

# GEORREFERENCIACIÓN DE LOCALIDADES: UNA GUÍA DE REFERENCIA PARA COLECCIONES BIOLÓGICAS



Convenio interadministrativo 13-014 (FA 005 de 2013) Instituto de Investigación de Recursos Biológicos Alexander von Humboldt - Fondo Adaptación

Subdirección de Servicios Científicos y Proyectos Especiales  
Instituto de Investigación de Recursos Biológicos Alexander von Humboldt  
Bogotá, D.C., 2015

# Georreferenciación de localidades: Una guía de referencia para colecciones biológicas

**Versión 3.0**  
Febrero de 2015

Proyecto: Insumos técnicos para la delimitación de ecosistemas estratégicos priorizados (páramos y humedales)

**Programa de Evaluación y Monitoreo del Estado de la Biodiversidad**  
Instituto de Investigación de Recursos Biológicos Alexander von Humboldt



Dairo Escobar

Srahyrlandy Rocío Díaz

Leidy Marcela Jojoa

Eduardo Rudas

Jhonny Saavedra



**Cítese como:**

Escobar D, Díaz SR, Jojoa LM, Rudas E, Albarracín RD, Ramírez C, Gómez JY, López CR, Saavedra J (2015). Georreferenciación de localidades: Una guía de referencia para colecciones biológicas. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt – Instituto de Ciencias Naturales, Universidad Nacional de Colombia. Bogotá D.C., Colombia. 95 p.

**ISBN:** Pendiente

**URI Persistente:** Pendiente

**Idioma:** Español

© Instituto de Investigación de Recursos Biológicos Alexander von Humboldt 2015

**Licencia:** Esta obra está bajo una licencia de *Creative Commons* Reconocimiento 4.0 Internacional. <http://creativecommons.org/licenses/by/4.0/>



## Índice de autores

### **Dairo Escobar**

Coordinador

Sistema de Información sobre Biodiversidad de Colombia - SiB Colombia

Instituto Humboldt

[descobar@humboldt.org.co](mailto:descobar@humboldt.org.co)

### **Srahyrlandy Rocío Díaz Sánchez**

Investigadora - Equipo de Georreferenciación

Instituto Humboldt

[sdiaz@humboldt.org.co](mailto:sdiaz@humboldt.org.co)

### **Leidy Marcela Jojoa**

Investigadora - Equipo de Georreferenciación

Instituto Humboldt

[ljojoa@humboldt.org.co](mailto:ljojoa@humboldt.org.co)

### **Eduardo Rudas**

Coordinador – Programa Informática de la Biodiversidad

Instituto de Ciencias Naturales, Universidad Nacional de Colombia

[erudasb@unal.edu.co](mailto:erudasb@unal.edu.co)

### **Jhonny Saavedra**

Líder - Infraestructura Institucional de Datos (I2D)

Instituto Humboldt

[jsaavedra@humboldt.org.co](mailto:jsaavedra@humboldt.org.co)

## Coautores

### **Rubén Darío Albarracín Caro**

Geógrafo - Programa Informática de la Biodiversidad

Instituto de Ciencias Naturales, Universidad Nacional de Colombia

[rdalbarracinc@unal.edu.co](mailto:rdalbarracinc@unal.edu.co)

### **Jorge Yovani Gómez Rodríguez**

Georreferenciador - Programa Informática de la Biodiversidad

Instituto de Ciencias Naturales, Universidad Nacional de Colombia

[jygomezr@unal.edu.co](mailto:jygomezr@unal.edu.co)

### **Carlos René López Candela**

Georreferenciador - Programa Informática de la Biodiversidad

Instituto de Ciencias Naturales, Universidad Nacional de Colombia

[candelopezcarlos@gmail.com](mailto:candelopezcarlos@gmail.com)

### **Carlos Ramírez**

Georreferenciador - Equipo de Georreferenciación

Instituto Humboldt

[cramirez@humboldt.org.co](mailto:cramirez@humboldt.org.co)

# Contenidos

INTRODUCCIÓN.....	9
ANTECEDENTES .....	11
<b>CAPÍTULO 1. BASES GENERALES DE LA GEORREFERENCIACIÓN DE REGISTROS BIOLÓGICOS</b>	<b>13</b>
1.1 ¿Que es la georreferenciación de localidades de registros biológicos? .....	14
1.2 ¿Por qué es importante georreferenciar los registros de las colecciones biológicas? .....	15
1.3 La georreferenciación en la estructura de trabajo de las colecciones: Una aproximación desde el IAvH .....	16
1.4 Factores que inciden en el estado de la localidad.....	18
1.4.1 Percepción del colector .....	18
1.4.2 Uso de Sistema de Posicionamiento Global (GPS) en la localización de registros biológicos .....	19
1.4.3 Cambio temporal de las entidades geográficas.....	21
1.5 Lineamientos para el manejo de bases de datos en la georreferenciación .....	22
1.5.1 Reducir la cantidad de datos de localidad a procesar.....	22
1.5.2 Mantener los campos que contienen los datos de la localidad original.....	23
1.5.3 Respetar la definición de los elementos .....	23
<b>CAPÍTULO 2: MÉTODO RADIO - PUNTO .....</b>	<b>24</b>
2.1 Métodos de georreferenciación y Método Radio-Punto .....	25
2.2 Incertidumbre de los datos geográficos según el Método Radio-Punto.....	26
2.2.1 Extensión de la localidad de referencia.....	27
2.2.2 Desconocimiento del Datum .....	29
2.2.3 Imprecisión en la medida de distancia.....	31
2.2.4 Incertidumbre por orientaciones .....	33
2.2.5 Imprecisión en la medida de las coordenadas.....	35

2.2.6 Incertidumbre por escala del mapa.....	36
<b>CAPÍTULO 3: PROTOCOLO DE GEORREFERENCIACIÓN DE REGISTROS BIOLÓGICOS .....</b>	<b>38</b>
3.1 Proceso general de la georreferenciación de localidades .....	39
3.2 Reconocimiento de la información .....	42
3.2.1 Revisión de elementos geográficos .....	42
3.2.2 Preparación del archivo de trabajo.....	42
3.3 Estandarización de localidades.....	44
3.3.1 Organización de la información .....	45
3.3.2 Aplicación de reglas y normalización de las localidades:.....	45
3.4 Clasificación de la localidad en niveles de calidad .....	49
3.4.1 Asignación de niveles de calidad de la localidad.....	51
3.5 Georreferenciación e incertidumbre .....	53
3.5.1 Consideraciones técnicas de la cartografía digital e insumos cartográficos.....	53
3.5.2 Criterios generales para la verificación de localidades .....	56
3.5.3 Criterios para la asignación de coordenadas e incertidumbre .....	59
3.6 Validación de datos .....	60
<b>CAPÍTULO 4. PROCEDIMIENTOS SEGÚN TIPOS DE LOCALIDAD .....</b>	<b>61</b>
4.1 Procedimientos de georreferenciación .....	62
4.2 Localidades con coordenadas (Nivel 1) .....	62
4.3 Localidades sin Coordenadas .....	64
4.3.1 Localidades del Nivel 2 .....	64
4.3.2 Localidades del Nivel 3 .....	70
4.3.3 Localidades del Nivel 4 .....	75
4.3.4 Localidades del nivel 5.....	76
4.3.5 Localidades del nivel 6.....	78
4.3.6 Localidades del nivel 7 .....	80

4.3 Tratamiento de la altura en la georreferenciación .....	81
<b>CAPÍTULO 5: HERRAMIENTAS ÚTILES PARA LA GEORREFERENCIACIÓN .....</b>	<b>84</b>
5.1 Cálculo automático del centroide.....	85
5.1.1 Elaboración del Workspace.....	85
5.1.2 Creación de centroides.....	87
5.2 Manejo de archivos raster en QGIS.....	89
5.3 Herramienta para la validación de localidades georreferenciadas .....	90
<b>BIBLIOGRAFÍA .....</b>	<b>92</b>
<b>OTROS RECURSOS .....</b>	<b>95</b>

## Índice de figuras

<b>Figura 1.</b> Localidad georreferenciada de un registros biológico a partir de los datos de etiqueta. Localidad: 6 km al Norte de Mariquita a lo largo de la ruta a Victoria, cerca al río Gualí. ....	14
<b>Figura 2.</b> La georreferenciación dentro de la dinámica de trabajo de las colecciones biológicas de IAvH.....	17
<b>Figura 3.</b> Temporalidad de las localidades con coordenadas en las colecciones biológicas del Instituto Humboldt. Las líneas punteadas indican las fechas de uso de la tecnología GPS y la influencia de este en el aumento de registros con coordenadas. ....	20
<b>Figura 4.</b> Temporalidad de los datos de localidad de los registros biológicos del IAvH desde el año 1851 hasta 2013. Las líneas punteadas marcan los principales cambios de la división político administrativa (DANE 2001), factor incide en la información de la localidad. ....	22
<b>Figura 5.</b> Métodos de georreferenciación: A) Método punto B) Método cuadrado C) Método polígono.....	25
<b>Figura 6.</b> Método Radio Punto A) Cálculo del centroide B) Cálculo de la incertidumbre. ....	26
<b>Figura 7.</b> Esquema de la incertidumbre por extensión de una entidad geográfica definida por un polígono. ....	27
<b>Figura 8.</b> Centroide de un río y extensión de incertidumbre. ....	28
<b>Figura 9.</b> Extensión de incertidumbre en localidades con coordenadas.....	28
<b>Figura 10.</b> Cálculo del centroide. A) Centroide calculado con el eje mayor y menor. B) Cálculo del centroide de un polígono irregular. C) Centroide calculado por un proceso automatizado.	29
<b>Figura 11.</b> Cambio en la localización de las coordenadas usando diferentes datum (Wieckzorek et al. (2001). ....	30
<b>Figura 12.</b> Desplazamiento en metros de una misma coordenada con diferentes datum para Colombia (MAGNA-SIRGAS y Bogotá) y algunos datum para Suramerica (Wgs 72, SAD69, Pico de Nieves y PDO Survery Datum 1993). ....	31
<b>Figura 13.</b> Uso de la distancia para describir un trayecto en campo. ....	32
<b>Figura 14.</b> Cálculo de incertidumbre por distancia según el Método de Wieckzoreck et al (2004). ....	33
<b>Figura 15.</b> Incertidumbre por orientaciones: A) Orientaciones cardinales, B) Orientaciones cardinales primarias y C) Orientaciones intercardinales secundarias.....	34

<b>Figura 16.</b> Cálculo de la incertidumbre integrando orientación y distancia mediante el teorema del coseno A) Incertidumbre solo por orientación B) Incertidumbre integrada. ....	35
<b>Figura 17.</b> Diagrama de flujo del proceso de georreferenciación de localidades de registros biológicos.....	40
<b>Figura 18.</b> Diagrama de flujo del proceso de preparación de archivo de trabajo y estandarización de localidades. ....	48
<b>Figura 19.</b> Proceso para la asignación de niveles según la información de la localidad. ....	52
<b>Figura 20.</b> Coordenadas próximas con la localidad y cálculo de incertidumbre por extensión.	63
<b>Figura 21.</b> Coordenadas que no coinciden con la descripción de la localidad. ....	64
<b>Figura 22.</b> Procedimientos de georreferenciación de las localidades del Nivel 2. ....	65
<b>Figura 23.</b> Localidad con distancia y orientación. A) Asignación de la coordenada B) Gráfica del cálculo de incertidumbre integrando, donde, a) 2 km b) 2.3 km y c) Radio de incertidumbre. ..	66
<b>Figura 24.</b> Localidad con distancia en una ruta. A. Asignación de coordenadas B. Cálculo de incertidumbre. ....	67
<b>Figura 25.</b> Entidad geográfica con orientación. A. Asignación de coordenadas B. Cálculo de incertidumbre. ....	68
<b>Figura 26.</b> Localidad que contiene una distancia desde un referente geográfico. El punto rojo marca el sitio para la asignación de las coordenadas y la esfera la extensión de la incertidumbre. ....	68
<b>Figura 27.</b> Asignación de coordenadas mediante la intersección de orientaciones ortogonales. ....	69
<b>Figura 28.</b> Tratamiento matemático para el cálculo de incertidumbre de localidades con orientaciones ortogonales .....	70
<b>Figura 29.</b> Procedimientos de georreferenciación de las localidades del Nivel 3. ....	71
<b>Figura 30.</b> Distancia recomendable para la georreferenciación en una vía. Las X rojas definen los puntos para georreferenciar a los lados de vía. ....	72
<b>Figura 31.</b> Entidades conectadas por una ruta. A) Asignación de coordenadas B). Extensión de la incertidumbre. ....	72
<b>Figura 32.</b> Asignación de coordenadas y evaluación de incertidumbre para una localidad no delimitada.....	73
<b>Figura 33.</b> Georreferenciación y evaluación de la incertidumbre para un registro sin datos de localidad y con reporte del municipio. ....	74

<b>Figura 34.</b> Procedimientos de georreferenciación de las localidades del nivel 4.....	75
<b>Figura 35.</b> Entidad geográfica que comparte varios municipios. A) Extensión total del río y división municipal B) Asignación de coordenadas y cálculo de incertidumbre.....	76
<b>Figura 36.</b> Procedimientos de georreferenciación de las localidades del nivel 5.....	77
<b>Figura 37.</b> Georreferenciación e incertidumbre de una localidad de distribución nacional.....	78
<b>Figura 38.</b> Procedimiento de georreferenciación de las localidades del Nivel 6. ....	79
<b>Figura 39.</b> Georreferenciación y cálculo de incertidumbre de registro donde el único referente espacial es el país. ....	79
<b>Figura 40.</b> Ejemplo de una localidad ambigua e inconsistente.....	80
<b>Figura 41.</b> Comparación del proceso de georreferenciación de una misma localidad con ausencia y presencia de la altura. La línea recta corresponde a una línea imaginaria que conecta las dos entidades geográficas y que se utiliza para georreferenciar. A) Localidad sin datos de altura B) Localidad con una dato de altura de 1200 m.s.n.m. ....	82
<b>Figura 42.</b> El sitio de colecta está en la ruta que conecta dos entidades. A) Vista general, el cuadro punteado define el sitio de interés B) Acercamiento del sitio de la colecta. ....	82
<b>Figura 43.</b> Georreferenciación y evaluación de la incertidumbre para un registro sin datos de localidad y con reporte del municipio. A. Sin dato de altura B. Con una altura de 3400 m.s.n.m. ....	83

## Índice de tablas

<b>Tabla 1.</b> Métodos para calcular la incertidumbre según el Método de Frazier. ....	33
<b>Tabla 2.</b> Incertidumbre por precisión de las coordenadas usando el sistema WGS-84 para 0° latitud (Basado en Wieczorek et al. 2004). ....	36
<b>Tabla 3.</b> Ejemplos del cálculo de incertidumbre por escala del mapa según estándar United States Geological Survey. ....	37
<b>Tabla 4.</b> Cálculo de incertidumbre por escala del mapa según error gráfico del IGAC .....	37
<b>Tabla 5.</b> Elementos para la georreferenciación y definición según los lineamientos del estándar Darwin Core. ....	43
<b>Tabla 6.</b> Elementos para documentar los productos de la estandarización y georreferenciación según el estándar Darwin Core (TDWG 2011). ....	44
<b>Tabla 7.</b> Reglas generales para estandarizar localidades de registros biológicos. ....	46
<b>Tabla 8.</b> Reglas específicas para estandarizar localidades. ....	47
<b>Tabla 9.</b> Niveles de calidad de las descripciones en localidades de registros biológicos. ....	49
<b>Tabla 10.</b> Casos frecuentes en el proceso de verificación y parámetros de incertidumbre. ....	58
<b>Tabla 11.</b> Parámetros de incertidumbre potenciales según los niveles de calidad .....	60
<b>Tabla 12.</b> Campos y formato de validación para ejecutar el script de verificación geográfica de registros biológicos. ....	91

## **INTRODUCCIÓN**

La toma de decisiones en conservación y la planeación ambiental de manera integral, requiere de la exploración y documentación de la biodiversidad del país, donde los datos primarios depositados en colecciones biológicas y museos juegan un papel preponderante como fuente de información (Guisan & Thuiller 2005; Overton et al. 2002; Gómez 1992). Las colecciones biológicas se pueden considerar como un archivo histórico natural de utilidad múltiple donde la información de cada registro puede ser la base de estudios taxonómicos, sistemáticos, ecológicos, filogenéticos, biogeográficos, genética de poblaciones y conservación (Tobar 2002).

Cuanto mayor y más precisos sean los datos, mayor será la capacidad para enfrentar diversos problemas (Muñoz et al. 2004). En este sentido, la georreferenciación entendida como el proceso de asignación de coordenadas espaciales a la descripción textual de un sitio, logra integrar un valor espacial que facilita la interpretación de la ocurrencia de los registros biológicos en el tiempo y en el espacio; posicionando a las colecciones como una fuente confiable de información primaria.

El método radio-punto, como método de georreferenciación, fue revisado y aplicado en los conjuntos de datos institucionales, para evaluarlo e incluir adaptaciones que se acoplen al estado actual de las bases de datos de las colecciones biológicas del país. La metodología además de asignar coordenadas a las localidades de registros biológicos, plantea la estimación de un valor de incertidumbre asociado al estado de los datos de la localidad, facilitando al investigador discriminar los datos.

Probablemente el más importante de los aspectos de la georreferenciación es la posibilidad de reproducir los procesos y un esfuerzo sustancial en la preparación de los datos, post-procesamiento y comprobación de la calidad (Murphey et al. 2004). En este sentido, se pretende abordar con detalle diferentes procedimientos desde el manejo de la base de datos, georreferenciación según los tipos de localidad, valoración de la incertidumbre y directrices para reportar los insumos cartográficos; con el fin de suministrar las herramientas para ejecutar esta metodología.

Por esta razón, el Instituto Humboldt (IAvH) y el Instituto de Ciencias Naturales (ICN) de la Universidad Nacional de Colombia, se han dado a la tarea de evaluar las bases y los métodos

de georreferenciación con el desarrollo de criterios metodológicos para su ejecución, cumpliendo con estándares internacionales y las necesidades nacionales.

Este documento provee una guía de referencia para realizar la georreferenciación de localidades asociadas a datos primarios sobre biodiversidad, estableciendo una serie de procedimientos o etapas, que incluyen la estandarización de las localidades, clasificación según la calidad de la descripción de la localidad, procedimientos de georreferenciación, evaluación de incertidumbre y finalmente la validación de los datos. Además de generar una posición crítica del alcance de la georreferenciación retrospectiva de registros biológicos; condición que genera limitantes, pero que como método es probablemente el mejor acercamiento para localizar geográficamente los datos de las colecciones.

## ANTECEDENTES

La información primaria sobre biodiversidad custodiada en diferentes colecciones y museos en todo el mundo, es cada vez más disponible y accesible a través de plataformas web (Beaman et al. 2004; Graham et al. 2004; Guralnick & Neufeld 2005). A través de esta iniciativa global donde se promueve la disponibilidad de los datos, se ha visto la necesidad de contar con datos de calidad, donde, la georreferenciación juega un papel preponderante.

Entre las iniciativas que han incluido la georreferenciación, además de una discusión amplia sobre las limitaciones que pueden tener estos datos en diferentes aplicaciones se encuentran: *Global Biodiversity Information Facility* (GBIF), *Australia's Virtual Herbarium* (AVH), *The Biological Collection Access Service for Europe* (BioCASE), *European Natural History Specimen Information Network*, *Fisnet* y *HerpNet*. Estas son solo algunas de las plataformas que desde hace más de una década se han dedicado a promover la publicación y acceso de los datos (Graham et al. 2004), al igual que la iniciativa VertNet financiada por *National Science Foundation*. A nivel nacional, el SiB Colombia ha promovido la gestión y acceso de los datos de biodiversidad.

