiiformes, Ardeidae, Ardea, herodias, Mexico, Baja California Sur, "Cayo Islet, just off S end of island", Chester C. Lamb				
Columbiformes, Columbidae, Zenaidura, macroura, Mexico, Baja California, Cerro Prieto 7 mi E, Chester C. Lamb				
Passeriformes, Emberizidae, Pipilo, aberti, Mexico, Baja California, Cerro Prieto 7 mi E, Chester C. Lamb				
Passeriformes, Emberizidae, Pipilo, aberti, Mexico, Baja California, Cerro Prieto 7 mi E, Chester C. Lamb				
Passeriformes, Mimidae, Toxostoma, crissale, Mexico, Baja California, Cerro Prieto 7 mi E, Chester C. Lamb				
Passeriformes, Troglodytidae, Cisticolus, palustris, Mexico, Baja California, Cerro Prieto 7 mi E, Chester C. Lamb				
Passeriformes, Tyrannidae, Sayornis, saya, Mexico, Chihuahua, Colonia Pacheco, H. H. Kimball				
Passeriformes, Cardinalidae, Phoebe, melanocephalus, Mexico, Chihuahua, Colonia Pacheco, H. H. Kimball				
Passeriformes, Cardinalidae, Phoebe, melanocephalus, Mexico, Chihuahua, Colonia Pacheco, H. H. Kimball				
Passeriformes, Mimidae, Mimus, polyglottos, Mexico, Chihuahua, Colonia Pacheco, H. H. Kimball				
Passeriformes, Mimidae, Toxostoma, curvirostre, Mexico, Chihuahua, Colonia Pacheco, H. H. Kimball				
Passeriformes, Tyrannidae, Contopus, pertinax, Mexico, Chihuahua, Colonia Pacheco, H. H. Kimball				
Passeriformes, Tyrannidae, Contopus, pertinax, Mexico, Chihuahua, Colonia Pacheco, H. H. Kimball				
Passeriformes, Tyrannidae, Tyrannus, vociferans, Mexico, Chihuahua, Colonia Pacheco, H. H. Kimball				
Passeriformes, Tyrannidae, Tyrannus, vociferans, Mexico, Baja California, Colorado River Delta, G. Bancroft Jr.				
Passeriformes, Emberizidae, Pipilo, ocai, Mexico, Puebla, Chalchicomula 8 mi NE, Charles G. Sibley				
Pelecaniformes, Phaethontidae, Phaethon, aethereus, Mexico, Nayarit, "20 mi off Mexican Coast", Harold H. Bailey				

c)

mosquitos : Tabla					
id	Estado	Municipio	Corr_SpecificLocality	Elevation	Fuen
6	Baja California	Mexicali	Mexicali, 60 km SW, Guadalupe hot springs	000300	Localidad
7	Baja California	Mexicali	Mexicali, 50 km SE, Cantú springs	000200	Localidad
8	Baja California	Mexicali	Mexicali, 60 km SW, Guadalupe hot springs	000300	Localidad
9	Campeche	Campeche	Lerma, 19.4 km S, 14.2 km S por la carretera Federal Libre pav	000000	Red de C
10	Campeche	Campeche	Lerma, 20.5 km S, 15.4 km S por la carretera Federal Libre pav	000000	Red de C
11	Campeche	Campeche	Lerma, 20.5 km S, 15.4 km S por la carretera Federal Libre pav	000000	Red de C
12	Campeche	Campeche	Etna, 13.9 km E del entroke la carretera Federal Libre pavime	000050	Localidad
13	Chiapas	Comitan de Dor	Comitan de Dominguez, 40 km NW por la carretera Federal Lib		Localidad
14	Chiapas	Pijijiapan	Pijijiapan, el jardin jerico		Localidad
15	Chiapas	Palenque	Palenque		Localidad
16	Chihuahua	Juarez	Juarez	001125	Localidad
17	Coahuila	General Ceped	La Rosa, 39 km W and 6 km W de Dulces Nombres; sobre la c	001150	Localidad
18	Coahuila	General Ceped	La Rosa, 59.30 km W and 4 km W del El Carmen; sobre la can	001170	Localidad
19	Distrito Federal	Xochimilco	Xochimilco	002300	Localidad
20	Durango	Durango	14.7 KM E OF FED RT 40 (KM 196 ON FED RT 45)	001900	Localidad
21	Durango	Durango	Victoria de Durango (Durango), park along river in W part of town	001900	Localidad
22	Durango	Pueblo Nuevo	El Salto, 40 kmW por la carretera Federal Libre No. 40 (pav. 2 c	002600	Localidad
23	Guanajuato	Acapulco de Ju	Puerto Marques, 3 km NE por la carretera Federal Libre No. 200	000010	Localidad
24	Jalisco	Puerto Vallarta	Boca de Tomatlan, 16 km N por la carretera Fedal Libre No. 200	000500	Localidad
26	Jalisco	Puerto Vallarta	Puerto Vallarta, 17.5 km S por la carretera Fedal Libre No. 200		Localidad
27	Jalisco	Puerto Vallarta	Puerto Vallarta, 17.5 km S por la carretera Fedal Libre No. 200		Localidad
28	Morelos	Tepoztlan	Tepoztlan, 13.6 km W por la carretera Federal libre (pavim. 2 c	001600	Localidad
29	Nayarit	San Blas	San Blas, 3 km E y 2 km S por la carretera Federal libre (pavim	000000	Localidad
30	Nayarit	San Blas	Matanchen, 6 km SE por la carretera Fedal Libre (pavim. 2 can	000000	Localidad