Entre las contribuciones más significativas para definir las directrices de la georreferenciación, se encuentran las propuestas de John Wieczorek (2001) del Museo de Zoología de Vertebrados de la Universidad de California en Berkeley y Paul C. Murphey et al. (2004) del Museo de Historia Natural de la Universidad de Colorado, en el marco del proyecto "Mountain and Plains Spatio-Temporal Database-Informatics Initiative (MaPSTeDI)" (Graham et al. 2004)

De estos planteamientos han surgido iniciativas mundiales para afrontar los retos de la georreferenciación y calidad de datos espaciales. En el 2001, *National Science Foundation* financia la iniciativa *Mammal Networked Information System* (MaNIS) que además de la georreferenciación incluye como parte integral de este proceso el reporte del dato de incertidumbre para lo cual se creó la calculador MaNIS con el fin de semi-automatizar su cálculo (Wieczorek 2001; Beaman et al. 2004), facilitando a los usuarios identificar los datos según las necesidades de resolución espacial (Wieczorek 2001). Posteriormente surgen otras propuestas direccionadas a automatizar los procesos, tales como BioGeoMancer (Beaman 2003) y GEOLocate (*Tulane University Biodiversity Research Institute*), aplicaciones que contienen herramientas para que los usuarios tengan acceso directo a la georreferenciación mediante

una interface de edición visual y una base de datos de topónimos para el rastreo de las localidades.

En Latinoamérica, uno de los procesos más trabajados ha sido el propuesto por la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) de México. Esta institución en el 2008 presenta un compendio de procedimientos para la georreferenciación, considerando los parámetros de incertidumbre planteados por el método radio-punto con algunas modificaciones. En Colombia, estos esfuerzos han sido puntuales en algunas colecciones del Instituto de Ciencias Naturales de la Universidad Nacional de Colombia y el Instituto Humboldt. En 1999, el Instituto Humboldt a través de la Unidad de Sistemas de Información Geográfica (UNISIG) y el programa de Inventarios de la Biodiversidad, inició un proceso de georreferenciación de localidades de registros biológicos y desarrolló un aplicativo para la captura y manejo de localidades para conformar un gacetero de localidades biológicas (Sua et al. 2004). Sin embargo, a raíz de la necesidad de contar con datos georreferenciados que documenten su estado de confiabilidad, se elaboró una revisión de la aplicación del Método Radio-punto, junto con modificaciones para dar solución al tratamiento metodológico ante la gran heterogeneidad de localidades presentes en las colecciones biológicas.

Por tal razón, el Instituto Humboldt (IAvH) y el Instituto de Ciencias Naturales (ICN) de la Universidad Nacional de Colombia, como entidades líderes en la gestión de información sobre biodiversidad en Colombia, suman esfuerzos para consolidar un protocolo de georreferenciación que se ajuste a las necesidades y limitaciones de las colecciones del país cumpliendo con estándares globales, asegurando calidad, uso y reúso de la información.

## **CAPÍTULO 1.**

# **BASES GENERALES DE LA GEORREFERENCIACIÓN DE REGISTROS BIOLÓGICOS**

## 1.1 ¿Que es la georreferenciación de localidades de registros biológicos?

El proceso de convertir la descripción textual de un sitio donde ha sido colectado un espécimen a coordenadas geográficas se conoce como georreferenciación (Wieczorek 2001; Proctor et al. 2004; Murphey et al. 2004). La ubicación es registrada de manera descriptiva, tal como, **“35 Km Vía Mocoa-Pitalito, Mocoa (Putumayo)”**; la cual no podría ser mapeada en los Sistemas de Información Geográfica (SIG), ni utilizada para hacer análisis espaciales. Por lo tanto, se requiere que la descripción sea traducida en coordenadas de latitud y longitud (véase Fig. 1). Este proceso de interpretación se puede considerar como georreferenciación retrospectiva (Proctor 2004), ya que no son coordenadas tomadas en campo sino de un proceso posterior.



**Figura 1.** Localidad georreferenciada de un registros biológico a partir de los datos de etiqueta. Localidad: 6 km al Norte de Mariquita a lo largo de la ruta a Victoria, cerca al río Gualí.

Con el avance en el uso de Sistemas de Posicionamiento Global (GPS) en las exploraciones biológicas, la localidad tiene coordenadas geográficas que facilitan su rastreo en cartografía digital. Sin embargo, raras veces vienen acompañadas del sistema de referencia configurado en el receptor GPS y de la precisión reportada por el dispositivo, para lo cual, la

georreferenciación entra a verificar la correspondencia de las coordenadas con respecto a la descripción de localidad y establecer el grado de incertidumbre para validar los datos.

Es importante que cada uno de estos procesos aseguren la exactitud de la ubicación del registro sin incrementar el error (Munt 2013), con el fin de valorar la confiabilidad de la localidad georreferenciada.

## 1.2 ¿Por qué es importante georreferenciar los registros de las colecciones biológicas?

Dentro de la problemática ambiental, gestión y conservación de recursos biológicos, la principal fuente de datos son las colecciones biológicas, no solo por su conjunto amplio de datos sino por la resolución temporal producto de las expediciones y trabajo en campo de los investigadores a través de la historia. De esta manera, se constituyen en una fuente de datos verificable de la ocurrencia de un espécimen en el tiempo (Proctor et al. 2004). Sin embargo para hacer uso de estos datos, se requiere de coordenadas geográficas en cada uno de los registros, constituyéndose en uno de los mayores retos que pueden tener las colecciones biológicas (Baker et al. 1998).

De acuerdo con el Registro Único Nacional de Colecciones Biológicas (RNC), existen 203 colecciones biológicas en el país que albergan cerca de **4,7 millones de ejemplares** principalmente de colecciones zoológicas y herbario<sup>1</sup>. La disponibilidad de estos registros georreferenciados facilitaría el análisis espacial de la biodiversidad además de generar un valor agregado a las colecciones en cuanto a la calidad de los datos que se enmarcan en los siguientes logros:

- ✓ Dar un valor geográfico a registros biológicos que anteriormente solo contaban con la descripción de la localidad.
- ✓ Posicionar las colecciones como una fuente de información confiable para análisis espaciales en el país.
- ✓ Documentar la incertidumbre de los datos permitiendo a los usuarios identificar los datos según su interés.

---

<sup>1</sup> <http://www.humboldt.org.co/servicios/registro-unico-nacional-de-colecciones-biologicas-rnc>

### **1.3 La georreferenciación en la estructura de trabajo de las colecciones: Una aproximación desde el IAvH**

Idealmente, una colección biológica requiere del trabajo dinámico y articulado entre curadores, técnicos y administradores de datos, quienes se encargan de mantener los especímenes y gestionar la información en la colección. Sin embargo, no toda colección tiene un administrador de datos y en ese sentido la recomendación primordial para asegurar la integridad de la base de datos, es mantener el contenido original y un buen manejo de la información generada en la georreferenciación. De lo contrario, una colección tendrá un valor limitado (Cato 1991).

En la colección del IAvH, la información asociada al espécimen es sistematizada en *Specify* por los auxiliares de la colección y hay un administrador de datos quién verifica la compatibilidad y calidad para ser integrada a la base de datos (véase Fig. 2).

Antes de empezar con la georreferenciación, se debe hacer el reconocimiento y evaluación del estado de la información contenida en los campos y verificar si existen datos duplicados, revisión que se hace sobre una copia de la base de datos.

En el proceso de georreferenciación de las localidades es fundamental mantener una sinergia continua con los curadores, debido a la existencia de datos que requieren una revisión de la información contenida en la etiqueta de los ejemplares, sobre todo aquellos con información dudosa.

Una vez las localidades han sido georreferenciadas, se ejecuta la migración de los datos georreferenciados asegurando siempre la integridad de estos.

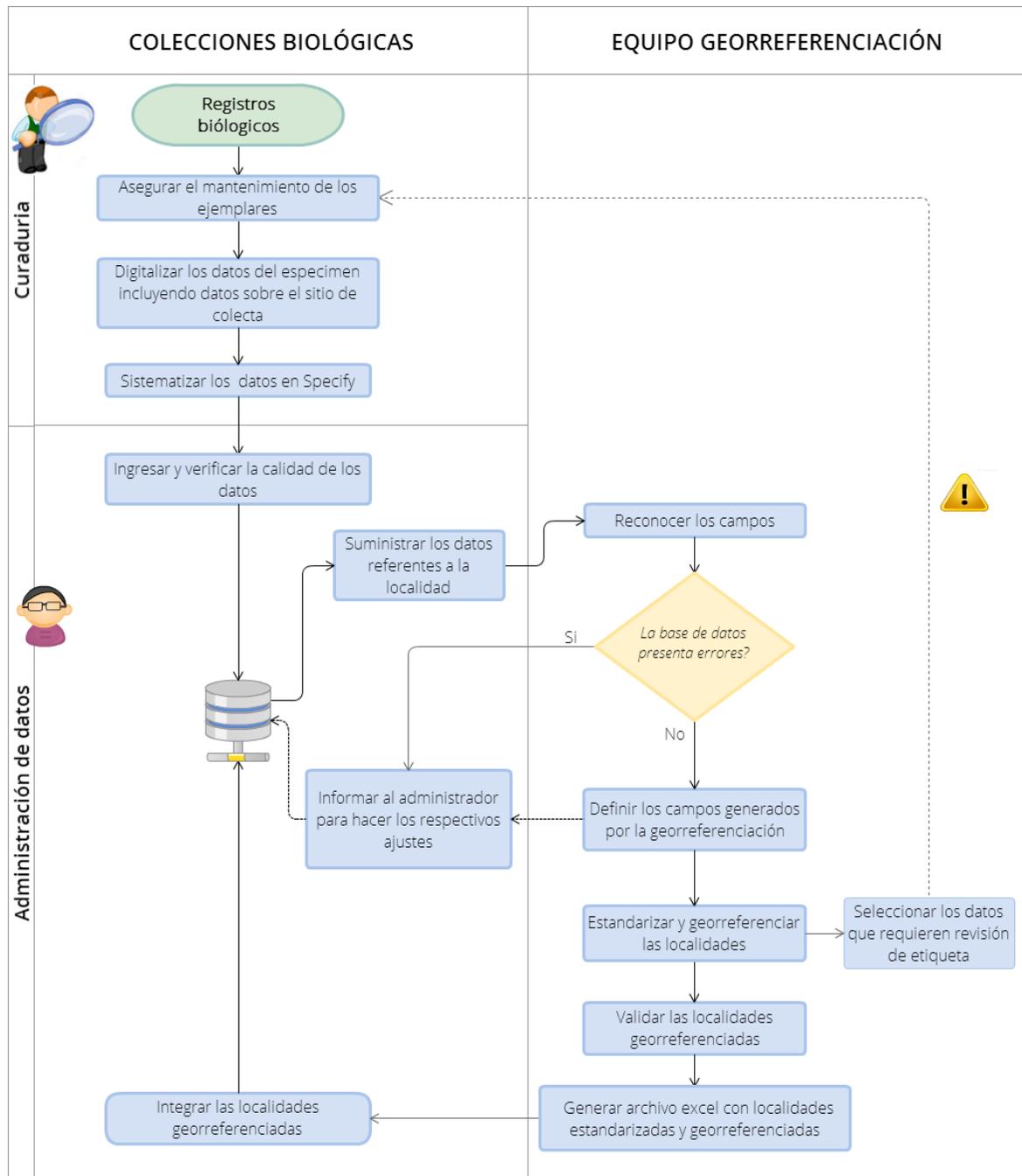


Figura 2. La georreferenciación dentro de la dinámica de trabajo de las colecciones biológicas de IAVH.

## 1.4 Factores que inciden en el estado de la localidad

La localidad es el punto de partida para la georreferenciación de registros biológicos, de ahí la importancia de tener clara su definición. La localidad es definida como la posición de una entidad en el espacio (Chapman 2005), o la representación verbal de su posición (Wieczorek 2001). Una definición más precisa para la georreferenciación es la propuesta por el estándar *Darwin Core*, donde la localidad es la descripción específica del lugar de colecta complementada por la geografía superior (país, departamento, municipio), altura, coordenadas, entre otros.

El estado de las localidades asociadas a registros biológicos es muy heterogéneo debido a múltiples factores. Los más relevantes por incidir en la calidad del conjunto de datos de la localidad son los siguientes:

- ✓ Percepción del colector
- ✓ Sistemas de Posicionamiento Global (GPS)
- ✓ Cambio temporal de las entidades geográficas

### 1.4.1 Percepción del colector

La descripción de una localidad donde fueron colectados los especímenes involucra cierto grado de subjetividad (CONABIO 2008). Las descripciones se reportan de maneras muy distintas, condicionadas por la percepción que tiene el colector del sitio de colecta, las necesidades respecto a los datos y las herramientas disponibles en campo (mapas, brújulas, GPS, etc.), influyendo en la calidad de la descripción. Esto ha generado una alta heterogeneidad en la calidad de los datos como registros sin datos de localidad que tienen como únicos referentes espaciales las entidades geográficas de mayor jerarquía. También es frecuente el uso de entidades que no se mantienen en el tiempo y no son verificables en la cartografía, tales como hábitat, tipos de ecosistemas, propiedades de particulares, etc.

Otro aspecto es el desconocimiento de los métodos usados para cierto reporte de datos, como distancias y orientaciones, que frecuentemente son una estimación del colector, condición que genera incertidumbre sobre su cálculo.

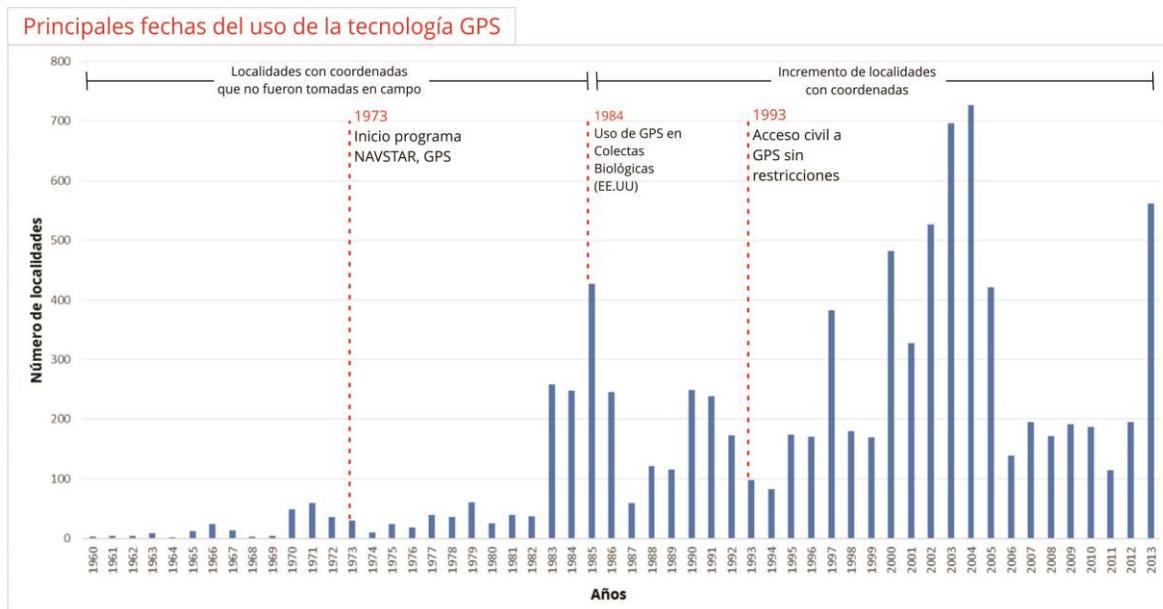
### **1.4.2 Uso de Sistema de Posicionamiento Global (GPS) en la localización de registros biológicos**

El Sistema de Posicionamiento Global (GPS) tuvo su aparición en 1973, pero solo hasta mediados de los 80's se tuvo acceso para uso civil con algunas restricciones en la precisión, solo hasta el año 1993 se declaró totalmente operable para el uso civil ( Bolfor & Etsfor 1999). Para exploraciones biológicas se tiene reportado que hasta mediados de 1980, el GPS comenzó a ser usado para registrar las coordenadas geodésicas del sitio de colecta (Proctor 2004).

Con el fin de evaluar la incidencia del GPS en el estado de las localidades con coordenadas. Se realizó un análisis temporal de la colección del IAvH desde 1960 hasta 2013. Se encontró que desde la disposición civil del GPS, hay un incremento en el número de localidades con coordenadas, aunque esto no necesariamente implica que todas las localidades hayan sido capturadas con GPS. Hay que tener en cuenta que para Colombia probablemente el acceso al GPS fue más limitado que en el ámbito internacional (véase Fig. 3).

También se identificaron localidades con coordenadas entre las décadas de los 60's y 70's, periodos donde no estaba a disposición el GPS, permitiendo inferir que la asignación de coordenadas fue posterior a la fase de campo y que solo se corrobora revisando las etiquetas o catálogos de los registros biológicos. Aunque no se puede desconocer el uso probable de mapas, un insumo frecuentemente en las salidas de campo pero que difícilmente puede tener la precisión obtenida por un GPS.

La georreferenciación en campo es una práctica estándar por la mayor parte de los colectores y junto con la descripción de la localidad, es la mejor combinación que confiere exactitud y precisión (Rhoads & Thompson 1992). Sin embargo, se ha generado la equivocada noción que solo con la captura de las coordenadas en terreno es suficiente, dejando a un lado la actividad de describir el sitio de colecta.



**Figura 3.** Temporalidad de las localidades con coordenadas en las colecciones biológicas del Instituto Humboldt. Las líneas punteadas indican las fechas de uso de la tecnología GPS y la influencia de este en el aumento de registros con coordenadas.

Efectivamente la disponibilidad del GPS, influyó en el incremento de localidades con coordenadas geográficas, permitiendo un rastreo cartográfico más eficaz del sitio de colecta. Sin embargo, factores como el desconocimiento del método de captura (GPS, mapa, gacetero), si fueron tomadas por GPS, la falta del reporte del datum y la precisión de las coordenadas confieren incertidumbre exigiendo una revisión de las coordenadas donde la descripción de la localidad permite realizar la verificación y sin esta es imposible.

Entre los factores que inciden en la incertidumbre de las localidades con coordenadas se encuentran los siguientes:

- ✓ Desconocimiento generalizado del método de captura de las coordenadas, impidiendo discriminar las coordenadas tomadas por GPS, de otras fuentes como gaceteros, mapas, Google Earth, etc.
- ✓ Desconocimiento del datum usado en la captura de coordenadas.
- ✓ Registros que solo tienen coordenadas y no tienen descripción de la localidad.

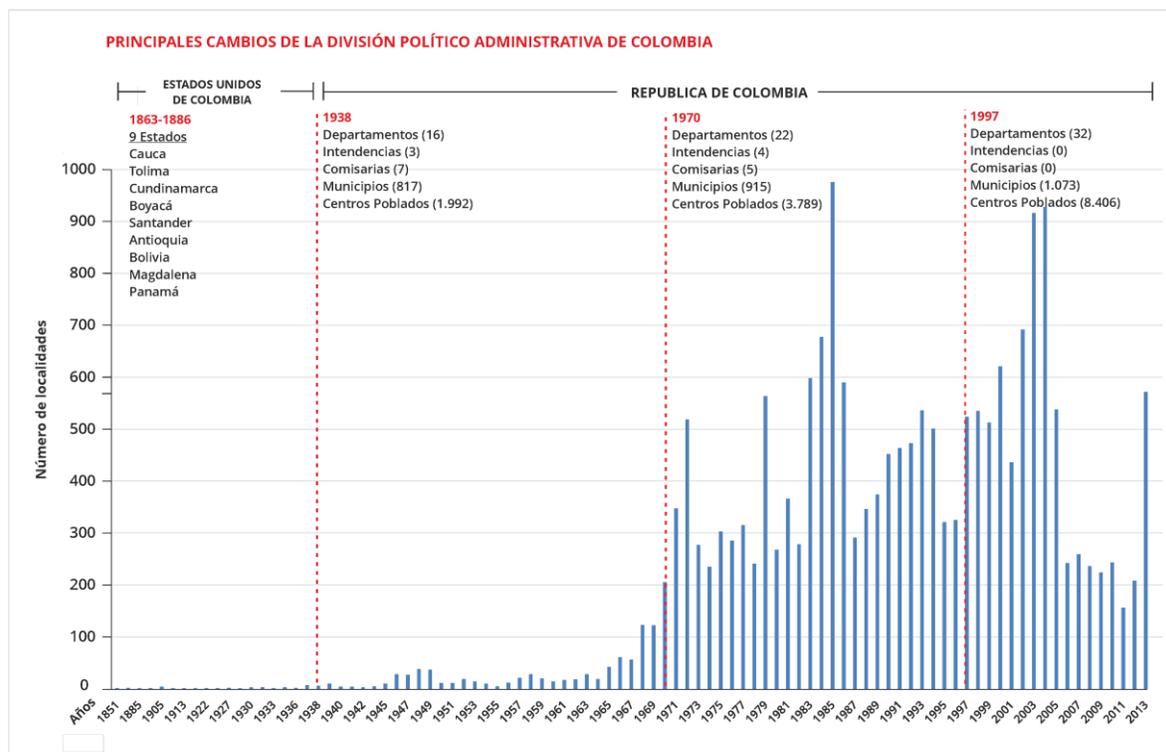
### 1.4.3 Cambio temporal de las entidades geográficas

Un registro para una colección es la evidencia de la presencia de un organismo en un tiempo y espacio determinado. Para evaluar la localidad de un registro biológico se requiere ubicar el espacio donde se colectó. Sin embargo, se debe tener en cuenta que las entidades geográficas cambian en el tiempo y probablemente en la actualidad esas entidades ya no cumplan con las mismas características que el colector utilizó para definir la localidad. Estos cambios pueden ser más relevantes a medida que la localidad es más antigua y cuyo rango temporal puede datar desde los inicios de las colecciones biológicas en el país.

En Colombia, desde finales del siglo XVII las ciencias naturales recibieron un gran adelanto marcado por los naturalistas Mutis (1783-1808), Triana (1850s), Caldas (1769-1816), Alexander von Humboldt (1769-1859), Jean-Bastiste Boussingault (1802-1887), José Manuel Restrepo (1781-1863), algunos más contemporáneos como: Cuatrecasas (1940s), Vergara y Velasco (1920s) entre otros, quienes alimentaron en sus inicios lo que hoy son las colecciones.

El rastreo cartográfico para estas localidades en registros antiguos puede verse limitado, debido a la escasa información en estos ejemplares (Rojas et al. 2002). Por lo tanto, entre más antigua sea la localidad, más difícil será encontrar los insumos cartográficos e identificar el cambio político-administrativo de la fecha con respecto al actual. Seguramente para las localidades más antiguas, el historial del colector, revisión de mapas de la época, bitácoras y el espécimen mismo sean herramientas que ayuden a la georreferenciación de estos datos. Estas localidades históricas requieren un mayor esfuerzo en la georreferenciación ya que permiten visualizar el valor histórico de las colecciones y sobretodo estos datos son la evidencia de la biodiversidad del pasado, probablemente de ecosistemas que ya no existen.

En la Figura 4, se ha retratado la temporalidad de los datos de localidad de los registros biológicos del IAvH que datan desde 1854, junto con las fechas cuando hubo fuertes cambios de la división-político administrativa y que podrían usarse como parámetro para la revisión de la equivalencia actual.



**Figura 4.** Temporalidad de los datos de localidad de los registros biológicos del IAvH desde el año 1851 hasta 2013. Las líneas punteadas marcan los principales cambios de la división político administrativa (DANE 2001), factor incide en la información de la localidad.

## 1.5 Lineamientos para el manejo de bases de datos en la georreferenciación

A continuación, se presentan los criterios para el manejo de las bases de datos que se enmarcan dentro de las buenas prácticas para asegurar la integridad de la base de datos y lograr mayor operatividad en la georreferenciación.

### 1.5.1 Reducir la cantidad de datos de localidad a procesar

La georreferenciación opera sobre los datos específicos de ubicación del sitio de colecta, donde cada localidad tiene un identificador (ID) que define al dato como único. Por lo tanto, no debería haber una localidad con los mismos datos en todos sus campos. Sin embargo, es posible que haya duplicidad generada en el momento de realizar la sistematización o migración por diferencias menores (ortografía, sintaxis, etc.).

La duplicidad aumenta el tiempo dedicado a la georreferenciación. Por tal razón, es necesario hacer un diagnóstico preliminar para identificar la duplicidad y reducir la dimensionalidad de los datos, mediante la aplicación de las correcciones pertinentes. En el caso particular de las colecciones del IAvH, el 21 % de las localidades de la base de datos contenía localidades duplicadas, que al ser corregidas redujeron con certeza varios meses de trabajo.

### **1.5.2 Mantener los campos que contienen los datos de la localidad original**

Todo proceso de georreferenciación debe mantener siempre los campos que contienen la información de la localidad original, es decir, aquella que refleja el contenido de la etiqueta. Esto permite enriquecer la base de datos y tener elementos de chequeo con la información original (GBIF 2008).

Una alternación de la información original implica inversión de tiempo y recursos para la revisión de las etiquetas y catálogos antes de iniciar el proceso de georreferenciación. Para esto se recomienda que cualquier resultado de un proceso de georreferenciación, se almacene en campos adicionales y no trabajar sobre los campos originales.

### **1.5.3 Respetar la definición de los elementos**

Generalmente, las colecciones biológicas poseen una base de datos con una serie de elementos o campos estandarizados que contienen la información de los especímenes, en cuanto a la taxonomía, sitio de colecta, colector, etc. El estándar Darwin Core<sup>2</sup>, ampliamente usado para la estructuración e intercambio de datos primarios sobre biodiversidad, posee una serie de elementos que pueden usarse para almacenar la información producto de un proceso de georreferenciación. Se recomienda su implementación con el fin de mantener la información estandarizada e interoperable con iniciativas nacionales y globales. Adicionalmente, facilita la interpretación y uso de la información que se genera durante la georreferenciación.

---

<sup>2</sup> Elementos de Darwin Core: Una guía de referencia rápida.

[http://www.sibcolombia.net/c/document\\_library/get\\_file?p\\_l\\_id=33134&folderId=33501&name=DLFE-301.pdf](http://www.sibcolombia.net/c/document_library/get_file?p_l_id=33134&folderId=33501&name=DLFE-301.pdf)

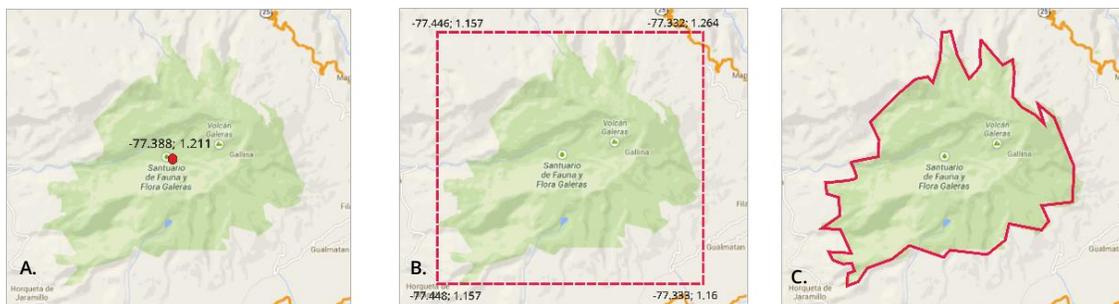
## **CAPÍTULO 2.**

### **MÉTODO RADIO - PUNTO**

## 2.1 Métodos de georreferenciación y Método Radio-Punto

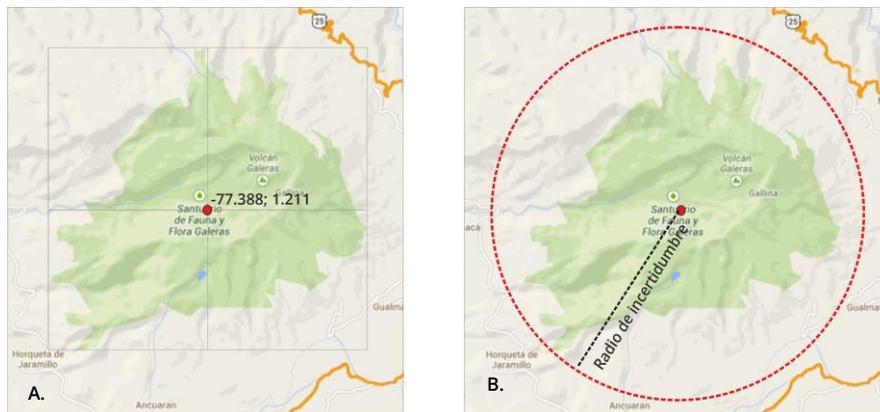
Varios métodos se han propuesto para georreferenciar registros biológicos. El método más conocido y asociado con la georreferenciación es el método-punto (*Point method*), el cual consiste en asignar un punto georreferenciado por localidad (Wieczorek et al. 2004). Sin embargo, este ignora la dimensión del área probable del sitio de colecta. El método polígono (*Shape method*) se puede considerar el más preciso de los métodos al plasmar la forma del área, aunque dentro de un proceso masivo de georreferenciación es poco operativo, ya que implica una combinación de formas para representar la entidad geográfica; además, por cada forma hay un conjunto de coordenadas cuyo almacenamiento en una base de datos es más conflictivo (Wieczorek et al. 2004).

Otro método que ha sido explorado, es el método cuadrado (*Bounding Box*) que requiere de cuatro vectores para definir el área, es más operativo pero menos preciso que el método polígono, además de ser complejo en el manejo de las bases de datos (véase Fig. 5) (Wieczorek et al. 2004).



**Figura 5.** Métodos de georreferenciación: A) Método punto B) Método cuadrado C) Método polígono.

El método radio-punto describe el sitio de colecta con un punto georreferenciado y alrededor de este una circunferencia para definir el sitio probable de colecta que integra la descripción de la localidad e incertidumbres asociadas (Wieczorek 2001; CONABIO 2008). Desde el punto de vista de manipulación de información, es más operativo para el manejo de la base de datos, al requerir solo un punto georreferenciado (latitud, longitud) y eficaz debido a que la incertidumbre va integrada en un solo atributo definido en una longitud en metros (véase Fig. 6).



**Figura 6.** Método Radio Punto A) Cálculo del centroide B) Cálculo de la incertidumbre.

## 2.2 Incertidumbre de los datos geográficos según el Método Radio-Punto

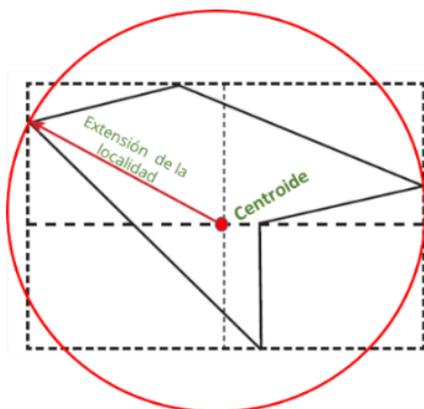
El cálculo de la incertidumbre provee una disposición clave para determinar el uso de los datos y calidad de los mismos (Chapman & Wieczorek 2006). Como se dijo anteriormente, el método radio-punto tiene una opción simple para registrar el máximo error mediante el uso de un solo atributo expresando en longitud en metros. Se tienen en cuenta 6 parámetros, todos con la capacidad de ser cuantificados y cuyos valores se suman para que la incertidumbre sea expresada en un solo valor. Estos parámetros han sido ampliamente documentados (Wieczorek et al. 2004), sin embargo se considera pertinente hacer una revisión de los mismos para abordar algunas adaptaciones que confieren mayor operatividad en su uso; dependiendo del tipo de localidad se definen los parámetros a utilizarse, sin embargo esto se abordará a profundidad en el capítulo 4 de este documento.

- ✓ Extensión de la localidad de referencia
- ✓ Desconocimiento del Datum
- ✓ Imprecisión en la medida de la distancia
- ✓ Incertidumbre por orientación
- ✓ Incertidumbre por coordenadas
- ✓ Escala del mapa

### 2.2.1 Extensión de la localidad de referencia

Todo lugar es un espacio finito que puede ser medido en cuanto a su extensión. Cuando la descripción de la localidad reporta una o varias entidades geográficas, estas se interpretan para ubicar el sitio potencial de colecta delimitado por una circunferencia que se extiende hasta el límite más lejano. Los referentes que se pueden encontrar son cabeceras municipales, veredas, corregimientos, lagunas, quebradas, caños, ríos, etc.

Frecuentemente los puntos de referencia utilizados en las descripciones de localidad pueden cambiar en el tiempo (casa de correo, palacio de justicia, plaza principal, tamaño de la ciudad o pueblo) y probablemente los mapas actuales no reflejen la extensión de esos lugares donde se tomaron las muestras (Chapman & Wieczorek 2006). Por tal motivo, se considera el centroide de la entidad geográfica el punto de partida para medir la incertidumbre y que se extiende hasta el límite más lejano de la entidad (véase Fig. 7).



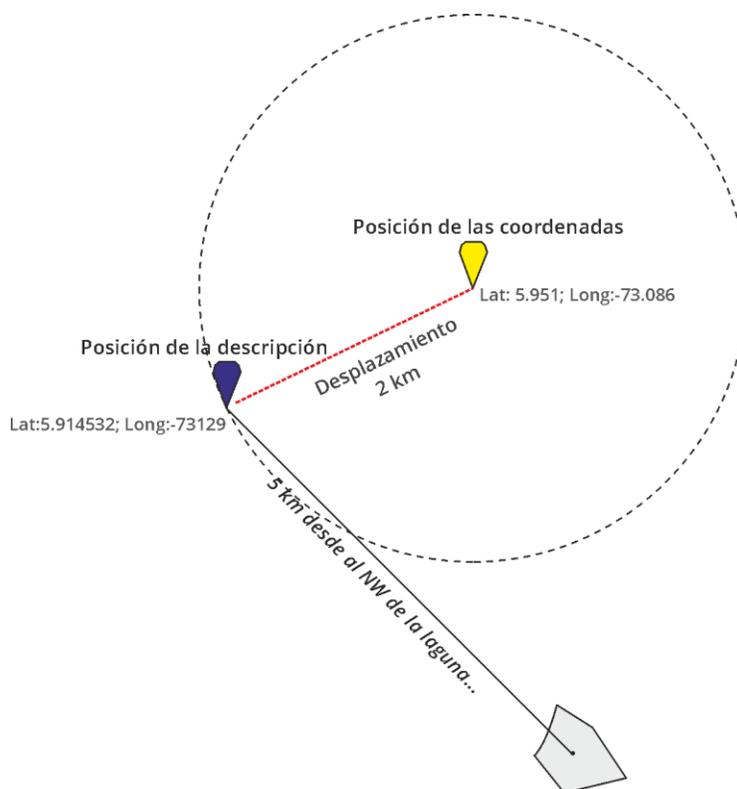
**Figura 7.** Esquema de la incertidumbre por extensión de una entidad geográfica definida por un polígono.

La interpretación de la incertidumbre por extensión puede ser más difícil al tratarse de localidades que tienen una forma irregular (por ejemplo: un camino o un río sinuoso), donde, la extensión es la distancia entre el punto medio hasta el punto más alejado de los extremos del vector (Chapman & Wieczorek 2006) incluyendo una amplia área que no corresponde al sitio más probable de colecta. Para esto se recomienda limitar el área usando la altura o intersectando el área con la división político-administrativa más específica (véase Fig. 8).



**Figura 8.** Centroide de un río y extensión de incertidumbre.

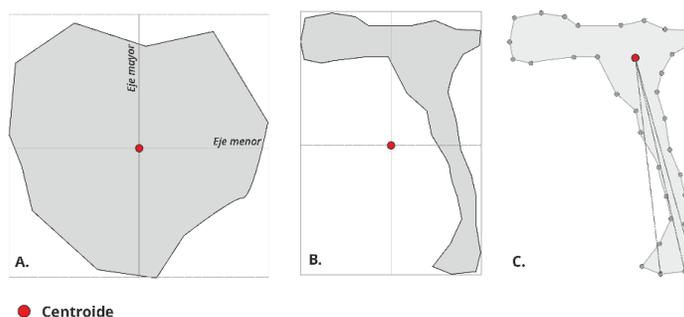
Una adaptación de este parámetro se presenta en el proceso de verificación para las localidades que tienen coordenadas, donde se hace la evaluación de la correspondencia entre la coordenada y la descripción de la localidad. Se aplica cuando el punto dado por las coordenadas y el punto definido por la descripción son próximos (véase Fig. 9), siendo la distancia que los separa el valor de incertidumbre por extensión.



**Figura 9.** Extensión de incertidumbre en localidades con coordenadas.

## Concepto de centroide

Un centroide es el centro geométrico de un polígono. La manera clásica para el cálculo del centroide se basa en trazar un recuadro que abarque la entidad geográfica, identificar eje mayor y menor, y en la intersección de ambos ejes se obtiene el centroide (CONABIO 2008). Este método tiene un limitante para entidades geográficas de forma irregular ya que el centroide se puede ubicar por fuera del área (véase Fig.10). Una manera más operativa y precisa es mediante el cálculo automático del centroide con herramientas SIG, las cuales determinan la localización del centroide a partir de triangulaciones con los límites más lejanos de la entidad geográfica. En el capítulo 5 de este documento se explica este procedimiento haciendo uso de herramientas con el software libre Quantum-GIS.



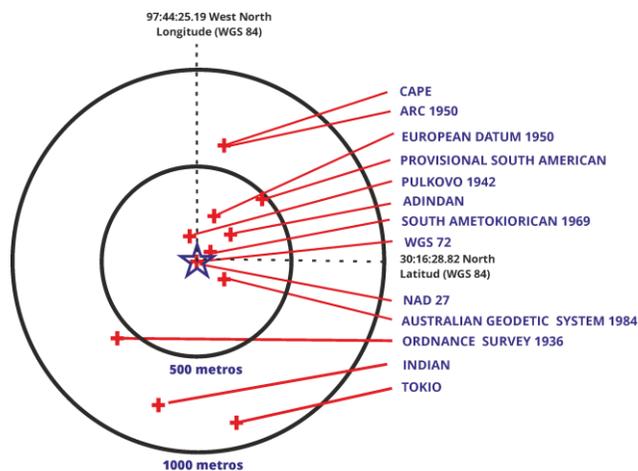
**Figura 10.** Cálculo del centroide. A) Centroide calculado con el eje mayor y menor. B) Cálculo del centroide de un polígono irregular. C) Centroide calculado por un proceso automatizado.

### 2.2.2 Desconocimiento del Datum

Rara vez las coordenadas geográficas están registradas junto con la información de datum geodésico, incluso en aquellas localidades que han sido tomadas con GPS (Chapman & Wiecek 2006). El desconocimiento de este dato genera imprecisión en la localización de las coordenadas añadiendo un error adicional a la georreferenciación (véase Fig. 11). Por lo tanto, una misma coordenada capturada con GPS y visualizada en cartografía con diferentes datum puede presentar desplazamientos de varios cientos de metros.

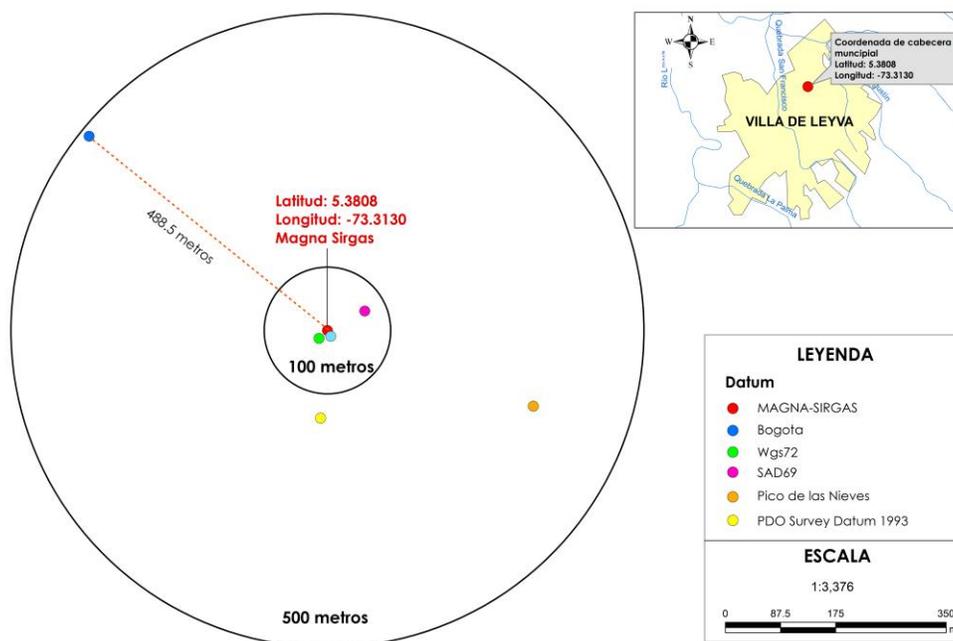
En Colombia, los datum más utilizados para captura de coordenadas son: WGS-84 (World Geodetic System, 1984), MAGNA SIRGAS (Marco Geocéntrico Nacional de Referencia, densificación en Colombia del Sistema de Referencia Geocéntrico para las Américas) y Datum

BOGOTÁ. La diferencia entre MAGNA SIRGAS y WGS 84 es mínima (~2cm), ya que la exactitud de las coordenadas es compatible. Sin embargo, entre MAGNA SIRGAS y Datum BOGOTÁ existe un desplazamiento de aproximadamente 500 metros en la posición de las coordenadas de un mismo punto (IGAC 2004).