Figura 2. Ejemplos de los diferentes formatos de tablas que se georreferencian. a) Base de mamíferos, b) Base de aves, c) Base de mosquitos

NORMALIZACIÓN

La normalización de la información constituye la primera etapa de la georreferenciación, es un paso esencial pues permite reducir significativamente el volumen de las bases de datos por georreferenciar, y en consecuencia, la carga de trabajo. Por lo general, las tablas de información original presentan registros duplicados debido a que se colectaron varios organismos en una misma localidad o porque ésta fue visitada por colectores diferentes.

Los duplicados se pueden clasificar en dos categorías: **registros redundantes** que son aquellos especímenes que fueron colectados en la misma localidad (por lo tanto la descripción de localidad es la misma); y **falsos duplicados** que corresponden a registros presentan una sintaxis distinta pero se refieren al mismo sitio. Ver figura 3.

	Localidad
redundantes {	0.25 mi E state boundary, 8 mi SW Calvillo
	0.25 mi E state boundary, 8 mi SW Calvillo
redundantes con errores de captura {	0.33 mi NW Ciudad de Los Ninos, 7 mi WSW Aguas
	0.33 mi NW Cuidad de Los Minos, 7 mi WSW
	0.33 mi S and 1.25 E Rincon de Romos
	0.33 mi S and 1.25 mi E Rincon de Ramos
	0.33 mi S, 1.25 mi E Rincon de Romos
falsos duplicados {	Rincon de Romos, 5 mi W
	0.5MI W RICON DE ROMOS
	0.5 mi S La Labor
	0.5 mi N Presa Calle
	0.5 mi S LaLabor

Figura 3. Campo con los datos originales referentes a la descripción de la localidad.

Los registros redundantes y falsos duplicados pueden ser eliminados de diferente forma, con una simple agrupación de tablas en un manejador de bases de datos y/o con la normalización, el realizar una u otra o ambas dependerá del número de registros que contenga la tabla de localidades a georreferenciar y de la duplicidad de localidades.

Es conveniente mencionar que las agrupaciones deben realizarse siguiendo una metodología precisa para evitar el problema de la pérdida de relaciones entre los registros de las tablas agrupadas. (anexo 03).

La normalización esta constituida por la Unificación y la Atomización de los datos.

Unificación

Consiste en eliminar los registros redundantes que presentan diferencias en sintaxis. El tener un orden alfabético de las descripciones permite identificar los registros idénticos que difieran entre sí solamente por errores de captura; de los registros idénticos se selecciona el registro que presente la sintaxis más adecuada Ver figura 4.

Desc_Localidad	Descripcion_homog
0.5 mi W of Jalisco-Nayarit border, Mex Hwy 15	0.5 mi W of Jalisco-Nayarit Border, MX 15, Nayarit, MX
0.5 mi W of Jalisco-Nayarit Border, MX 15, Nayarit, MX	0.5 mi W of Jalisco-Nayarit Border, MX 15, Nayarit, MX
0.5 mi. W of Jalisco-Nayarit border, Mex Hwy 15	0.5 mi W of Jalisco-Nayarit Border, MX 15, Nayarit, MX
8 km al SE del Rancho El Mortero municipio de Mezquitic.	8 km al SE del Rancho El Mortero, municipio de Mezquitic
8 km al SE del Rancho El Mortero, municipio de Mezquitic	8 km al SE del Rancho El Mortero, municipio de Mezquitic
8 km SE del Rancho El Mortero, Mezquitic Mun.	8 km al SE del Rancho El Mortero, municipio de Mezquitic
8 km SE of Rancho El Mortero, Mezquitic Mun	8 km al SE del Rancho El Mortero, municipio de Mezquitic
8 km SE of Rancho El Mortero, rd to Mezquitic	8 km SE of Rancho El Mortero, rd to Mezquitic
Coalcoman	Coalcomán
Coalcoman.	Coalcomán

Figura 4. Procedimiento para sustituir registros redundantes con errores de captura. Se copia la columna de descripciones, se identifican los errores de captura y se homogenizan los registros.

Criterios de Unificación

- Se conserva el idioma original.
- De los registros idénticos se selecciona el registro que presente la sintaxis más adecuada.
- Se eliminan puntos al final de la descripción.
- Si se sabe que la palabra o el nombre lleva acento, se elige la que esta acentuada y se corrigen las demás.

Atomización

- La siguiente etapa consiste en desagregar cada descripción de localidad en sus distintos elementos, en un proceso denominado **Atomización**, esta etapa permitirá la eliminación de los falsos duplicados además de estructurar de manera ordenada

la información, permite concentrar a todas las descripciones que se refieran a un mismo sitio facilitando la búsqueda de rasgos y la asignación de coordenadas. La atomización facilitará, eventualmente, el diseño de un nomenclátor de descripciones de localidades georreferenciadas.

Cada elemento de las descripciones se organiza en un campo específico en la tabla de trabajo; los elementos que deben ser identificados y desagregados en cada descripción son los siguientes:

1	<i>RefGeoSup</i>	Referencia geográfica de orden superior al rasgo principal de referencia.
2	<i>LocalidadPrincipal</i>	Nombre del rasgo geográfico principal.
3	<i>Distancia</i>	Distancia recorrida del rasgo geográfico principal al sitio, unidades de distancia abreviada (p. ej. m,km,mi) y anotaciones adicionales a la distancia (p. ej. aproximado, cerca).
4	<i>Dirección</i>	Dirección cardinal y anotaciones adicionales a ésta.
5	<i>ViaDeAcceso</i>	Nombre de la vía de acceso y anotaciones adicionales.
6	<i>Distancia_ort</i>	Distancias ortogonales: unidades y dirección cardinal (p. ej. 10 km N, 5 km E) .
7	<i>Referencia_comp.</i>	Referencias complementarias para la ubicación de una descripción de localidad.
8	<i>2aLocalidadRef</i>	Rasgo de referencia secundario.
9	<i>Informacion_adicional</i>	Información que no aporte elementos a la ubicación geográfica de un sitio.
10	<i>Altitud_ato</i>	Dato de altitud contenido en la descripción original.
11	<i>Estado_ato</i>	Nombre del estado si está contenido en la descripción original.
12	<i>Mpio_ato</i>	Nombre del municipio si está contenido en la descripción original.
13	<i>Observ_ato</i>	Observaciones particulares sobre la atomización del registro.

Ejemplo 1.

Registros que son falsos duplicados

Estado	Municipio	Descripción de localidad
México	SLP	Sierra Fría, 3.5 mi N Cerro del Maíz
México	SLP	3 1/2 mi N Cerro del Maíz
México	SLP	3.5 mi N Cerro del Maíz,
México	SLP	Sierra Fría, 3.5 mi Norte Cerro del Maíz
México	SLP	S. Fria, 3.5 mi N Cerro del Maíz

Las anteriores *descripciones de localidad* se refieren al mismo *rasgo geográfico* (el cerro), del cual parte la misma dirección y distancia; la descripción atomizada para todos los registros del ejemplo 3 quedaría de la siguiente manera:

1	2	3	4	5	6	7	8	9	10	11	12	13
Sierra Fría	Cerro del Maíz	3.5 mi	N									

Durante este proceso se deben observar ciertos criterios básicos que permitirán mantener la integridad de la información de la descripción original:

- se debe respetar el idioma en el que viene escrita la información original.
- no se debe eliminar información original, el campo llamado '*información_adicional*' puede incluir cualquier dato que esté incluido en la descripción de la localidad y que no encaje en ninguno de los campos previos.
- respetar estrictamente el orden sugerido de la atomización.
- Cualquier corrección ortográfica o sintáctica que sea necesaria (y de la que el analista tenga certeza que es correcta) debe realizarse durante la atomización.

En el proceso de atomización hay que tomar en cuenta que existen rasgos geográficos homónimos a nivel país, estado y/o municipio, y que es frecuente que las *descripciones de localidad* no cuenten con la información de las divisiones administrativas (especialmente la municipal), lo que dificulta la definición de si es o no un *rasgo geográfico* homónimo, por ejemplo:

Estado	Municipio	Descripción de Localidad	
Oaxaca	--	Ayutla, 3 KM NE	} Localidades homónimas a nivel de país
Jalisco	--	Ayutlla, 7 km N, 18 km W	
Guerrero	--	Ayutla; 3 km NE	
Chihuahua	--	Camargo	
Tamaulipas	--	Camargo	
Chihuahua	Chinipas	Las Delicias	} Localidades homónimas a nivel de entidad
Chihuahua	Guerrero	Las Delicias	
Chihuahua	Balleza	Las Delicias	} Localidades homónimas a nivel de municipio
Chihuahua	Balleza	Las Delicias	

Cuando la *descripción de localidad* no incluya la información de estado y municipio asociada se debe realizar la atomización cuidadosamente, y elegir las descripciones más congruentes y claras en cuanto a su sintaxis; una vez en la etapa de asignación de coordenadas se tendrá que acudir a los nomenclátors y cartografía para revisar si existen o no homonimias.

Durante la atomización el analista debe realizar una interpretación *a priori* de la estructura de la descripción (es decir, sin consultar la cartografía), y asignar el rasgo principal de referencia; si durante el proceso de asignación de coordenadas se revisa la cartografía y se observa que el segundo rasgo de referencia de una descripción es más adecuado para georreferenciarla, será necesario reestructurar la atomización. (anexo 04).

Ejemplo 2:

- Rancho El Milagro 40 km al NE de S.L. Potosí, carretera S.L.P.- Río Verde, Mpio. Villa Zaragoza. Veg. Encinar; suelo amarillo con piedra caliza

1	2	3	4	5	6	7	8	9	10	11	12	13
	San Luis Potosí	40 km	NE	carretera S.L.P.- Río Verde			Rancho El Milagro	Veg. Encinar; suelo amarillo con piedra caliza			Villa Zaragoza	

Si en la cartografía se puede encontrar la localidad 'Rancho El Milagro' (durante el proceso de asignación de coordenadas), entonces la atomización se corregiría de este modo:

1	2	3	4	5	6	7	8	9	10	11	12	13
	Rancho El Milagro						40 km al NE de San Luis Potosí, carretera San Luis Potosí - Río Verde	Veg. Encinar; suelo amarillo con piedra caliza			Villa Zaragoza	

Ejemplo 3:

- Paso de los Bueyes, Río Santiago, 12 km al E de Mojarra, brecha a Huajimic. Selva mediana subcaducifolia.