**Figura 11.** Cambio en la localización de las coordenadas usando diferentes datum (Wieczorek et al. (2001).

Se realizó un análisis del valor del desplazamiento en metros de las coordenadas usando los datum más comunes para Suramérica y Colombia. Para Colombia el mayor valor de desplazamiento se presenta entre el datum MAGNA-SIRGAS y el antiguo datum BOGOTA, registrándose 485.5 metros de diferencia entre ambos, un valor aproximado a los 500 metros establecidos por el Instituto Geográfico Agustín Codazzi (véase Fig. 12)



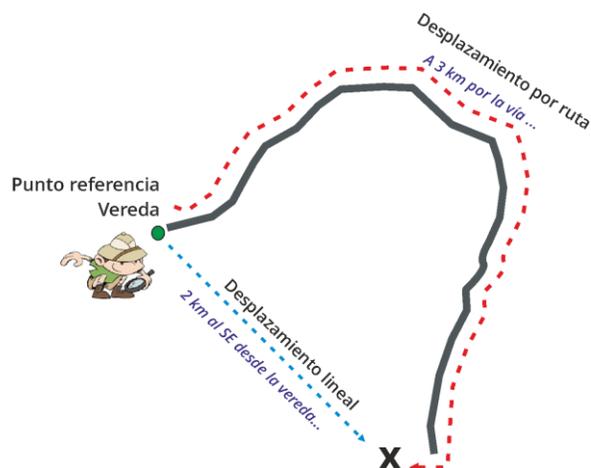
**Figura 12.** Desplazamiento en metros de una misma coordenada con diferentes datum para Colombia (MAGNA-SIRGAS y Bogotá) y algunos datum para Suramérica (Wgs 72, SAD69, Pico de Nieves y PDO Survey Datum 1993).

Teniendo en cuenta lo anterior se establece un valor de incertidumbre de 500 metros por incertidumbre por datum para las coordenadas sin información del datum. Se espera que a medida que se mejoren las buenas prácticas en la capturar de coordenadas, entre ellas el reporte del datum, los nuevos registros que ingresen tengan una incertidumbre de 0 km.

**Nota:** Si un registro biológico presenta una adecuada captura de coordenadas y reporte de todos los datos informativos del receptor GPS, no es necesario someterlo al proceso de georreferenciación, únicamente se da paso a la verificación de las entidades político-administrativas.

### 2.2.3 Imprecisión en la medida de distancia

El uso de la distancia en la localidad generalmente está asociado a una ruta u orientación, por ejemplo: *A 3 km por la vía hacia Túquerres* o *a 2 km al SE de la Vereda El Rosal* (véase Fig. 13). Generalmente se desconoce el método de cálculo, que puede ser por percepción que tiene el colector sobre la distancia recorrida, sea lineal o por una vía, y poco frecuente el empleo de un equipo especializado para su medición. Esta condición genera una disminución o aumento de la distancia reportada y que se interpreta como incertidumbre.



**Figura 13.** Uso de la distancia para describir un trayecto en campo.

Existen dos métodos para calcular la incertidumbre por distancia, uno propuesto por Wieczorek et al. (2004) y otro por Frazier et al. (2004).

El método de Wieczorek plantea evaluar la distancia según su grado de precisión mediante la presencia de número decimales y su proporción en una fracción, para distancia grandes con múltiplos de potencia 10 establece una relación de 0.5 veces como potencia de incertidumbre. Este método tiene limitantes, en la figura 13 se muestra una regla con diferentes valores, observe que la distancia de 140 millas al ser convertida a kilómetros se transforma en una cifra con decimales, generando una incertidumbre con un nivel de precisión mucho más alto. Por otra parte la incertidumbre en un valor grande al tener decimales disminuye significativamente, si se tiene en cuenta que en la mayoría de los casos la distancia es tomada por percepción, es de esperar que valores grandes contenga un error más significativo que distancias más pequeñas (véase Fig. 14).

La propuesta sugerida por Frazier et al. (2004), recomienda que para distancias con múltiplos de 10 (10, 100, 200) o fracciones como 25 o 75, se puede usar el 15% del valor de la distancia como incertidumbre (Chapman & Wieczorek 2006), independientemente del tipo de unidad sean metros, kilómetros, millas, etc. (véase Tabla 1). Este método es mucho más operativo, ya que no es afectado por una conversión de la unidad de medida y mantiene una coherencia entre las distancias grandes y pequeñas, de ahí que se recomiende este método para su cálculo.

MÉTODO POR WIEKZORECK



Figura 14. Cálculo de incertidumbre por distancia según el Método de Wiekzoreck et al (2004).

Tabla 1. Método para calcular la incertidumbre según el Método de Frazier.

Valor de distancia	Método de Frazier 15% de la unidad
2000 m	300 m
9.5 km	1.4 km
20 km	1.5 km
30 km	4.5 km
140 mi	33.8 km
200 mi	48.28 km

### 2.2.4 Incertidumbre por orientaciones

Los puntos cardinales son un parámetro de uso frecuente en la orientación espacial por parte de los colectores. Sin embargo, dependiendo de su especificidad hay un rango de variación hacia las orientaciones próximas. En este sentido, la incertidumbre por orientación se interpreta como la mitad de la diferencia en grados entre las orientaciones cercanas.

#### Orientaciones cardinales primarias

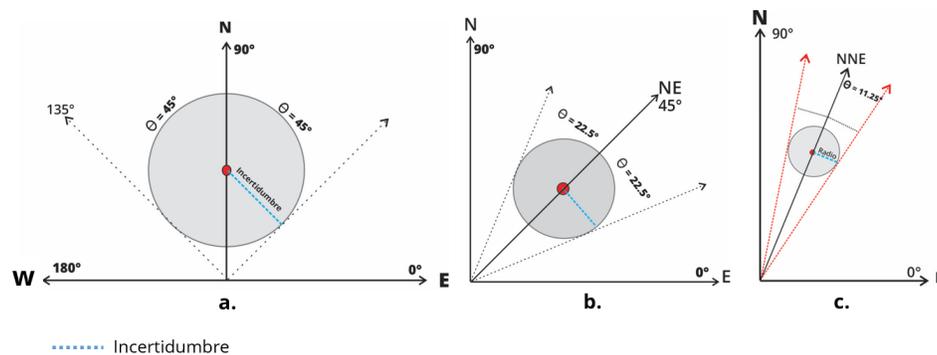
Este tipo de orientaciones (N, S, E, W) manejan una amplia área de incertidumbre con un ángulo de 45° a los dos lados desde la entidad geográfica. Si la orientación es *Norte* o *Sur* existe la probabilidad de que el sitio de colecta se haya desviado al *Este* u *Oeste*, de la misma manera, si la orientación es *Este* u *Oeste* el rango de variación ocurre entre el *Norte* o *Sur* (véase Fig. 15).

### Orientaciones intercardinales primarias

Este tipo de orientaciones (NE, SE; SW y NW) disminuyen el área de incertidumbre a  $25^\circ$  hacia las orientaciones próximas debido a la precisión de la orientación (véase Fig. 15).

### Orientaciones intercardinales secundarias

Las orientaciones intercardinales secundarias (NNE, ENE, ESE, SSE, SSW, WSW, WNW y NNW) son las direcciones medias entre los puntos cardinales y puntos intercardinales primarios con una diferencia de  $22.5^\circ$ , que equivalen a un incertidumbre de  $11.25^\circ$ . El uso de este tipo de orientaciones en campo es poco frecuente, ya que se requiere de una brújula o dispositivo digital de orientación para tener ese grado de incertidumbre (véase Fig. 15).



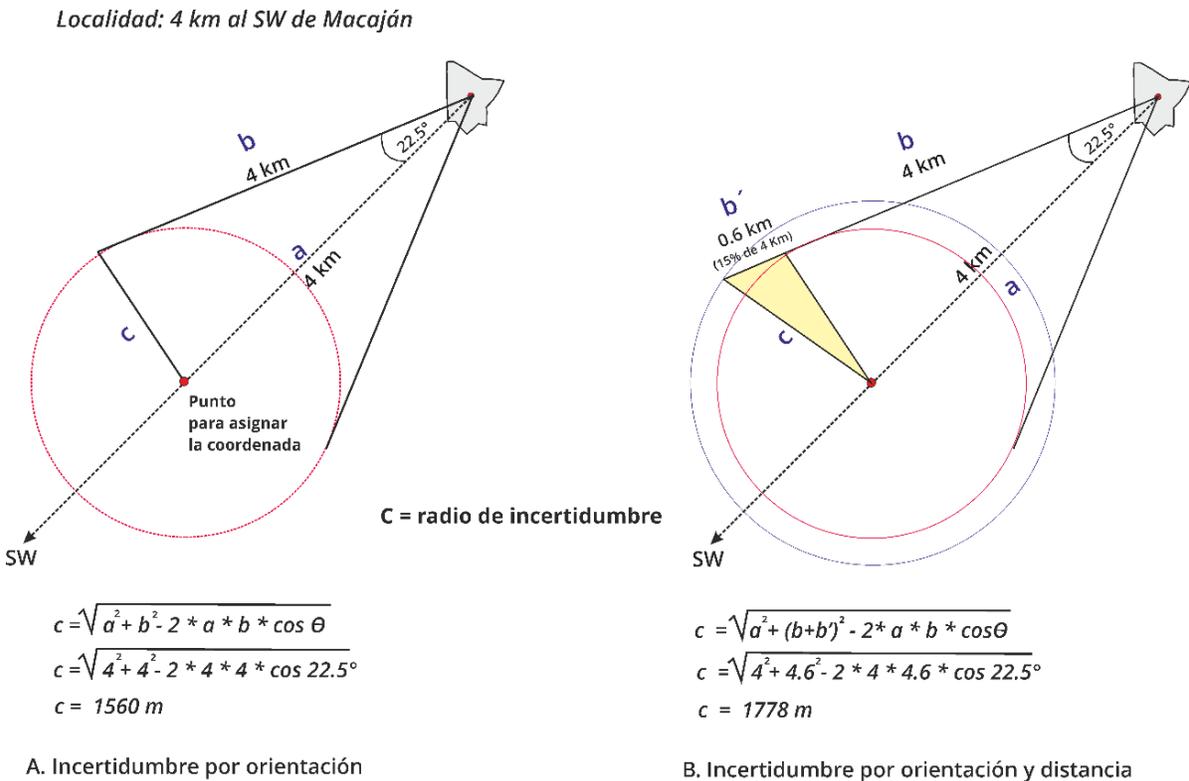
**Figura 15.** Incertidumbre por orientaciones: A) Orientaciones cardinales, B) Orientaciones cardinales primarias y C) Orientaciones intercardinales secundarias.

Para medir la incertidumbre en localidades cuya orientación está asociada a una distancia, se involucra tanto la distancia como la orientación, ejemplo: **4 km al SW de Macaján**. En aras de la simplificación y evitar una serie de pasos desde los SIG se propone usar el teorema del coseno, que permite resolver triángulos oblicuángulos donde se conocen 2 lados y el ángulo que conforman. Donde, los lados **a** y **b** están definidos por la distancia que reporta la localidad y el coseno del ángulo (**cos  $\theta$** ) para la incertidumbre en grados según el tipo de orientación.

**Teorema del coseno:**

$$c = \sqrt{a^2 + b^2 - 2ab \cos \theta}$$

En la figura 16A se aplica el Teorema del coseno para un ejemplo en particular. Para calcular la incertidumbre únicamente por orientación, **a** y **b** tienen el mismo valor y conforman un ángulo de 22.5°. Sin embargo, la localidad lleva implícita la incertidumbre por distancia que es el 15% del valor de la unidad, valor que se suma a los 4 km generando un aumento del radio de incertidumbre (véase Fig. 16B). De esta manera se obtiene la incertidumbre integrada por orientación y distancia (véase Fig. 16B).



**Figura 16.** Cálculo de la incertidumbre integrando orientación y distancia mediante el teorema del coseno A) Incertidumbre solo por orientación B) Incertidumbre integrada.

**2.2.5 Imprecisión en la medida de las coordenadas**

Este tipo de incertidumbre está asociada al número de decimales reportados, es decir, entre más dígitos sean registrados mayor es la precisión de las coordenadas y menor la

incertidumbre, y viceversa. Este cálculo se debe hacer a partir de las coordenadas originales, ya que la conversión puede incrementar la incertidumbre.

La Tabla 2 muestra la incertidumbre para diferentes niveles de precisión de las coordenadas para la latitud ecuatorial usando el sistema elipsoidal WGS-84. La incertidumbre para la línea ecuatorial es la equivalencia más cercana para Colombia, país ubicado entre 4° 13' 30" latitud Sur y 12° 27' 46" latitud Norte (IGAC 1996).

**Tabla 2.** Incertidumbre por precisión de las coordenadas usando el sistema WGS-84 para 0° latitud (Basado en Wieczorek et al. 2004).

Formato de coordenadas	Latitudes/Grados	Ejemplos	Incertidumbres
<b>Grados, minutos y segundos</b>	1 grado	4°	156.904 m
	1,0 minutos	4° 2'	2.615 m
	1,0 segundo	4° 2' 1"	44 m
	0,1 segundo	4° 2' 1,1"	5 m
	0,01 segundos	4° 2' 1,11"	1 m
<b>Grados, minutos decimales</b>	0,1 minutos	4° 3,1'	262 m
	0,01 minutos	4° 3,11'	27 m
	0,001 minutos	4° 3,111'	3 m
<b>Grados decimales</b>	0,1	2,1	15.691 m
	0,01	2,11	1.570 m
	0,001	2,111	157 m
	0,0001	2,1111	16 m
	0,00001	2,11111	2m

El proyecto de georreferenciación MaNIS/HerpNet/ORNIS tiene a disposición la calculadora MaNIS<sup>3</sup>, diseñada para evaluar todos los tipos de incertidumbre o uno en específico. Sin embargo, la ejecución puede ser dispendiosa para procesos masivos de georreferenciación. Para el uso de esta calculadora se requiere del control de las opciones de ejecución, para lo cual se ha dispuesto de un manual específico para este cálculo<sup>4</sup>.

## 2.2.6 Incertidumbre por escala del mapa

La ubicación del punto exacto de la descripción de una localidad se hace sobre un mapa digital. La exactitud del mapa depende de los datos originales utilizados para elaborar el mapa

<sup>3</sup> MaNIS/HerpNet/ORNIS Georeferencing Guidelines. <http://manisnet.org/GeorefGuide.html>

<sup>4</sup> Uso de la calculadora MaNIS para calcular la incertidumbre por coordenadas. [http://www.sibcolombia.net/c/document\\_library/get\\_file?p\\_l\\_id=33134&folderId=53536&name=DLFE-2105.pdf](http://www.sibcolombia.net/c/document_library/get_file?p_l_id=33134&folderId=53536&name=DLFE-2105.pdf)

y la resolución en que ha sido impreso o visualizado (Wieczorek 2001), por lo tanto el cálculo de la incertidumbre por escala del mapa resulta de la multiplicación del error gráfico máximo permisible por la escala de mapa.

$$\text{Incertidumbre} = \text{error gráfico} \times \text{escala del mapa}$$

Para la cartografía con el estándar *United States Geological Survey*, para escalas mayores a 1:20.000, el error gráfico en milímetros es de 0,846 mm (1/30 pulgadas), y para escalas de 1:20.000 o menores el error en milímetros es de 0.508mm (1/50 pulgadas) (USGS 1999) (véase Tabla 3).

**Tabla 3.** Ejemplos del cálculo de incertidumbre por escala del mapa según estándar *United States Geological Survey*.

Escala	Error gráfico	Incertidumbre
1:5.000	0,846 mm	4,2 m
1:10.000	0,846 mm	4,8 m
1:20.000	0,508 mm	10,16 m
1:100.000	0.508 mm	50,8 m
1:500.000	0.508 mm	254 m

Para la cartografía de Colombia, el estándar del Instituto Geográfico Agustín Codazzi (IGAC) plantea que el 90% de los puntos extraídos del mapa deben concordar dentro de 0,5 mm a la escala del mismo respecto a la posición del punto en el terreno. El error gráfico según el parámetro "error medio cuadrático" corresponde a 0,30 mm (IGAC 1994), valor que aplica para toda la cartografía (véase Tabla 4).

**Tabla 4.** Cálculo de incertidumbre por escala del mapa según error gráfico del IGAC

Escala	Error gráfico	Incertidumbre
1:5.000	0,30 mm	1,5 m
1:10.000	0,30 mm	3 m
1:20.000	0,30 mm	6 m
1:100.000	0,30 mm	30 m

De no haber la certeza o desconocer el estándar que tiene la cartografía, se establece un valor de 1 mm de error (Wieczorek 2001), el cual se multiplica con la escala del mapa.

## **CAPÍTULO 3.**

# **PROTOCOLO DE GEORREFERENCIACIÓN DE REGISTROS BIOLÓGICOS**

### 3.1 Proceso general de la georreferenciación de localidades

La propuesta para la georreferenciación de registros biológicos se basa en el Método Radiopunto con algunas adaptaciones en los procedimientos y lleva una secuencia de 5 etapas, que incluyen el reconocimiento de la información, la estandarización de la localidad, clasificación de la descripción en niveles según la calidad de la información, georreferenciación y validación de los datos generados. La conectividad de estas etapas junto con los procesos intermedios se representa en el siguiente diagrama de flujo.

La ejecución de estas etapas confiere mayor calidad al conjunto de datos sin incrementar el error. A continuación, se precisa de manera general la importancia de cada etapa en la georreferenciación de registros biológicos.

#### **Etapa 1:** Reconocimiento de la información de localidad

Inicialmente, se debe realizar la identificación de todos los datos que hacen parte de la localidad y los elementos o campos que deben ser adicionados para consignar la información generada. Este proceso permite mantener un orden en la documentación de los datos y a su vez obtener una mirada general del estado de la base de datos.

#### **Etapa 2:** Estandarización de localidades

En la base de datos es común encontrar varias descripciones con el mismo contenido pero con diferente sintaxis, gramática, escritura u orden. Se han establecido una serie de reglas generales y específicas para homogenizar las localidades, obteniendo una reducción en el volumen de información; de esta manera se mejora la calidad de los datos y hace de la georreferenciación un proceso más eficiente.

#### **Etapa 3:** Clasificación de la descripción de localidad en niveles de calidad

Las localidades son muy diversas desde localidades con una correcta descripción fácilmente rastreable en cartografía, hasta localidades ambiguas, inconsistentes e incluso registros que no tienen datos de localidad y los únicos referentes son las entidades político administrativas. Con el fin de adaptar los procesos de georreferenciación al estado de las localidades, se proponen 7 niveles que se estructuran según los datos informativos que contiene la localidad.