1	2	3	4	5	6	7	8	9	10	11	12	13
Río Santiago	Paso de los Bueyes						12 km al E de Mojarra, brecha a Huajimic	Selva mediana subcaducifolia				

Si durante la asignación de coordenadas no se encontró la localidad 'Paso de los bueyes', se corrige la atomización de esta manera:

1	2	3	4	5	6	7	8	9	10	11	12	13
	Mojarras	12 km	E	brecha a Huajimic			Paso de los Bueyes, Río Santiago	Selva mediana subcaducifolia				

Ejemplo 4:

- Agua Cruz, Mun. Jalacingo

1	2	3	4	5	6	7	8	9	10	11	12	13
	Agua Cruz										Jalacingo	

Si durante la asignación de coordenadas no se encontró la localidad 'Agua Cruz', y por lo tanto no se pudo georreferenciar la atomización se corrige de esta manera:

1	2	3	4	5	6	7	8	9	10	11	12	13
Mun. Jalacingo	Agua Cruz											

Al finalizar el procedimiento previo se debe obtener una tabla de *descripciones de localidad únicas*, a la cual se le agregaran los campos necesarios para registrar la información asociada al proceso de asignación de coordenadas.

Preparación de la tabla de trabajo

Al concluir la normalización de los datos se debe haber obtenido una tabla que contenga los campos atomizados y los campos asociados de divisiones nacionales y administrativas (país, entidad y municipio). Esta tabla deberá contener registros únicos y además debe agregársele un identificador único (id), el cual será utilizado en el formulario para registrar cada descripción de localidad georreferenciada.

Es preciso que a esta tabla se le agreguen los campos necesarios para registrar la información asociada al proceso de georreferenciación, de esta manera se obtiene la tabla de trabajo final. Los campos que se deben agregar son los siguientes:

Id_Formulario	identificador único para registrar en el formulario.
No_Georref	Explicación de la razón por la que el registro no pudo ser georreferenciado.
Analista	Persona que hizo la georreferenciación del registro.
Fecha	Fecha de georreferenciación del registro.

Los campos en donde se registran las coordenadas, la incertidumbre, las observaciones y las fuentes cartográficas están incluidos en la tabla asociada al formulario de Access, por lo que no deben agregarse a la tabla de trabajo.

La figura 5 muestra la estructura de una tabla de trabajo; se puede observar, en este ejemplo, que se conservan los nombres de los campos originales asociados a la descripción de localidad ('country', 'fst_political', 'snd_political' y 'elevation'). (anexo 03).

	Nombre del campo	Tipo de datos
►	ID_2	Autonumérico
	country	Texto
	fst_political	Texto
	snd_political	Texto
	elevation	Texto
	RegGeoAdicional	Texto
	LocalidadPrincipal	Texto
	Distancia	Texto
	Direccion	Texto
	ViaDeAcceso	Texto
	Distancia_ort	Texto
	Referencia_comp	Texto
	2aLocalidadRef	Texto
	Altitud_ato	Texto
	Estado_ato	Texto
	Mpio_ato	Texto
	Id_Formulario	Número
	No_Georref	Texto
	Analista	Texto
	Fecha	Fecha/Hora
	Observaciones	Memo

Figura 5. Ejemplo de la estructura una tabla de trabajo finalizada.

La agrupación de tablas y la integración de los resultados

Durante el proceso de normalización se realizan diversas agrupaciones de la tabla que se está depurando con lo cual se van obteniendo, progresivamente, tablas con menor número de registros. Al finalizar la asignación de coordenadas a las descripciones de la última tabla obtenida (denominada 'tabla de trabajo') será preciso integrar las coordenadas y su información asociada a la tabla original de la cual se partió.

Para que este proceso se realice de manera eficaz es preciso que se puedan establecer las ligas adecuadas entre las descripciones de localidades agrupadas, a lo largo de las distintas tablas obtenidas; es decir, las coordenadas asignadas a una descripción de la tabla de trabajo deben poder integrarse a cada uno de los registros redundantes presentes en la tabla original (si es que existieran descripciones redundantes o falsos duplicados para esa descripción en particular).

Una descripción se relaciona en todas las tablas agrupadas a través de identificadores (id's). En el diseño de bases de datos, se denomina *Llave primaria* (PK) de una tabla al campo que contenga un identificador único por cada registro, y *Llave foránea* (FK) al campo de una tabla que contenga identificadores que pueden duplicarse. El procedimiento de agrupación de tablas en la normalización requiere que a cada tabla agrupada se le agregue una llave primaria y una llave foránea para establecer las ligas que permitan integrar los resultados de la georreferenciación. (anexo 03).

MANEJO Y USO DEL MATERIAL

Cartografía

La información cartográfica es una herramienta esencial durante el proceso de georreferenciación, ya que sin ella sería muy difícil poder ubicar las coordenadas de las diferentes descripciones de localidad. Es por eso que se mencionarán algunas de las características de la Cartografía impresa y la Cartografía Digital.