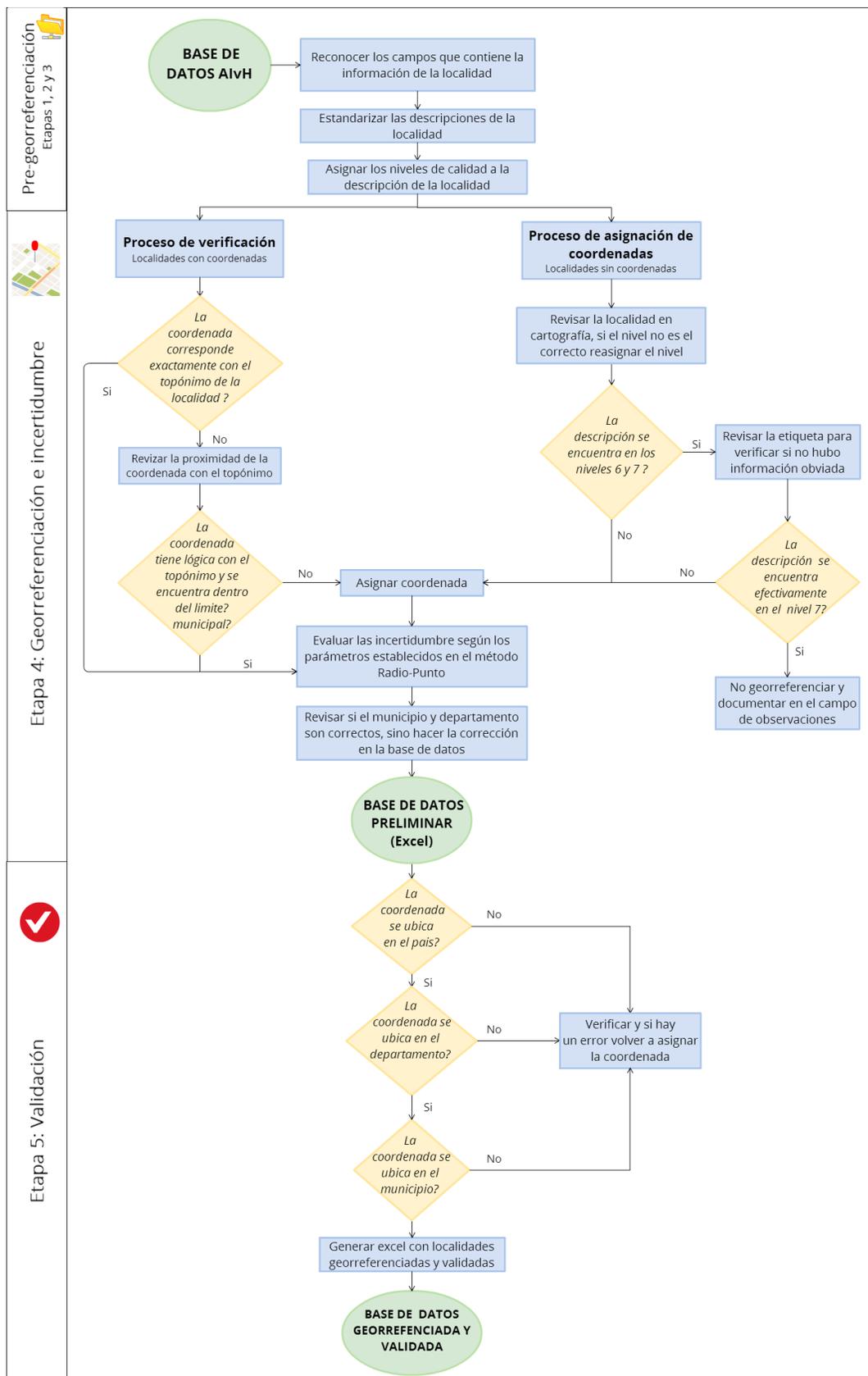


Figura 17. Diagrama de flujo del proceso de georreferenciación de localidades de registros biológicos

#### **Etapa 4:** Georreferenciación e incertidumbre

Esta etapa busca por una parte evaluar la concordancia entre las coordenadas y la descripción para el conjunto de datos que ya cuenta con coordenadas pero que requiere ser verificado. Existe otro conjunto de datos que no tiene coordenadas donde el proceso de georreferenciación busca asignar coordenadas al sitio probable de colecta según la descripción.

La evaluación de la incertidumbre se hace para todos los datos y permite dar un valor cuantitativo de confiabilidad al dato georreferenciado facilitando al investigador discriminar los datos según su estado de incertidumbre. Este valor puede integrar varios parámetros como ausencia del datum, estado de las coordenadas, escala del mapa, unidad de medida, orientación y extensión de la entidad geográfica de referencia.

#### **Etapa 5:** Validación

En el proceso de georreferenciación pueden ocurrir errores en la asignación o digitación de la información, por esta razón se realiza la validación. Para asegurar la calidad del producto se evalúa la correspondencia de las coordenadas asignadas respecto a la geografía superior, haciendo referencia al país, departamento y municipio.

Esta propuesta se ha adaptado al estado de las localidades en las colecciones biológicas del país, con una posición más flexible en su ejecución, con el fin de rescatar la información geográfica asociada a un registro biológico. Un caso de ello es la georreferenciación de un registro cuando no tiene datos de localidad y solo contiene la división político administrativa o entidades geográficas de gran extensión, aunque contiene una alta incertidumbre se justifica la georreferenciación mientras se reporte la incertidumbre, y cuyo valor le permite al investigador hacer uso o no del dato de georreferenciado según el detalle de análisis que requiera.

## 3.2 Reconocimiento de la información

### 3.2.1 Revisión de elementos geográficos

El reconocimiento de los elementos se hace directamente desde la base de datos según la estructura del modelo de datos que maneje la colección. En muchos casos se manejan archivos *Excel* o aplicaciones diseñadas para el almacenamiento y digitalización de datos primarios sobre biodiversidad. El resultado de la consulta debe ser exportado a *Excel* con el fin de generar un archivo de trabajo y no afectar el contenido de la base de datos.

La Tabla 5, contiene los campos donde se puede encontrar información asociada a la localidad y de utilidad para la georreferenciación. Existen otros elementos que no hacen parte de la información geográfica de la localidad pero deben ser revisados ya que pueden ser útiles como: hábitat, comentarios, colección y fecha del evento.

La colección de referencia puede ser relevante cuando se trata de localidades en ecosistemas acuáticos (laguna, río, mar, etc.), ya que la interpretación de la localidad tendría algunas variaciones para definir el sitio más de colecta. En el caso de la fecha del evento facilita identificar si existieron cambios administrativos y evaluar la equivalencia actual.

### 3.2.2 Preparación del archivo de trabajo

La información generada en la estandarización y georreferenciación debe documentarse en elementos diferentes a los originales, para esto se debe adicionar nuevos elementos a la base de datos con el fin de mantener la integridad de la información original. Se recomienda usar el estándar *Darwin Core* como guía para documentar la información producto de la georreferenciación (véase Tabla 6). De esta forma, los datos pueden ser compartidos a través del Sistema de Información de Biodiversidad (SiB Colombia) y globales como la Infraestructura Mundial para Datos sobre Biodiversidad (GBIF).

**Tabla 5.** Elementos para la georreferenciación y definición según los lineamientos del estándar Darwin Core.

Elemento	Ejemplo	Descripción
localityID	3105	Identificador único por cada localidad.
País	Colombia	Información político administrativa que no pertenece a datos específicos del sitio de colecta.
Departamento	Amazonas	
Municipio	Leticia	
Localidad original ( <i>localityName</i> )	Estación Biológica El Zafire	Descripción textual original del lugar.
Latitud Original ( <i>VerbatimLatitude</i> )	04° 00' S	La latitud y longitud textual de la ubicación sin modificaciones.
Longitud Original ( <i>VerbatimLongitude</i> )	69° 54' W	
Latitud Decimal ( <i>decimalLatitude</i> )	-4.0583333	Coordenadas en grados decimales de la ubicación, pueden ser generadas por transformaciones o georreferenciación.
Longitud decimal ( <i>decimalLongitude</i> )	-69.9125	
Elevación máxima en metros ( <i>maximum ElevationInMeters</i> )	100	El límite inferior del rango de elevación en metros (altitud, por encima del nivel del mar).
Elevación máxima en metros ( <i>maximum ElevationInMeters</i> )	200	El límite superior del rango de elevación en metros (altitud, generalmente por encima del nivel del mar).
Hábitat ( <i>hábitat</i> )	Bosque	Descripción del hábitat en el que ocurrió el evento. La información contenida no pertenece a la localidad, pero en ocasiones puede contener información sobre el sitio de colecta.
Código de la colección ( <i>CollectionCode</i> )	Entomología	El nombre, acrónimo, código alfanumérico, o iniciales que identifican la colección o conjunto de datos del que procede el registro.
NationalParkName	Reserva Forestal del Rio Calderón	Área protegida a la que pertenece.
Comentarios de georreferenciación ( <i>georeferenceRemarks</i> )	Coordenadas tomadas por mapa	Anotaciones o comentarios sobre métodos de captura o procedimientos de georreferenciación.
Fecha del evento ( <i>eventDate</i> )	1998-03-14	La fecha o el intervalo cuando se produjo el evento. En la georreferenciación es fundamental para identificar localidades antiguas, las cuales pueden tener una variación de las entidades geográficas o división político administrativa con respecto a la actual.

**Tabla 6.** Elementos para documentar los productos de la estandarización y georreferenciación según el estándar Darwin Core (TDWG 2011).

Elementos	Definición
Localidad ( <i>locality</i> )	Localidad producto de la estandarización.
Elevación mínima en metros ( <i>minimumElevationInMeters</i> )	Dato de elevación mínima producto de la georreferenciación.
Elevación máxima en metros ( <i>maximumElevationInMeters</i> )	Dato de altitud máxima producto de la georreferenciación.
Latitud decimal ( <i>decimalLatitude</i> )	Coordenadas en grados decimales de la ubicación en latitud producto de la georreferenciación.
Longitud decimal ( <i>decimalLongitude</i> )	Coordenadas en grados decimales de la ubicación en longitud producto de la georreferenciación.
Datum geodésico ( <i>geodeticDatum</i> )	El elipsoide, datum geodésico, o sistema de referencia espacial (SRS) usado en la georreferenciación.
Incertidumbre en metros ( <i>coordinateUncertaintyInMeters</i> )	El uso de este campo tiene una adaptación ya que no corresponde netamente a la incertidumbre por coordenadas sino que incluye la suma de los parámetros de incertidumbre según los datos informativos de la localidad. Tenga en cuenta que cero no es un valor válido para este elemento al menos que la localidad no haya podido ser georreferenciada.
Georreferenciado por ( <i>georeferencedBy</i> )	Nombre de la persona que realizó la georreferenciación.
Fecha de georreferenciación ( <i>georeferenceDate</i> )	Fecha expresada en día, mes y año de la georreferenciación
Protocolo de georreferenciación ( <i>georeferenceProtocol</i> )	Descripción del método de georreferenciación usado y/o nombre del protocolo que contiene los métodos.
Fuentes de georreferenciación ( <i>georeferenceSources</i> )	Una lista de los mapas, gaceteros, u otros recursos utilizados para georreferenciar la ubicación, lo suficientemente específica como para permitir que cualquier persona en el futuro utilice los mismos recursos.
Comentarios de la georreferenciación ( <i>georeferenceRemarks</i> )	Notas o descripciones con información que no es producto exacto de la georreferenciación pero que puede contribuir a su interpretación. Por ejemplo, la documentación del por qué una localidad no ha sido georreferenciada, uso de métodos particulares, aclaraciones sobre la localidad, etc.

### 3.3 Estandarización de localidades

El proceso de estandarización de localidades se entiende como la aplicación de una serie de parámetros que definen la manera correcta de describir una localidad. Su desarrollo mejora la calidad de las descripciones y reduce el volumen de los datos facilitando su tratamiento en el proceso de georreferenciación.

La estandarización de las localidades se puede dividir en dos fases:

### 3.3.1 Organización de la información

Esta fase hace referencia a la manera correcta de organizar los datos geográficos que contiene la localidad llevando un orden jerárquico que va desde lo general a más específico. Esto requiere identificar cuál es el tipo de información relevante, es decir, aquella que permite ubicar el sitio exacto de la colecta. Por ejemplo: accidentes geográficos (montes, valles, acantilados, etc.), entidades administrativas inferiores a municipio (corregimiento municipal y/o departamental, veredas, caseríos, resguardos, etc.), rutas (vías, carreteras, caminos, vías ferroviarias, etc.), afluentes hídricos (ríos, quebradas, etc.), áreas protegidas, entre otros; los cuales pueden estar asociados a distancias y orientaciones.

La información sobre hábitat, suelos, ecosistemas y zonas de vida no deben ser documentados en el campo de localidad ya que son rasgos que no son verificables en el tiempo, y su utilidad puede ser relativa en la georreferenciación.

### 3.3.2 Aplicación de reglas y normalización de las localidades

Un problema frecuente en las bases de datos es la duplicidad de localidades generada por errores gramaticales, mal uso de mayúsculas, puntuación, idioma y sintaxis. Al normalizar o estandarizar las localidades, se reduce significativamente el número y facilita su manejo.

Se han establecido reglas generales que incluyen los criterios básicos que se deben aplicar a todas las descripciones y reglas específicas que se enfocan en casos puntuales (véase Tabla 7 y 8).

**Tabla 7.** Reglas generales para estandarizar localidades de registros biológicos.

Regla	Ejemplo
<p><b>Verificación del nombre oficial de entidades geográficas</b></p> <ul style="list-style-type: none"> <li>✓ Primera letra en mayúscula e incluye revisión de fuentes oficiales<sup>5</sup>.</li> </ul>	<p>Original/ Depurada</p> <ul style="list-style-type: none"> <li>✓ cundinamarca / Cundinamarca</li> <li>✓ Juan de acosta / Juan de Acosta</li> </ul>
<p><b>Jerarquía en la descripción de la localidad</b></p> <ul style="list-style-type: none"> <li>✓ Estructura que va desde lo general a lo específico.</li> <li>✓ NO modificar el sentido de la descripción, ni agregar ni quitar información.</li> <li>✓ La jerarquía se aplica si la descripción contiene más de una entidad geográfica.</li> </ul>	<ul style="list-style-type: none"> <li>✓ Localidad: 2 km al O del Manantial en el Cerrejón, entre la carretera del Manantial y Albania</li> <li>✓ Localidad depurada: El Cerrejón, entre la carretera del Manantial y Albania, 2 km al W del Manantial</li> </ul>
<p><b>Ortografía</b></p> <ul style="list-style-type: none"> <li>✓ Uso de comas (,) para separar referentes espaciales. Se omiten otros signos de puntuación como: punto y coma (;) dos puntos (:), puntos seguidos (...) y punto final (.), excepto para en el uso de siglas que contengan punto (.).</li> <li>✓ Letra mayúscula para iniciar la frase.</li> <li>✓ No se usa punto final al final de la frase.</li> <li>✓ Se mantiene las tildes (´) si su uso es correcto.</li> <li>✓ Los paréntesis () se usan para hacer aclaraciones u observaciones que complementan la descripción.</li> <li>✓ Las comillas se usan para referirse a citas textuales o descripciones ambiguas, no se deben usar para designar nombres propios. Para esto es suficiente con usar mayúscula inicial.</li> </ul>	<ul style="list-style-type: none"> <li>✓ Localidad: Entre la vereda Juntas y vereda Pico de oro, río combeima. Localidad depurada: Río Combeima, entre la vereda Juntas y vereda Pico de Oro</li> <li>✓ Carretera de Pasto a Sibundoy, en el km (no se ve en etiqueta)</li> </ul>
<p><b>Datos que NO hacen parte de la localidad</b></p> <ul style="list-style-type: none"> <li>✓ Si la descripción contiene información sobre hábitat, suelos, áreas protegidas se debe hacer el traslado al campo correspondiente.</li> </ul> <p><b>Nota:</b> si el área protegida es una dato fundamental para entender la descripción se mantiene, por ejemplo: A 5 km al NE del PNN Pisba.</p>	<ul style="list-style-type: none"> <li>✓ Localidad: Isla la corota, Santuario de flora y fauna la Corota.</li> <li>✓ Localidad depurada: Isla la Corota.</li> </ul> <p><b>Nota:</b> el área protegida se traslada al campo correspondiente.</p>
<p><b>Descripciones sin datos de localidad:</b> Descripciones que no poseen referentes espaciales específicos del sitio de colecta.</p> <ul style="list-style-type: none"> <li>✓ Si hay el reporte de municipio como único referente, se debe escribir en el campo de localidad "Sin datos de localidad" y trasladar el municipio a su respectivo campo.</li> </ul>	<ul style="list-style-type: none"> <li>✓ Localidad: Tunja</li> <li>✓ Localidad depurada: Sin datos</li> <li>✓ Municipio: Tunja</li> </ul> <p><b>Nota:</b> Tunja se trasladó al campo del municipio</p>

<sup>5</sup> **NOTA:** IGAC y DANE son las fuentes oficiales para Colombia, su uso depende de los criterios e insumos usados para el diseño de la estructura de datos geográficos en la base de datos en las colecciones biológicas.

**Tabla 8.** Reglas específicas para estandarizar localidades.

<b>Regla</b>	<b>Ejemplo</b>
<p><b>Siglas</b></p> <ul style="list-style-type: none"> <li>✓ Se admite el uso de siglas para nombres de entidades institucionales (Corporaciones, institutos, organismos de gestión pública) y descripciones de sociedades mercantiles (sociedad anónima (S.A), sociedad limitada (Ltda.), sociedad por acciones simplificada (S.A.S), entre otros.</li> <li>✓ Se respeta el uso de puntos siempre que sea una fuente oficial.</li> </ul> <p><b>Siglas de áreas protegidas</b></p> <ul style="list-style-type: none"> <li>✓ No se usan siglas para los nombres de las áreas protegidas, ni para sus categorías.</li> </ul>	<ul style="list-style-type: none"> <li>✓ S.A, S.A.S y Ltda.</li> <li>✓ INDERENA</li> </ul> <ul style="list-style-type: none"> <li>✓ Localidad: RNN Nukak</li> <li>✓ Localidad depurada: Reserva Natural Única Nukak</li> </ul>
<p><b>Abreviaturas</b></p> <ul style="list-style-type: none"> <li>✓ Se recomienda el uso de abreviaturas para las unidades de distancias rigiéndose por el sistema internacional de unidades y orientaciones cardinales por sus siglas en inglés.</li> <li>✓ No se admiten abreviaturas para nombres de entidades geográficas de división político administrativa y accidentes geográficos.</li> </ul>	<ul style="list-style-type: none"> <li>✓ Incorrecto: 10kms. 50 mts. / Correcto: 10 km (kilómetros), 50 m (metros), 30 s. Sin usar puntos y separados por un espacio del número.</li> <li>✓ Norte (N), Sur (S), Este (E), Oeste (W), Noreste (NE), Sureste (SE), Noroeste (NW), Suroeste (SW), Nor-noreste (NNE), Nor-noroeste (NNW), Este-noreste (ENE), Este-sureste (ESE), Oeste-noroeste (WNW), Oeste-suroeste (WSW), Sur-suroeste (SSW) y Sur-sureste (SSE).</li> <li>✓ Incorrecto: Vda. Congo./ Correcto: Vereda Congo</li> <li>✓ Incorrecto: Hda. Cotové / Correcto. Hacienda Cotové</li> <li>✓ Incorrecto: Qda. La Agudelo / Correcto: Quebrada La Agudelo</li> <li>✓ Incorrecto: Ca. / Correcto: cerca, cerca de</li> <li>✓ Incorrecto: prox / Correcto: próximo</li> <li>✓ Incorrecto: alr. / Correcto: alrededores</li> </ul>
<p><b>Idioma</b></p> <ul style="list-style-type: none"> <li>✓ Las descripciones en otro idioma deben ser traducidas al español.</li> <li>✓ Se mantiene los nombres propios del idioma de origen.</li> <li>✓ La traducción se hará lo más cercana a la descripción original, manteniendo las reglas generales y específicas de la estandarización.</li> </ul>	<p>Inglés: About 4 km E. of Roche on the Roche - Tabaco road</p> <p>Español: Entre la carretera Roche y Tabaco, cerca de 4 Km al E de Roche</p>
<p><b>Uso de conectores</b></p> <p>Se utilizan cuando la expresión de la localidad no es suficientemente clara para su interpretación. Se recomiendan los siguientes conectores:</p> <p><b>Entre:</b> punto medio entre dos entidades geográficas y se usa anteponiéndolo a la primera entidad geográfica.</p>	<ul style="list-style-type: none"> <li>✓ Localidad: Silvania-Fusagasugá</li> <li>Localidad depurada: Entre Silvania y Fusagasugá</li> </ul>

Regla	Ejemplo
<b>Y:</b> Se ubica entre las dos entidades geográficas y permite hacer más explícita la delimitación del sitio de colecta.	

Hasta el momento se han abordado dos etapas claves en la georreferenciación de registros biológicos, (i) el reconocimiento de los elementos y preparación de archivo, y (ii) el proceso de estandarización de la descripción de localidad. En la figura 18 se encuentra la secuencia de estos dos procesos y fases intermedias.

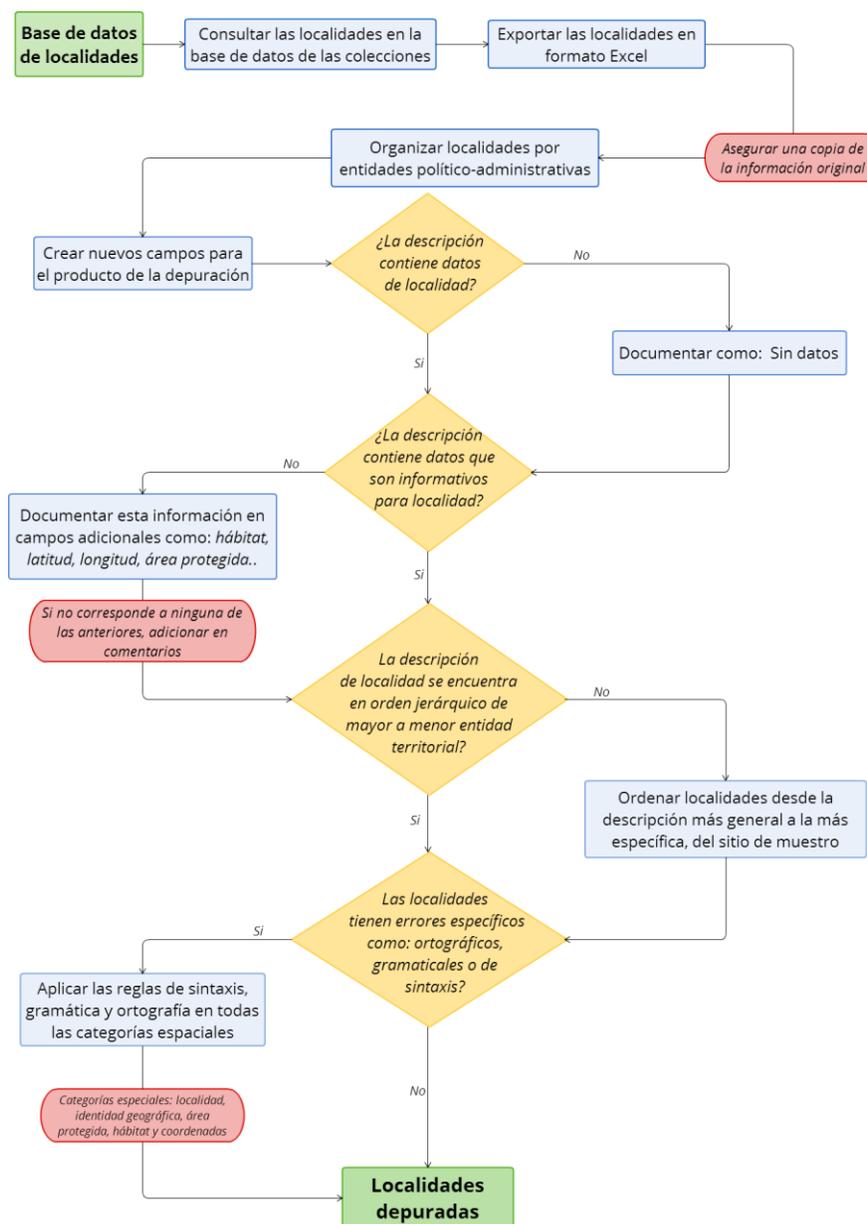


Figura 18. Diagrama de flujo del proceso de preparación de archivo de trabajo y estandarización de localidades.

### 3.4 Clasificación de la localidad en niveles de calidad

El contenido de las descripciones de la localidad puede variar según la calidad de la descripción, incidiendo en los procedimientos y criterios específicos en la georreferenciación.

Desde un enfoque teórico, metodológico y didáctico, se ha elaborado una clasificación basada en los datos informativos que puede tener la localidad. La clasificación de localidades lleva una estructura numérica del 1 al 7, los cuales se diferencian de acuerdo al tipo de información que contiene la localidad. El nivel con mayor calidad es el 1, el cual contiene localidades con coordenadas permitiendo identificar el punto exacto de la colecta. Los niveles subsiguientes no tienen coordenadas, a medida que aumenta el nivel, las localidades son menos precisas y de menor calidad hasta el nivel 7 que corresponde a información dudosa o ambigua que no puede ser georreferenciada.

Estos niveles se han denominado *Niveles de calidad de la descripción de la localidad*, los cuales son el eje de los procesos subsecuentes para la asignación y verificación de coordenadas, y evaluación de la incertidumbre (véase Tabla 9).

**Tabla 9.** Niveles de calidad de las descripciones en localidades de registros biológicos.

Nivel	Descripción	Ejemplos
1	Localidades con coordenadas que fueron tomadas en campo con GPS o que fueron inferidas a partir de información secundaria como gaceteros, bases de datos de topónimos, bitácoras, mapas digitales o análogos, etc <sup>6</sup> .  En este nivel se realiza un proceso de verificación de coordenadas para evaluar si las coordenadas coinciden con la división político administrativa y el topónimo preciso del sitio de colecta descrito en la localidad.	<ul style="list-style-type: none"> <li>Vereda las flores, cueva el Arrayán (Zapatoca, Santander)</li> <li>Coordenadas 6° 48' 3.5" N, 73° 16' 11.68" W</li> </ul>
2	Localidades sin coordenadas con descripciones que contienen entidades geográficas asociadas a distancias y/u orientaciones, sea por una ruta o por aire (distancia lineal), condición que facilita identificar el sitio puntual de colecta.  Los atributos geográficos que se pueden encontrar son rutas (vías principales y secundarias, caminos, etc.), formaciones hidrográficas (ríos, quebradas, humedales, etc.), accidentes geográficos (cerros, cuevas, valles, islas, etc.), y/o entidades administrativas inferiores o iguales a municipio (corregimiento, caseríos, resguardos, veredas, etc.).	<ul style="list-style-type: none"> <li>2km al ENE de Uribia</li> <li>Desde Zaragoza a 22 Km al S y 25 km al O</li> <li>Sobre la vía entre Tabaco y Roche, 1,5 km al SE de Roche</li> <li>Al N del PNN Pisba</li> </ul>

<sup>6</sup> **NOTA:** No se hace una discriminación según el método de captura, debido a que en la mayoría de bases de datos de las colecciones del país no está documentado. Por lo tanto no es posible tener la certeza que un dato sea de GPS o de otra fuente.

Nivel	Descripción	Ejemplos
3	<p>Descripciones de localidades que tienen las mismas entidades del Nivel 2, sin embargo, no cuentan con datos de distancia ni orientación. Frecuentemente, están relacionados a un punto entre entidades geográficas, rutas cuyos conductores pueden ser ríos y vías, intersección entre dos rutas o ríos, o simplemente hacen mención a entidades político administrativos inferiores a municipio (veredas, corregimientos, cabeceras municipales), accidentes geográficos locales (cerros, caño, colinas, isla, etc.) y formaciones hidrográficas (ríos, quebradas, arroyo, lago, etc.).</p> <p>Esta categoría también incluye todas las descripciones que aparecen como "sin datos de localidad" y tienen dato de municipio, pero no aplica para los "sin datos de localidad" con reporte del departamento o país.</p>	<ul style="list-style-type: none"> <li>• Vía entre Pasto y Túquerres.</li> <li>• Entre Silvania y Fusagasugá</li> <li>• Finca La Paz, margen derecha río Guaviare</li> <li>• Vereda el Pandango</li> <li>• Municipio - Sin datos de localidad</li> </ul>
4	<p>Localidades sin coordenadas con descripciones que no especifican el sitio exacto de colecta sino un área que puede abarcar varias entidades administrativas como municipios o departamentos. Estas localidades puede hacer referencia a áreas de jurisdicción como: parques naturales, resguardos, reservas, etc., o áreas que corresponden a una delimitación natural como páramos, humedales, sabanas, entre otros; además de documentar el dato de municipio y/o departamento.</p> <p>También se han incluido a este nivel los registros que no tienen datos de localidad pero reportan el departamento, siendo el departamento el referente espacial usado para georreferenciar</p>	<ul style="list-style-type: none"> <li>• (Huila, San Agustín) Río Magdalena, 1700 m</li> <li>• (Popayán) Región de Sate</li> <li>• (Leticia) PNN Amacayacu</li> <li>• (Santander) Meseta de los caballeros Departamento - Sin datos de localidad</li> </ul>
5	<p>Localidades con descripciones muy generales que no hacen referencia al sitio de colecta sino a entidades geográficas a escala nacional. Contiene áreas de amplia extensión como: regiones naturales (Andes, Llanos, Caribe, Amazonia), ríos (Río Magdalena, Río Cauca, etc.) y vías principales (vía panamericana).</p> <p>Este grupo de localidades no documentan el departamento y municipio, solo tienen el dato de país.</p>	<ul style="list-style-type: none"> <li>• (Colombia) Orilla de río Magdalena</li> <li>• (Colombia) Llanos Orientales</li> <li>• (Colombia) Río Apaporis</li> </ul>
6	<p>Descripciones con información insuficiente y la única certeza es su ubicación dentro del territorio nacional. Estos datos tienen una alta incertidumbre, sin embargo, se considera pertinente georreferenciarlas teniendo en cuenta su utilidad para análisis espaciales a nivel continental o global.</p> <p>En este nivel aplican todos los registros sin datos de localidad y país como único referente espacial.</p>	<ul style="list-style-type: none"> <li>• (Colombia) Borde de carretera de rocas</li> <li>• (Colombia) En todo el país</li> <li>• (Colombia) Relleno sanitario</li> <li>• País- sin datos de localidad</li> </ul>
7	<p>Descripción de localidades con información dudosa e inconsistente que no permite tener ninguna certeza de la ubicación. En este nivel se pueden ubicar aquellas localidades con homonimias sin resolver, signos de interrogación o incoherencia en la descripción.</p> <p>Las localidades ubicadas en este nivel no se georreferencian.</p>	<ul style="list-style-type: none"> <li>• Tal vez en la Amazonía</li> <li>• Vaupés, Guaviare</li> <li>• Meta Villavicencio??</li> <li>• Posiblemente del Valle</li> <li>• Sin dato de país - Pontevedra</li> </ul>

### 3.4.1 Asignación de niveles de calidad de la localidad

El proceso de asignación de niveles consiste en identificar el tipo de información que contiene la descripción y categorizarla en el nivel que corresponda. Para esto es necesario adicionar una columna en el archivo de trabajo, de tal manera que se pueda georreferenciar por grupos de localidades. Esto no es un proceso obligatorio en la georreferenciación, ya que depende de los intereses del investigador (colección en particular, grupo taxonómico, región, ecosistema, etc.), pero seguramente aquellas localidades con mejor calidad en la descripción le permitirá rastrear con mayor precisión el sitio de colecta y dejar para una fase posterior aquellas localidades conflictivas que pueden requerir un mayor tiempo optimizando el proceso.

Esta clasificación se debe hacer sobre un elemento o campo adicional, el cual no hace parte del estándar de publicación de datos del *Darwin Core*, por lo tanto, no será un elemento que se publique, pero se recomienda que el nivel asignado sea mencionado en el campo de comentarios de georreferenciación (*georeferenceRemarks*).

La figura 19 muestra la estructura lógica para asignar los niveles de clasificación de calidad de la localidad.

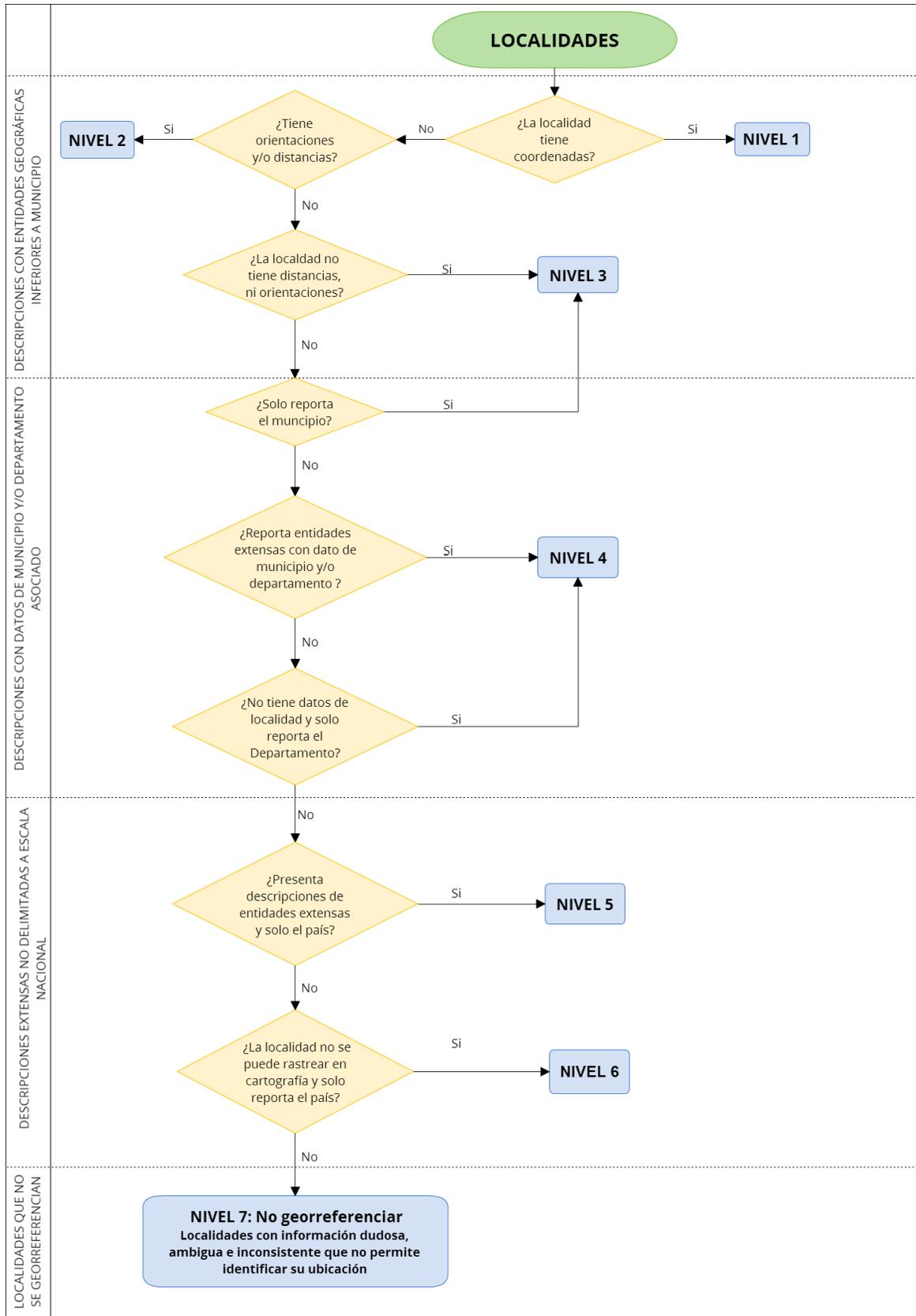


Figura 19. Proceso para la asignación de niveles según la información de la localidad.

## 3.5 Georreferenciación e incertidumbre

Hay dos procedimientos principales en la georreferenciación, el proceso de verificación y asignación de coordenadas. La **Verificación** se aplica para las localidades que tienen coordenadas, y el proceso de **Asignación de coordenadas** para las localidades que solo cuentan con la descripción de la localidad.

Estos dos procesos llevan implícito un margen de error determinado por el estado de los datos y que se traduce en incertidumbre. La incertidumbre se expresa como una medida de longitud en metros y se puede interpretar como el sitio más probable de colecta.

El cálculo de la incertidumbre integra 6 parámetros explicados al inicio de este documento. La aplicación de estos tipos de incertidumbre depende de los datos informativos que contenga la descripción (orientaciones, distancias, rutas, etc.), es decir, no siempre se calculan los mismos parámetros para todas las localidades.

- ✓ Extensión de la localidad de referencia
- ✓ Desconocimiento del Datum
- ✓ Imprecisión en la medida de la distancia
- ✓ Incertidumbre por orientación
- ✓ Incertidumbre por coordenadas
- ✓ Escala del mapa

### 3.5.1 Consideraciones técnicas de la cartografía digital e insumos cartográficos

Para el uso de cartografía digital se requiere de un SIG, que consiste en la integración organizada de hardware, software y datos geográficos. Las herramientas SIG están diseñadas para capturar, almacenar y analizar la información geográfica.

El rastreo de localidades en cartografía requiere de diferentes fuentes de información. Por ejemplo: diccionarios geográficos, visores web de múltiples cartografías, imágenes de satélite, modelo digital de elevación, entre otros.

Estos insumos se pueden dividir en dos tipos de información: vectorial y raster. Los datos vectoriales se pueden presentar como puntos, líneas y polígonos, los cuales representan

entidades geográficas y topónimos. Los datos raster, se usan para la representación de datos continuos, como el Modelo Digital de Elevación (DEM), ampliamente usado para verificar datos de altitud.

A continuación, se hace una revisión de los insumos cartográficos que pueden ser usados en la cartografía digital. Además de algunas precisiones sobre el uso del datum y consideraciones importantes para dar confiabilidad a los procedimientos de la georreferenciación.

**Cartografía digital:** La cartografía debe contener información de la división político administrativa (país, departamentos y municipios), infraestructura (vía férrea, red de vías primarias, secundarias y terciarias), hidrografía (ríos, quebradas, lagunas y ciénagas) y nombres de lugares para las entidades inferiores a municipio o accidentes geográficos (corregimientos, veredas, fincas, cerros, entre otros).

La escala de la cartografía a usar depende del nivel de detalle que se realiza el proceso. En escalas detalladas (1: 25.000) se puede visualizar mayor cantidad y definición de los referentes espaciales mientras que a escala 1:100.000 se abarca un nivel general de los referentes más representativos a nivel nacional. No se recomienda realizar georreferenciación a escalas mayores a 1:100.000 ya que la representación de las entidades geográficas es muy general confiriendo un mayor grado de incertidumbre al proceso.

**Gaceteros:** Hacen referencia a base de datos con nombres geográficos, la mayoría proporcionan el punto de ubicación de los sitios suministrando el dato de la coordenada y su respectivo datum. Gaceteros internacionales como *Geonet Names Server* (GNS), cuentan con una amplia base de datos de nombres geográficos para diferentes países y coordenadas en datum WGS84 permitiendo descargar toda la información de un país.

El uso de este tipo de herramientas es de gran utilidad porque permiten realizar búsquedas avanzadas y precisas con los nombres de entidades geográficas descritas en la localidad, visualizando a través de puntos la localización de los sitios.

**Modelo digital de elevación (DEM):** Es una representación visual de altitud de la superficie terrestre respecto al nivel medio del mar. Estos valores se presentan en archivos tipo raster, útiles para identificar y verificar los datos de elevación reportados en la localidad.

Actualmente la NASA tiene a disposición un modelo de elevación mundial de uso libre través del portal Reverb-ECHO<sup>7</sup>, donde muestra datos de altitud cada 30 metros, siendo posible extraer datos de determinadas regiones o países enteros. Se recomienda el uso de este modelo a través de su descarga para Colombia<sup>8</sup>.

**Visores geográficos:** Actualmente existen plataformas web con visores de múltiple cartografía, imágenes satelitales y repositorios de mapas que permiten hacer búsquedas de lugares a nivel mundial. Entre los más destacados y funcionales para el proceso de georreferenciación se encuentran Google Earth y Google Maps.

**Uso de datum en la georreferenciación:** Es indispensable antes de iniciar el proceso de georreferenciación, definir parámetros estándar respecto al uso del datum, con el fin de generar datos replicables y reducir al máximo la transformación de los datos por parte del usuario.

Se propone el uso del datum WGS 84 porque es el más usado a nivel global y captura de coordenadas con GPS. Si se manejan diferentes tipos y/o cartografía con datum diferentes, las coordenadas se pueden ver desplazadas y cometer errores en su ubicación.

En el caso de la verificación de las coordenadas, se requiere importar los datos de las coordenadas a una aplicación SIG. Si las coordenadas cuentan con un datum diferente al WGS 84, la verificación se hace sobre el datum original, y si no reportan el datum, se define el uso del datum WGS 84. La cartografía usada para realizar la verificación, debe manejar el mismo datum para que la ubicación de las coordenadas, de lo contrario se pueden desplazar con respecto a su posición original.

**Nota:** Cuando se trata de localidades antiguas, los insumos cartográficos son muy limitados y seguramente los insumos anteriormente documentados no sean los más idóneos para la revisión cartográfica de la localidad. Para un mejor acercamiento se requiere de la revisión de fuentes bibliográficas, libretas de campo, bitácoras, mapas históricos, entre otros.