Características de la Cartografía impresa

Se pueden medir distancias y obtener coordenadas de los diferentes rasgos de manera manual, pero esta forma lleva más tiempo que el hacerlo con un software especial y con cartografía digital. Otra desventaja es que no se puede sobreponer la información de mapas con escalas diferentes.

Características de la Cartografía Digital

La información cartográfica digital es principalmente de dos tipos: en conjunto de datos vectoriales (puntos, líneas y polígonos) e imágenes (.tif, .sid, entre otras); dentro de la cartografía digital se incluye el modelo digital del terreno (mdt)

Ventajas: permite la sobreposición de diferentes temas de información en diferentes escalas, siempre y cuando contengan la misma proyección geográfica. Se pueden realizar acercamientos de una determinada zona.

En conjuntos de datos vectoriales se pueden realizar búsquedas de información, mientras que las imágenes sólo se pueden visualizar.

Ubicación de la información cartográfica utilizada

Es conveniente que la información cartográfica que se utilice para georreferenciar se encuentre contenida en un lugar específico, y pueda consultarse de manera simultánea entre varios usuarios (anexo 06).

Softwares

Los softwares más utilizados en el proceso de georreferenciación son: Arcview, Biótica, ya que en ellos es posible organizar la información por medio de proyectos, realizar consultas, búsquedas, así como medir distancias y obtener coordenadas (anexo 05).

ASIGNACIÓN DE COORDENADAS E INCERTIDUMBRE

La segunda etapa del proceso está constituida por la asignación de coordenadas e incertidumbre a los registros normalizados. En la Conabio se utilizan cuatro técnicas para georreferenciar localidades: 1) por medio del formulario de *Access*, 2) con el *software* de *Arcview*, 3) con el sistema de información Biótica y 4) por medio de cartografía impresa.

En esta primera sección del capítulo se abordará la asignación de coordenadas con las cuatro técnicas mencionadas y posteriormente se especificará el uso del formulario para determinar la incertidumbre.

Georreferenciación por medio de Cartografía impresa

La georreferenciación se realiza por medio de cartas escala 1: 50,000 o 1: 250,000 (escalas disponibles en Conabio), editadas por el INEGI; utilizando un curvímeter digital.

La opción de georreferenciar por medio de cartografía impresa se utiliza básicamente cuando la información digital es insuficiente para localizar un sitio; los casos más frecuentes en donde se presenta esta situación se refieren a descripciones de localidades relacionadas con caminos rurales o ríos, o con elementos que carecen de una cartografía detallada. (anexo 08, Sección: Cartografía Impresa)

Georreferenciación por medio de *ArcView*

Arcview es utilizado básicamente para la georreferenciación de las descripciones de localidades del tipo 7, es decir, las que deben resolverse recorriendo una ruta definida por algún rasgo geográfico (carretera, autopista, río, etc.). Pero este *software* también permite resolver otros procesos de la georreferenciación, como la diferenciación de homónimos, la estimación de la extensión de los rasgos, medición de distancias para definir incertidumbres, y principalmente el manejo de la cartografía para ubicar localidades. (anexo 08, Sección: *ArcView*)

En *Arcview* es posible:

- a) Organizar y manipular la información cartográfica y nomenclatural. La cartografía y los nomenclátors que posee Conabio para realizar la georreferenciación de localidades está ordenada en proyectos de trabajo de *Arcview* para cada entidad del país. En cada proyecto se puede encontrar información de: localidades, toponimias por entidad federativa³, división estatal, división municipal, un modelo digital del terreno, aeropuertos, cuerpos de agua, estaciones climáticas, estaciones hidrométricas, puertos, puentes, cruces de carreteras, red de carreteras, ríos, Islas, observatorios climatológicos, toponimia de cartas de INEGI en escala 1: 50,000 y en 1: 250,000.
- b) Búsqueda automatizada y selección de localidades o rasgos geográficos por medio de consultas. Esta herramienta además permite visualizar los elementos seleccionados mostrando su ubicación espacial, y ayuda a discernir entre rasgos homónimos para escoger aquél que corresponda a una descripción determinada.
- c) Medición de distancias. *Arcview* tiene una herramienta que permite medir distancias sobre la cartografía, pudiendo medirse en línea recta o siguiendo un rasgo de la cartografía. Esta función se utiliza también para definir las extensiones máximas de los rasgos.
- d) Obtención de las coordenadas de un punto. Una función de *Arcview* permite trazar una distancia y obtener las coordenadas del punto final.

Georreferenciación por medio de Biótica (versión 4.1)

El Sistema de Información Biótica permite realizar funciones similares a las descritas para el *software* de *ArcView* que permiten realizar la georreferenciación de localidades del tipo 7 y también del 9, es decir, *por aire*. Estas funciones se localizan en el módulo de SIG de Biótica. (anexo 08, Sección: Biótica).

³ Esta información corresponde tanto a nombres de localidades como a *rasgos geográficos*, como cerros, mesetas, ríos, infraestructura de aspectos sociales, áreas de protección de flora y fauna silvestres, entre otros.

- 1) La información cartográfica y nomenclatural también está organizada en proyectos por entidad.
- 2) Permite consultar localidades o rasgos y ubicarlos espacialmente.
- 3) Obtener las coordenadas de un punto a partir de una dirección y distancia definidas previamente; y también medir distancias sobre la cartografía obteniendo las coordenadas finales.