<sup>7</sup> [http://reverb.echo.nasa.gov/reverb/#utf8=%E2%9C%93&spatial\\_map=satellite&spatial\\_type=rectangle](http://reverb.echo.nasa.gov/reverb/#utf8=%E2%9C%93&spatial_map=satellite&spatial_type=rectangle)

<sup>8</sup> Manual de descarga del Modelo Digital de Elevación – DEM

### 3.5.2 Criterios generales para la verificación de localidades

Unas coordenadas defectuosas se pueden detectar por medio de la verificación entre las descripciones de la localidad respecto a las coordenadas (Soberón & Peterson 2004). Como se describe anteriormente, aun si la localidad tiene coordenadas y permiten un rastreo más preciso de la localidad, se requieren revisar ya que se pueden cometer errores en la mala manipulación de GPS, falta del datum, digitalización, entre otros.

Para la verificación se requiere seguir los siguientes pasos:

- ✓ **Localización de la coordenada en cartografía:** Consiste en identificar el punto exacto donde se ubica la coordenada reportada en la localidad. Se debe tener en cuenta el tipo de coordenadas documentado, que pueden ser geográficas (latitud, longitud) o planas proyectadas (X, Y). En este proceso se debe evitar no realizar transformaciones, con el fin de identificar el punto exacto de ubicación.

Para las coordenadas es importante verificar la orientación y los signos negativos en los campos de latitud (Norte o Sur) y longitud (Este u Oeste) para evitar una desviación en la localización. Tenga en cuenta que los valores límites de latitud están entre -90 a +90 y de longitud entre -180 a +180, con latitudes positivas en el hemisferio norte y longitudes positivas al oriente del Meridiano de Greenwich.

Para coordenadas UTM se debe tener en cuenta el huso y zona para una correcta localización. La ausencia de estos datos pueden causar indeterminación en la localización geográfica (Fernández-Coppel 2001).

- ✓ **Rastreo de la localidad en cartografía:** Se realiza ubicando los referentes espaciales de la localidad en la cartografía. Su búsqueda se puede hacer con la ayuda de gaceteros, bases de datos, imágenes de satélite o visores web.

Comúnmente las localidades se encuentran ligadas a los datos de país, departamento y municipio, información que tiene que ser revisada y corregida si es el caso. Es frecuente encontrar que la localidad es correcta pero el municipio asignado es erróneo, esto pasa frecuentemente ya que en campo es difícil reconocer los límites municipales.

- ✓ **Verificación de las coordenadas:** Se evalúa la correspondencia espacial de la coordenada con respecto a la ubicación de la localidad en el mapa. En este sentido la descripción de la

localidad juega un papel importante porque esta información permite rastrear la localidad. Sin embargo, muchas localidades no tienen descripción de sitio de colecta, por lo tanto, no hay un referente espacial específico de la localidad que permita realizar la verificación del sitio exacto de colecta. Debido a esto se plantea los casos en los cuales se pueden validar las coordenadas (véase Tabla 10).

- ✓ **Incertidumbre:** Los parámetros de incertidumbre que se tiene en cuenta en la verificación son la incertidumbre por coordenadas, desconocimiento de datum y en algunos casos específicos la incertidumbre por extensión. En la Tabla 10, se especifican con mayor detalle los parámetros asociados a los casos más frecuentes.
- ✓ **Altura:** Este dato permite que el proceso de verificación de las coordenadas sea más minucioso. La altura solo puede ser utilizada cuando el dato de la altura se encuentre en el área probable del sitio de colecta, teniendo en cuenta los siguientes criterios:
  - Si la altura definida por las coordenadas no difieren en más de 100 metros con respecto a la altura original, el dato de altura definido por las coordenadas se conserva.
  - Si el dato de altura dado por las coordenadas difiere en más de 100 metros con respecto a la altura original, la coordenada georreferenciada se desplaza al dato de la altura, siempre y cuando, este dato se ubique en el área probable del sitio de colecta. De lo contrario, reubicar la coordenada a un punto distante puede implicar un mayor error que mantener la coordenada.

**Tabla 10.** Casos frecuentes en el proceso de verificación y parámetros de incertidumbre.

<b>Con datos de localidad</b>	<b>Parámetros de incertidumbre</b>	<b>Estado</b>
<p><b>Caso 1</b></p> <p>Coordenadas que coinciden con el topónimo.</p>	<ul style="list-style-type: none"> <li>✓ Desconocimiento de Datum (si no está reportado)</li> <li>✓ Incertidumbre por coordenadas</li> </ul>	Se validan las coordenadas.
<p><b>Caso 2</b></p> <p>Coordenadas que no coinciden con el topónimo pero son próximas.</p> <p>Por el momento no hay un parámetro de distancia máxima para evaluar la validez de las coordenadas con respecto a la descripción. Sin embargo, se propone que si las coordenadas y el punto que marca la descripción se encuentran dentro del municipio, las coordenadas se mantienen, con la salvedad que de encontrarse las coordenadas en el límite o cerca a este, se considere el contexto de la descripción, y no se descarte la validez de estas.</p>	<ul style="list-style-type: none"> <li>✓ Desconocimiento del Datum</li> <li>✓ Incertidumbre por coordenadas</li> <li>✓ Incertidumbre por extensión: se calcula la distancia entre las coordenadas y el punto que marca la descripción de la localidad.</li> </ul>	Se validan las coordenadas.
<p><b>Caso 3</b></p> <p>Coordenadas que no coinciden en absoluto con el topónimo</p>	<ul style="list-style-type: none"> <li>✓ Se evalúan en la georreferenciación según el estado y descripción de la localidad.</li> </ul>	No se validan las coordenadas y se georreferencia nuevamente.
<b>Sin datos de localidad</b>	<b>Parámetros de incertidumbre</b>	<b>Estado</b>
<p><b>Caso 4</b></p> <p>Descripciones sin datos de localidad, con reporte de la división político administrativa y las coordenadas se ubican dentro del municipio reportado.</p>	<ul style="list-style-type: none"> <li>✓ Incertidumbre por coordenadas.</li> <li>✓ Desconocimiento del Datum</li> <li>✓ Incertidumbre por extensión: Si la altura es coherente en el área de trabajo, la extensión es la distancia que hay desde las coordenadas a la cota o pixel más cercano del modelo de elevación. Si no hay altura o esta no es correcta, la extensión va hacia el límite más lejano del municipio.</li> </ul>	Se validan las coordenadas.
<p><b>Caso 5</b></p> <p>Descripciones sin datos de localidad y con coordenadas que no se ubican dentro del municipio ni departamento reportado.</p>	<ul style="list-style-type: none"> <li>✓ Se calcula según la georreferenciación (incertidumbre por extensión de la entidad de referencia e incertidumbre pro escala del mapa)</li> </ul>	No se validan las coordenadas y se georreferencia a partir de la entidad de referencia más específica.

### 3.5.3 Criterios para la asignación de coordenadas e incertidumbre

El proceso de asignación de coordenadas se realiza a partir de la información de la localidad de aquellas localidades que no tienen coordenadas. Cada procedimiento se encuentra ligado al **nivel de calidad de la descripción** y según el tipo de información se precisan los mínimos procedimientos metodológicos para georreferenciar.

- ✓ **Rastreo cartográfico de la localidad:** Consiste en identificar las entidades geográficas que son informativas para hallar el sitio de la colecta. De no haber datos de localidad, se debe identificar la entidad político administrativa más específica.
  
- ✓ **Asignación de las coordenadas:** Una vez se ha identificado el sitio de colecta, se asigna la coordenada teniendo en cuenta los siguientes parámetros.
  - a. **Coordenadas expresadas en grados decimales:** Este sistema se define por dos atributos, latitud y longitud, sistema más conveniente para la georreferenciación ya que tiene aplicabilidad a escala mundial y puede ser transformado fácilmente en otros sistemas de coordenadas, así como de un dato a otro (Wieczorek et al. 2004).
  - b. **Número de decimales:** Un aumento o disminución de los decimales puede implicar un desplazamiento en su ubicación y precisión. Por tal razón, las coordenadas deben incluir el mayor número de decimales de precisión dado por el origen de las coordenadas (Chapman & Wieczorek 2006), respetando el número de dígitos reportado en la captura.
  
- ✓ **Altura:** El uso de la altura permite afinar el proceso de georreferenciación, siempre y cuando el dato se encuentre en el área probable del sitio de colecta que se define por la descripción, si la altura no se reporta en el sitio es preferible no utilizar este dato y netamente incluir el dato de altura dado por el punto georreferenciado según la localidad.
  
- ✓ **Evaluación de la incertidumbre:** Según la información de la localidad se deben escoger los parámetros de incertidumbre. Las fuentes de incertidumbre potencial para cada nivel se encuentran en la Tabla 11. Observe que las localidades del nivel 2 al 7, no incluyen las incertidumbres por coordenadas y ausencia del datum, justamente porque no contiene coordenadas, característica que agrupa a las localidades del nivel 1.

**Tabla 11.** Parámetros de incertidumbre potenciales según los niveles de calidad

Niveles	Extensión de localidad	Desconocimiento del datum	Incertidumbre por coordenadas	Escala del mapa	Incertidumbre por orientación	Imprecisión en la distancia
1	X	X	X			
2	X			X	X	X
3	X			X		
4	X			X		
5	X			X		
6	X			X		
7	No se georreferencian					

X: Aplica para casos particulares, según la información de localidad y su ubicación en cartografía.

### 3.6 Validación de datos

La validación es el último paso dirigido a corregir posibles errores generados en la georreferenciación, sea en la captura de las coordenadas o digitación de estas. Se realiza con el fin de garantizar la fiabilidad de los datos georreferenciados antes de ser ingresados a la base de datos.

Este proceso consiste en identificar la correcta localización de las coordenadas, evaluando su correspondencia con la entidades de división político administrativa desde la más general a específica (país, departamento y municipio). De haber localidades que no corresponden, se debe proceder a identificar el error y corregirlo. Si las coordenadas coinciden con las entidades político administrativas, se entienden como validadas y pueden ser ingresadas a la base de datos.

Este proceso se puede hacer de manera manual, sin embargo se propone realizarlo de manera masiva para un conjunto de datos, haciendo uso de la herramienta “Verificación geográfica de registros biológicos para Colombia y países vecinos” desarrollada por el Laboratorio de Biogeografía y Bio-acustica (LABB) del Instituto Humboldt (veáse capítulo 5).

## **CAPÍTULO 4.**

### **PROCEDIMIENTOS SEGÚN TIPOS DE LOCALIDAD**

## 4.1 Procedimientos de georreferenciación

El objetivo de este capítulo es contextualizar los procedimientos metodológicos de la georreferenciación para casos concretos que responden a la alta heterogeneidad del estado de las localidades y que se han agrupado en niveles de calidad. Cada nivel contiene los parámetros para asignar las coordenadas y evaluar la incertidumbre, incluso recomendaciones de cómo abordar ciertos casos. En la parte final se ha dado un énfasis en el uso de la altura de manera independiente por la importancia de este parámetro para mejorar la georreferenciación.

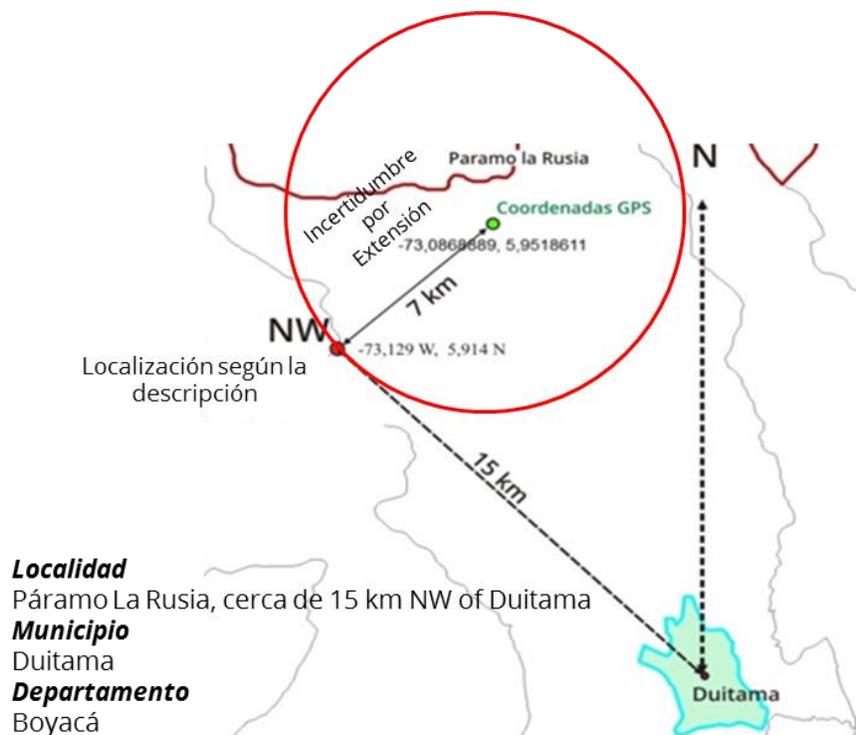
Para todos los diferentes tipos de localidades abordados, se usó de manera general la cartografía IGAC a escala 1:100.000 con un error gráfico de 0.30 mm. Según los planteamientos de los parámetros de incertidumbre compilados en el Capítulo 2, la incertidumbre por escala del mapa al utilizar esta cartografía es de 30 metros.

## 4.2 Localidades con coordenadas (Nivel 1)

Son localidades que por tener coordenadas se facilita identificar la localización del sitio de colecta. Sin embargo, en muchos de los casos las coordenadas no van acompañadas de datos esenciales, como: fuente de captura (gacetero, GPS, mapa, entre otros) y tipo de datum, razón por cual se requiere verificar su localización en cartografía.

En este nivel el procedimiento a implementar consiste en verificar la localización de las coordenadas con respecto a la descripción, los casos más frecuentes son los siguientes:

1. **Coordenadas próximas al sitio de la descripción:** En este caso se aceptan las coordenadas y se calcula la extensión de incertidumbre desde el punto de la coordenada hasta la localización de la localidad (véase Fig.20)



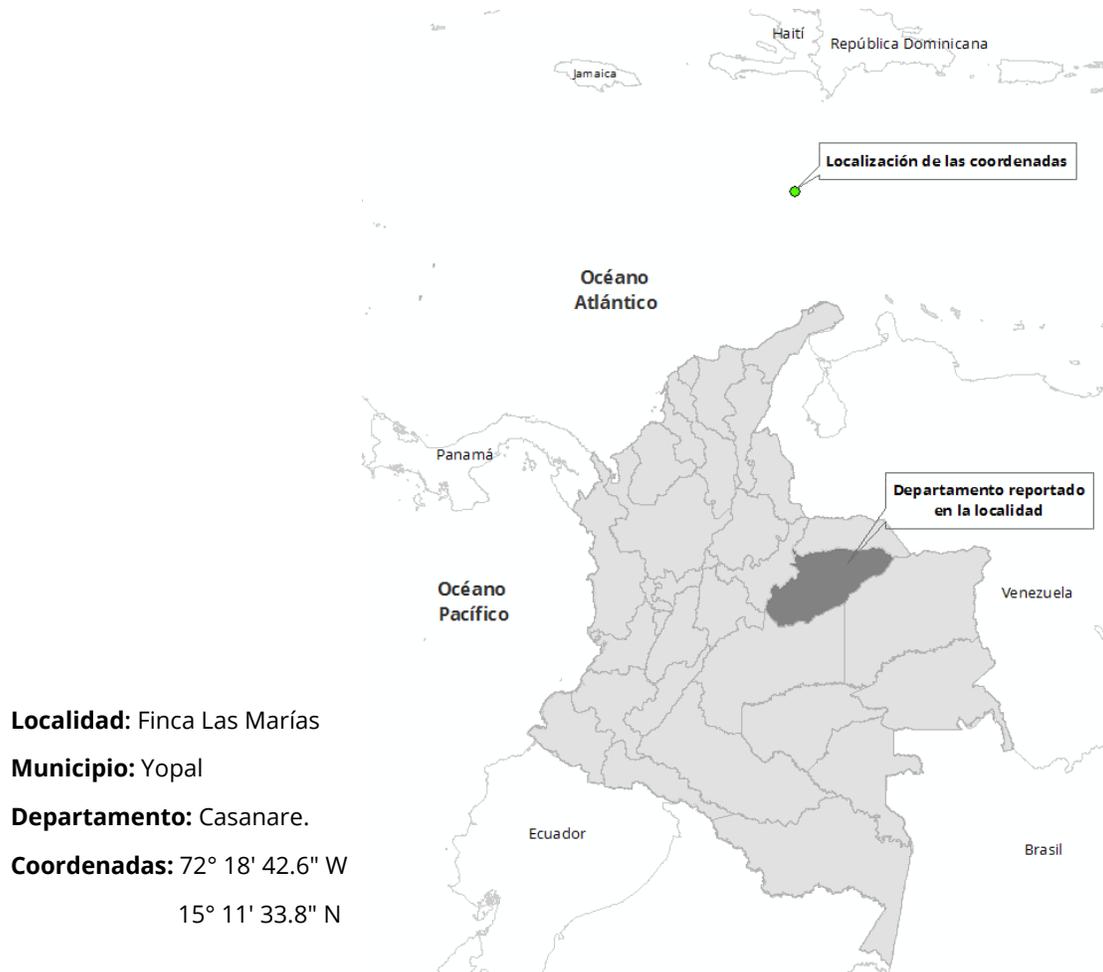
**Figura 20.** Coordenadas próximas con la localidad y cálculo de incertidumbre por extensión

Los parámetros de incertidumbre son:

- ✓ Incertidumbre por extensión: 7 km
- ✓ Incertidumbre por coordenadas: 1 m
- ✓ Incertidumbre por datum: 500 metros

**Nota:** La incertidumbre por datum se calcula solo si las coordenadas no reportan el tipo de datum usado para su captura, en caso de contar con esta información no se asume este valor de incertidumbre.

2. **Coordenadas que no coinciden con la localidad:** En este caso la ubicación de la coordenada no corresponde con la localidad descrita, por lo cual se requiere georreferenciar nuevamente siguiendo los parámetros de georreferenciación e incertidumbre que se aplican para las localidades sin coordenadas agrupados entre los nivel 2 y 7 según corresponda (véase Fig. 21).



**Figura 21.** Coordenadas que no coinciden con la descripción de la localidad.

## 4.3 Localidades sin Coordenadas

La característica principal de este conjunto de datos es que no tienen coordenadas y la georreferenciación se hace netamente a partir de la descripción de la localidad, a partir del estado de la información se consideran los procesos metodológicos para georreferenciar y evaluar la incertidumbre.

### 4.3.1 Localidades del Nivel 2

Estas localidades son aquellas que tienen información detallada y rastreada en cartografía (orientaciones y distancias) permitiendo mayor exactitud en la ubicación del sitio de colecta y por consiguiente menor grado de incertidumbre (véase Fig. 22). Las localidades más

frecuentes son: distancias en rutas, distancias en orientaciones o entidades geográficas puntuales que contienen netamente orientación o distancia. En caso de la distancia, esta puede estar reportada por un solo dato o por rangos. Si se trata de rangos siempre se trabaja con el dato promedio calculado, es decir, si la localidad específica “entre 2 km y 4 km”, el valor es 3 km.

La información potencial de estas localidades son entidades de referencia, orientaciones y distancias, por lo tanto, los parámetros incertidumbre asociados son: i) incertidumbre por orientación, ii) incertidumbre por la entidad geográfica de referencia iii) incertidumbre por unidad de medida e iv) incertidumbre por escala del mapa. Este último es un parámetro inherente a la localidad y que depende de la cartografía que se esté usando.