Georreferenciación por medio del formulario de Access

La base de datos en Access con formato de formulario es una herramienta que se diseñó fundamentalmente para la asignación de la incertidumbre en la georreferenciación, pero también cumple una función básica para asignar las coordenadas a varios tipos de descripciones. En esta sección se explicarán exclusivamente las utilidades del formulario para asignar las coordenadas, y en la parte final del capítulo como se mencionó en el párrafo anterior, se retomará el formulario para explicar el cálculo de la incertidumbre.

Las descripciones de localidades del tipo 4, 5, 6 , 8 y 9 presentados en la tabla 1 (ubicada en la sección ‘Los tipos de descripciones de localidades’) pueden resolverse a través del formulario, capturando los datos de cada descripción en las casillas de captura.

El tipo de descripción de localidad número 4 se refiere a aquellas que ya poseen coordenadas, en este caso, el formulario se utiliza para asignarle únicamente la incertidumbre. En el tipo 5, compuesto solamente por el nombre de una localidad o rasgo, se puede asignar la coordenada de origen del mismo y adicionar la incertidumbre de su extensión. En la descripción del tipo 6, se puede agregar la distancia mencionada a la extensión de la propia localidad para definir su incertidumbre, asignando las coordenadas de la misma localidad. Para las descripciones del tipo 8 el formulario posee campos para ingresar medidas ortogonales y definir las coordenadas del sitio. Finalmente, para las descripciones del tipo 9 que son aquellas definidas *por aire*, se

puede ingresar la coordenada de origen, la distancia y la dirección definidas en la descripción y el formulario determina las coordenadas del sitio.

En el formulario existen cuatro campos básicos con los cuales se pueden asignar las coordenadas del sitio a todos los tipos de localidades mencionados previamente (ver la sección ‘El formulario para el cálculo de la incertidumbre’ donde se describen todos los campos del formulario):

- a) Extensión de la localidad de referencia
- b) Rosa de los vientos
- c) Distancia y
- d) Distancias ortogonales

Dependiendo de sus características, se utilizan los campos correspondientes capturando la información propia para cada descripción. Con el botón ‘Calcula’ se obtienen las *coordenadas sitio*, es decir, las coordenadas del punto de colecta. (anexo 08, Sección: Uso del Formulario)

Casos especiales de georreferenciación

Existen diversos casos de descripciones de localidades que presentan características particulares y que requieren de la utilización de criterios especiales para georreferenciarlos, es decir, se pueden clasificar dentro de los nueve tipos descritos en la tabla 1, pero poseen ciertos elementos en la descripción que requieren procedimientos específicos para asignarles las coordenadas o la incertidumbre. Por ejemplo, la georreferenciación a partir de un cerro, de una bahía, de una corriente de agua, un punto intermedio entre dos localidades, etc.

Dentro de la metodología adoptada en Conabio para la georreferenciación de localidades se han establecido criterios para resolver los denominados ‘casos especiales’; el objetivo central de la documentación de estos criterios es que se realicen de la misma manera los diferentes tipos de casos, pues la aplicación de criterios

diferentes puede llegar a ser igualmente válida, sin embargo el objetivo es homogenizar los procedimientos (está establecido en la propia metodología del proyecto MaNIS que un criterio para georreferenciar puede ser válido si éste es documentado apropiadamente en los registros en que haya sido utilizado).

La utilización de algún criterio especial para georreferenciar debe registrarse en el campo de observaciones de la tabla de trabajo; los casos especiales que se presentan más frecuentemente tienen asignada una clave numérica que pertenece a las claves que se manejan para clasificar las observaciones sobre la georreferenciación. En el anexo 09 se presenta una guía gráfica que describe con detalle los diferentes tipos de 'casos especiales' y presenta el criterio convenido en cada caso. En el anexo 10 se encuentra la lista de todas las observaciones documentadas que sirven para especificar los criterios utilizados en la asignación de coordenadas a los registros.

Registros homónimos y registros no georreferenciados

Dentro de los registros que deben ser clasificados como no georreferenciados están incluidos los registros que tengan homónimos en sus descripciones que no pudieron ser resueltos y los tres primeros tipos de descripciones presentadas en la tabla 1. La causa por la que un registro no pudo ser georreferenciado debe ser explicada en el campo correspondiente de la tabla de trabajo: '*No_georref*'.

Homonimias

Dentro de las bases de datos es muy frecuente la existencia de registros homónimos, los cuales requieren de un análisis cuidadoso para evaluar su posible asignación de coordenadas. Las localidades o rasgos geográficos que presenten homonimias no pueden georreferenciarse, a menos que la descripción contenga algún otro elemento que permita diferenciar ese sitio del resto de los sitios homónimos. Se consideran registros homónimos todos aquellos que presenten al menos una palabra igual dentro del nombre completo (excluyendo artículos, preposiciones, etc.), y que coincidan en el mismo municipio de una entidad.

Por ejemplo:

<i>localidad de referencia</i>	<i>localidades encontradas</i>
San Juan	San Juan Nepomuceno San Juan Buenaventura San Juan
Guadalupe	Rancho Guadalupe Mina de Guadalupe Guadalupe Cabeza rancho la Guadalupe

Todas las localidades o rasgos geográficos ubicados en la columna de la derecha son considerados homónimos, asumiendo que todos pertenecieran al mismo municipio.

Sólo deben georreferenciarse los registros que presenten localidades o rasgos homónimos cuando éstos estén ubicados de manera adyacente, pues es común que los asentamientos pequeños localizados en la periferia de grandes ciudades se denominen con el mismo nombre. En este caso se debe tomar como localidad de referencia aquella con el mayor número de habitantes (ver figura 6), y asignar una incertidumbre tal que abarque al homónimo de menor población.

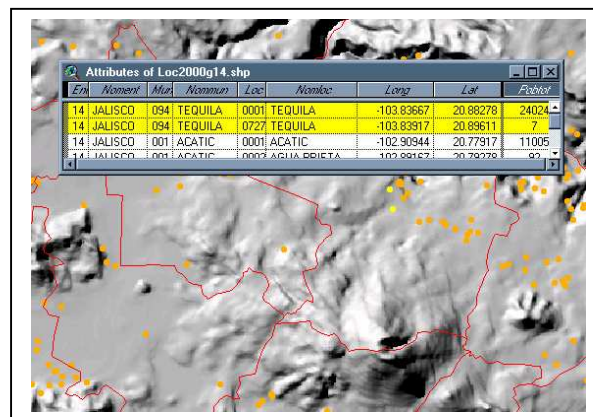


Figura 6. Localidades homónimas adyacentes, en este caso se debe tomar como localidad para georreferenciar la que posee el mayor número de habitantes.

Descripción de localidad dudosa (tipo 1)

Existen descripciones de localidades que poseen un signo de interrogación, el cual debe ser interpretado como la carencia de una certeza absoluta sobre el nombre correcto del sitio de colecta; debido a la incertidumbre explícita de la descripción misma, no se le pueden asignar coordenadas.

Por ejemplo:

‘San José?’

Descripción de localidad que no puede ser ubicada (tipo 2)

Las razones por la que una localidad no pueda ser ubicada son de dos tipos: 1) porque los datos contenidos en el propio registro son insuficientes o imprecisos y 2) porque no se cuenta con la cartografía o nomenclátors lo suficientemente detallados como para ubicar un sitio determinado.

Un ejemplo de una descripción con información insuficiente sería, por ejemplo:

‘San José, México’

Y un ejemplo de una descripción imprecisa sería:

‘Sierra Madre oriental’

Descripción de localidad comprobadamente inexacta (tipo 3)

En ocasiones, las referencias dadas para ubicar una descripción son incongruentes entre sí; esta aseveración se debe cotejar con la cartografía disponible y con los nomenclátors. Las incongruencias más frecuentes se refieren a localidades o rasgos que son ubicados en una división política o administrativa que no le corresponde.

Como una forma práctica de capturar en las tablas de trabajo las razones por las que no se asignaron coordenadas a las descripciones, se ha establecido una clave numérica para que sea utilizada durante el proceso de georreferenciación. Esta clave debe ser

sustituida durante el proceso o cuando se termine de georreferenciar una tabla, al integrar el trabajo con la información original.

Clave	Causa de no georreferenciación
1	Registro inconsistente.
2	Registro sin información suficiente o impreciso.
2A	Registro que presentó homónimos que no pudieron resolverse.

Ver anexo 10 Guía de claves de observaciones sobre la georreferenciación.

Localidades o rasgos nuevos

Es posible que durante el proceso de asignación de coordenadas se localicen nuevos elementos geográficos que no estén registrados en la cartografía o en los nomenclátors disponibles, es decir, localidades o rasgos que sean descritos con suficiente precisión en las tablas y que puedan identificarse como rasgos nuevos y ser registrados.

Existe una tabla de *Access* en la que se registran las localidades o rasgos nuevos que se van encontrando en el proceso de georreferenciación, se ubica en *N/formulario/georref.mdb/nuevos_rasgos*. En esta tabla se documentan todas las características de referencia del nuevo rasgo, como la entidad, el municipio, el nombre, las coordenadas, etc. Para determinar un nuevo rasgo se debe revisar cuidadosamente la cartografía disponible y la fuente de donde se extrae la nueva información, de forma que se tenga la certeza de que el rasgo es correcto y está bien ubicado. El registro de estos nuevos elementos enriquecerá progresivamente los nomenclátors disponibles.

VALIDACIÓN Y REPORTE

Integración de los resultados

Una vez que una tabla de trabajo ha sido georreferenciada es necesario integrar toda la información generada a los registros pertenecientes a la base de datos original, es

decir, las coordenadas asignadas, la incertidumbre y todos los comentarios realizados sobre el proceso. Esta información está contenida tanto en la tabla de trabajo final como en la tabla 'precisión' del formulario. Durante el proceso de depuración se fue reduciendo el volumen de registros con el fin de obtener sólo descripciones de localidades únicas; así que ahora es necesario ligar toda la información generada a la tabla original de la cual se partió.

Validación de las coordenadas

Las coordenadas que se obtuvieron a través de la georreferenciación son sometidas a una validación con el límite nacional de México y con los límites estatales, con la cual se verifica la correcta ubicación de cada coordenada dentro de los territorios respectivos. Si no se conoce el origen de los datos, es decir si son colectas acuáticas o terrestres, se opta por manejar el criterio de ubicar todas las coordenadas sobre tierra.

La validación se realiza importando la tabla de trabajo en un proyecto de *Arcview*, la cual se sobrepone con el límite nacional y los límites estatales. Aquellos puntos en los que se detecten irregularidades se identifican y se realiza la verificación de su georreferenciación.

El proceso detallado de la integración de los resultados está descrito en el Anexo 11, sección Integración de datos y validación de resultados, el procedimiento está basado en el manejador de bases *Access*.

Elaboración del reporte

Existe un formato para reportar la georreferenciación de las bases de datos, en el cual se registra la información básica que permita llevar el seguimiento y control sobre el trabajo realizado en el área. Los reportes pueden ser parciales o referidos a la finalización de una base de datos, dependiendo del volumen que presente cada base; las bases pequeñas se reportarían una vez que hayan sido finalizadas, mientras que en las bases grandes se pueden ir haciendo reportes parciales sobre el avance.

En el formato del reporte se deben especificar los datos principales de cada tabla: el nombre, la fecha en la que se trabajó, el nombre de los campos que se agregaron a los originales, el país de origen de los datos, el número total de registros de la base original, el número de registros a los que se les asignaron coordenadas, el total de registros no georreferenciados y el número de registros únicos.

El formato que se utiliza se encuentra en el anexo 11, Sección Elaboración de reporte.

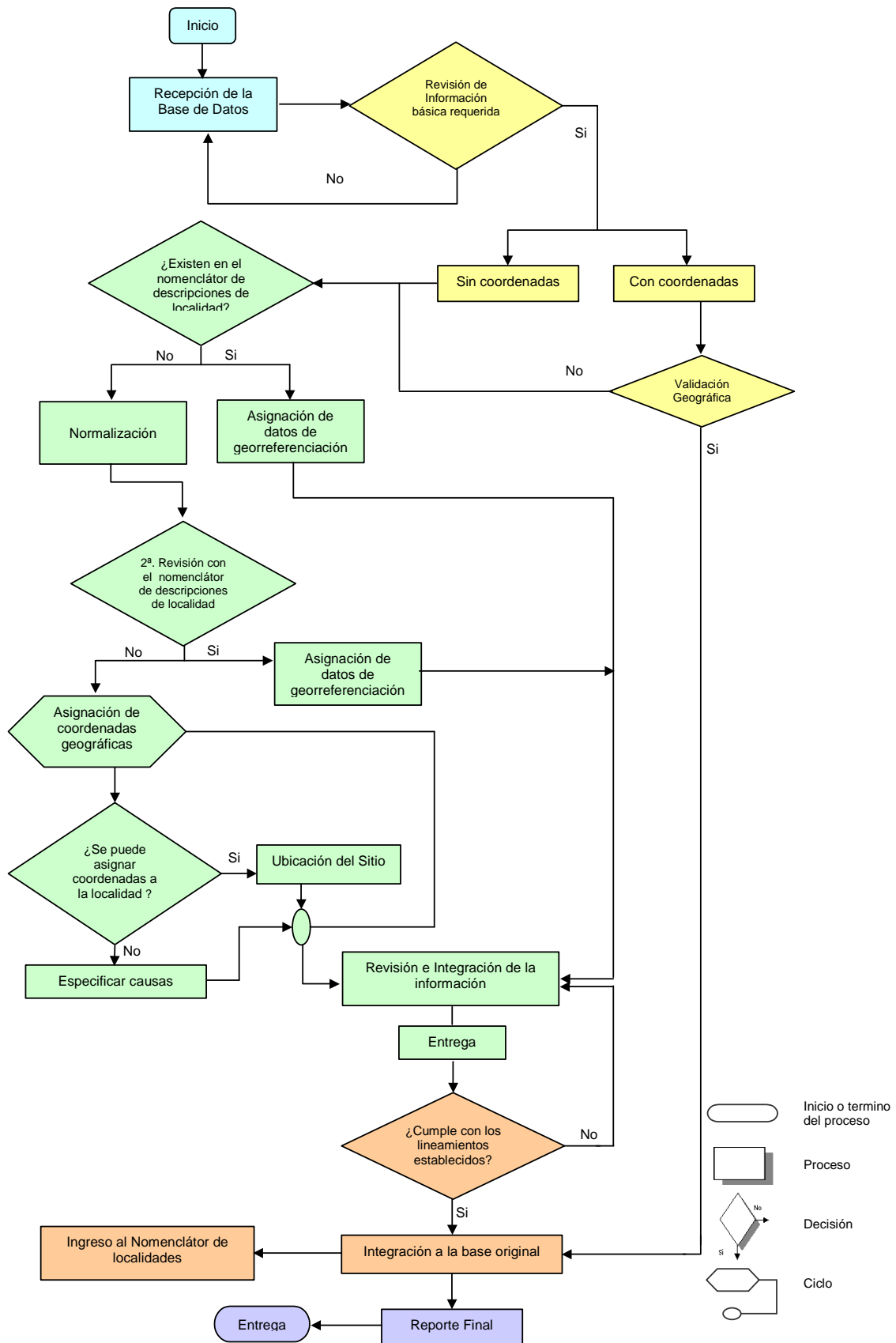


Fig. 8. Diagrama de flujo del proceso de Georreferenciación