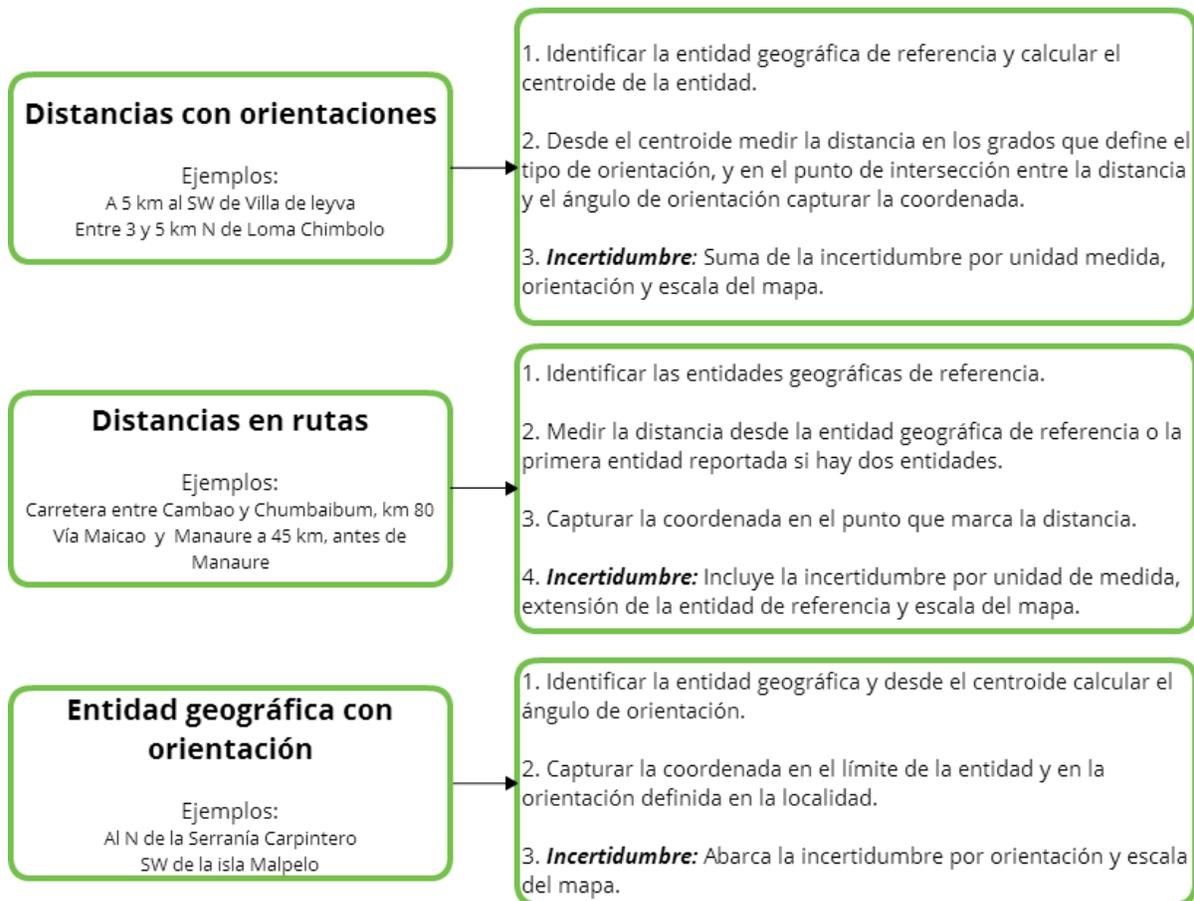


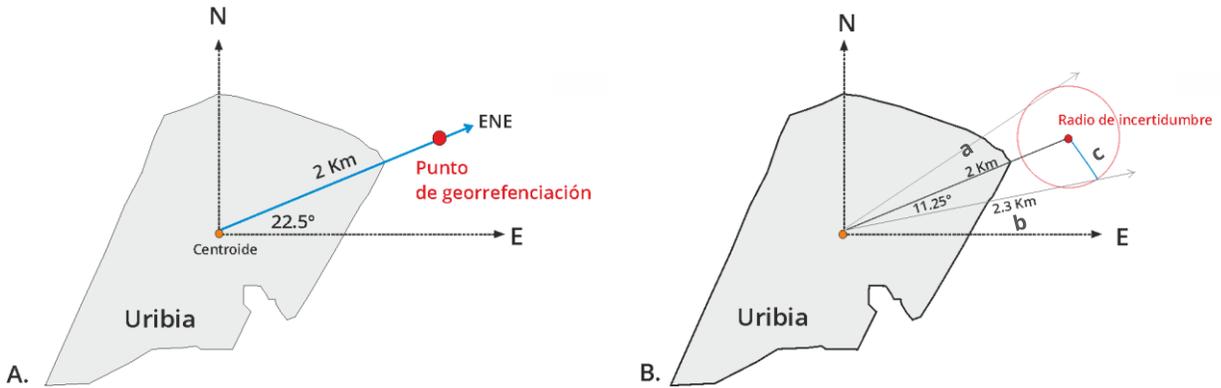
Figura 22. Procedimientos de georreferenciación de las localidades del Nivel 2.

Enseguida se abordan algunos ejemplos específicos de este nivel que explican con mayor detalle el proceso metodológico para la georreferenciación y cálculo de los parámetros de incertidumbre.

**Distancias en Orientaciones**

Para estas localidades la distancia y la orientación son los parámetros que definen la asignación de coordenadas y cálculo de la incertidumbre. La siguiente localidad tiene estas características, donde, el punto de georreferenciación se hace a los 2 km en la orientación ENE desde el centroide de la entidad geográfica (Véase Fig.23 A).

Localidad: 2 km al ENE de Uribía, departamento de la Guajira.



**Figura 23.** Localidad con distancia y orientación. A) Asignación de la coordenada B) Gráfica del cálculo de incertidumbre integrando, donde, a) 2 km b) 2.3 km y c) Radio de incertidumbre.

El cálculo integrado de la incertidumbre por orientación y distancia corresponde a la solución de un triángulo oblicuángulo, que se resuelve con el Teorema del coseno, donde, la distancia del lado **c** corresponde a la incertidumbre, el lado **a** equivale a 2 km y el lado **b** a 2.3 km. Este último que resulta de la sumatoria de la distancia (2 km) más el 15% del valor de la distancia (b'=0.3 km).

$$c = \sqrt{a^2 + (b + b')^2 - 2 * a * (b + b') * \cos \theta}$$

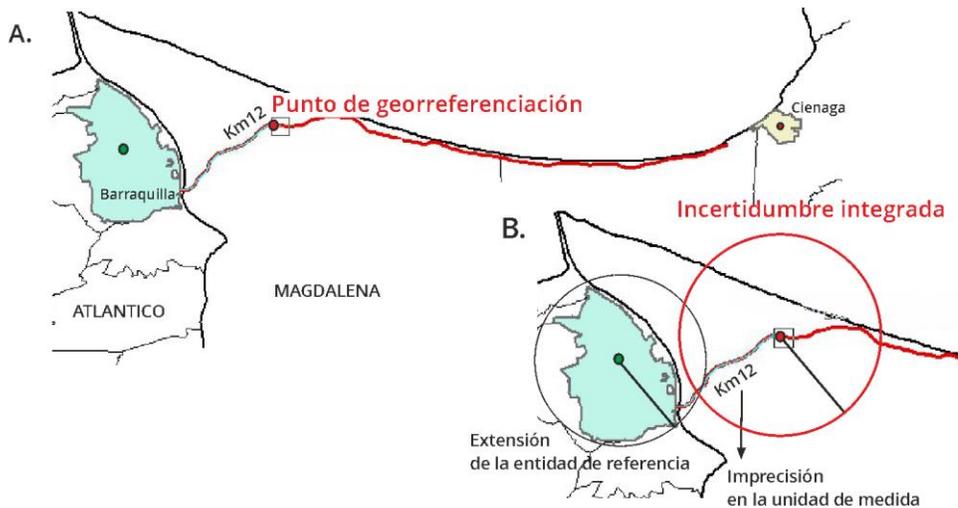
$$c = \sqrt{2^2 + (2 + 0.3)^2 - 2 * 2 * (2 + 0.3) * \cos 11.25^\circ} = 0.516 \text{ km} = 516 \text{ m}$$

En estas localidades, la incertidumbre por extensión no se usa, ya que sobre la misma área se solapa con la distancia generando sobrestimación de la incertidumbre.

### Distancias en rutas

En estas localidades el punto para la georreferenciación está definido por la distancia, valor que se mide desde el centroide de la entidad de referencia, tal como lo muestra el ejemplo de la Figura 24B. Sin embargo, cuando la distancia es pequeña ( $\sim \leq 4$  km) es probable que haya sido tomado desde los límites de la entidad geográfica pero cuya revisión se logra mediante la interpretación de la localidad en la cartografía.

Localidad: *Carretera desde Barraquilla a Ciénaga, Km 12*. Departamento del Magdalena.



**Figura 24.** Localidad con distancia en una ruta. A. Asignación de coordenadas B. Cálculo de incertidumbre.

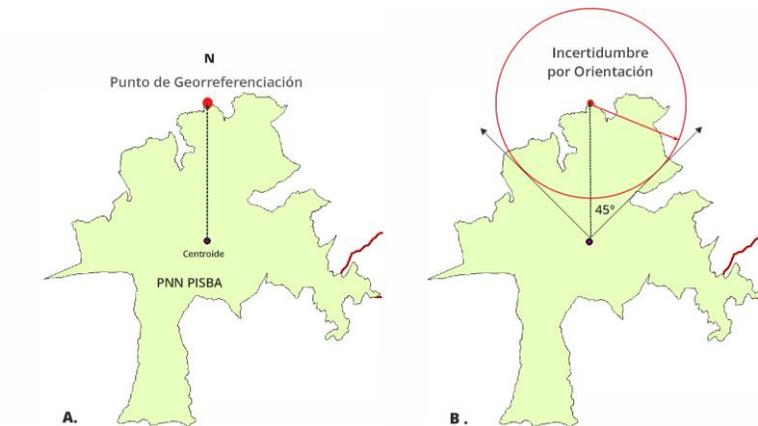
Los parámetros de incertidumbre que se incluyen para esta localidad son:

- ✓ Incertidumbre por distancia: 15% de 12km = 1.8 km = 1800 m
- ✓ Incertidumbre por extensión de la entidad de referencia: Equivale a la distancia desde el centroide hasta el límite más lejano de Barraquilla.
- ✓ Incertidumbre por escala del mapa: 30 m

### Entidad geográfica con orientación

Este tipo de localidades no tienen distancia pero hacen explícito que la colecta ocurre en una orientación específica que puede ser dentro de la entidad geográfica o afuera. La georreferenciación se hace en el límite de la entidad geográfica en la orientación descrita (véase Fig.25A)

Localidad: *Al N del PNN Pisba*



**Figura 25.** Entidad geográfica con orientación. A. Asignación de coordenadas B. Cálculo de incertidumbre.

La incertidumbre se define por:

- ✓ Incertidumbre por orientación: Equivale a la distancia en metros desde el punto de georreferencia hasta el límite definido por el ángulo de orientación, en este caso de 45°.
- ✓ Incertidumbre por escala del mapa: 30 m

### **Referente geográfico con solo la distancia**

En este tipo de localidades la distancia es el referente más preciso para ubicar el sitio de colecta, sin embargo, ante la ausencia de la orientación se asignan las coordenadas sobre el topónimo específico (véase Fig. 26).

Localidad: A 16 km del tapón; municipio de Cumaribo, departamento del Vichada



**Figura 26.** Localidad que contiene una distancia desde un referente geográfico. El punto rojo marca el sitio para la asignación de las coordenadas y la esfera la extensión de la incertidumbre.

Para la medida de incertidumbre, la distancia se interpreta como incertidumbre por extensión, ya que a 16 km a la redonda en cualquier orientación pudo haber ocurrido la georreferenciación. También se adiciona la incertidumbre por escala del mapa definida por la cartografía usada.

### Localidades con orientaciones ortogonales

La localidad se compone de dos orientaciones cada una con distancia, que al intersectarse permiten ubicar el punto de la colecta donde se hace la georreferenciación (véase Fig.27).

Localidad: Desde Zaragoza a 22 Km al S y 25 km al W, departamento de Antioquia.



**Figura 27.** Asignación de coordenadas mediante la intersección de orientaciones ortogonales.

Para el cálculo de la incertidumbre por distancia y orientación, se puede aplicar el *Teorema de Pitágoras*  $h = \sqrt{a^2 + b^2}$  que se usa para resolver triángulos rectángulos. Para este caso, **h** es la incertidumbre, **a** y **b** surgen de la sumatoria de las incertidumbres por distancia y extensión de la entidad de referencia (Zaragoza). El lado **a** se define por la descripción 25 km al W de Zaragoza y el lado **b** por 22 km al Sur de Zaragoza, la explicación de este cálculo se especifica continuación:

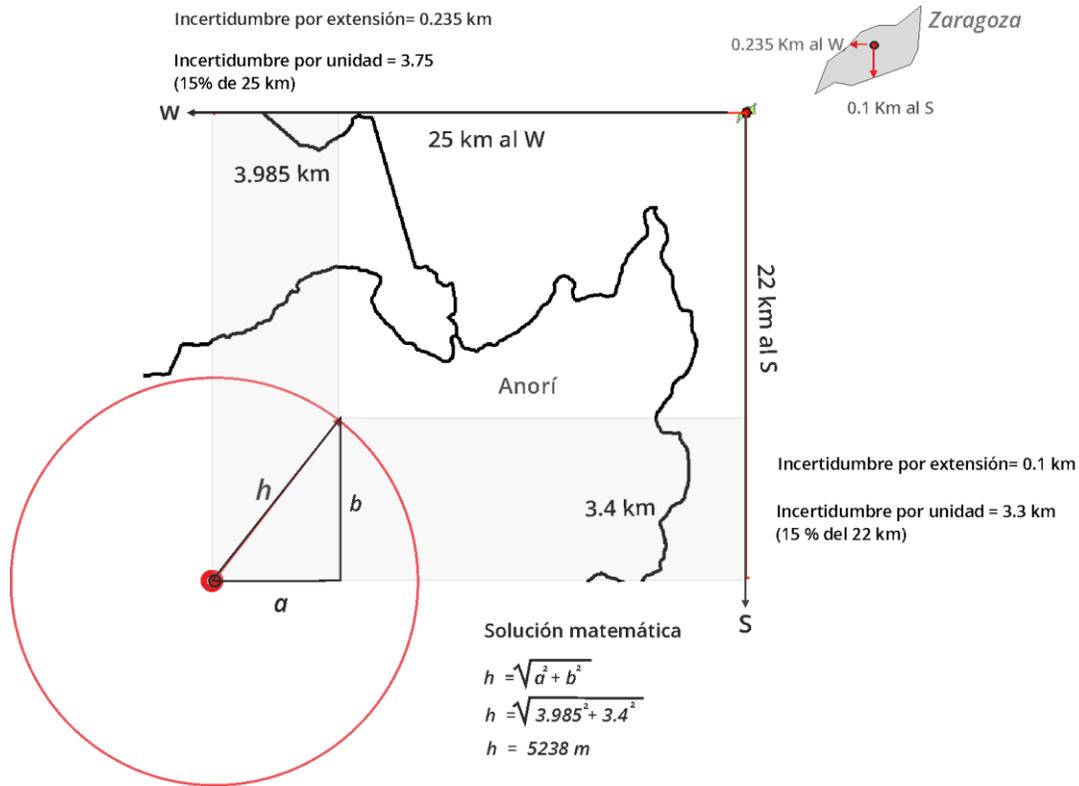


Figura 28. Tratamiento matemático para el cálculo de incertidumbre de localidades con orientaciones ortogonales

### 4.3.2 Localidades del Nivel 3

Abarca entidades geográficas específicas pero sin datos de distancias ni orientaciones que permitan hacer un rastreo cartográfico más puntual del sitio de colecta. Algunos tipos de localidades más frecuentes son: i) Localidades que describen entidades geográficas conectadas por rutas. ii) Entidades geográficas no delimitadas en la cartografía como veredas, cerros, altos, serranías, corregimientos, fincas, entre otros, y iii) Datos sin descripción del sitio de colecta y el único referente es el municipio.

Solo dos parámetros de incertidumbre se tienen en cuenta para estas localidades, la incertidumbre por extensión de la entidad de referencia y escala del mapa. La incertidumbre por extensión tiene algunas variaciones que dependen del tipo de localidad (véase Fig. 29).

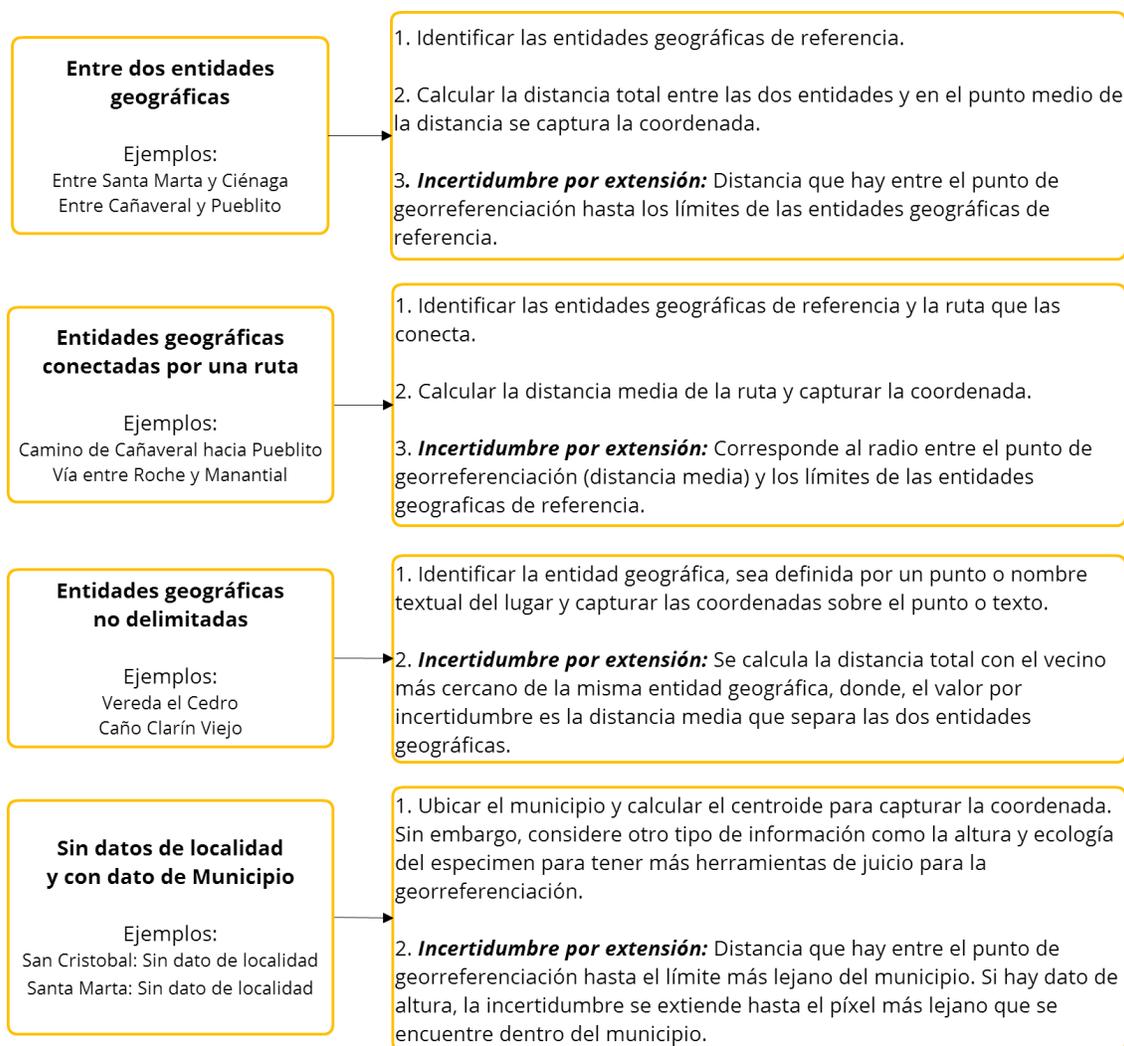
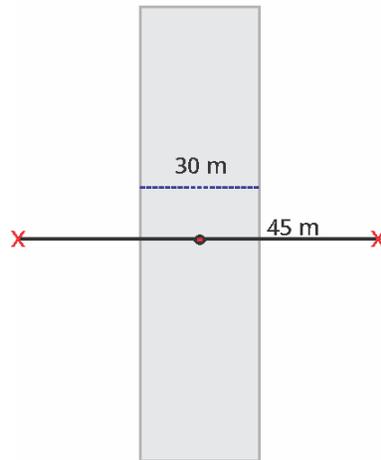


Figura 29. Procedimientos de georreferenciación de las localidades del Nivel 3.

### **Entidades geográficas conectadas por rutas o ríos**

La descripción de estas localidades contiene entidades geográficas conectadas por vías o ríos y se asignan las coordenadas sobre la entidad de conexión. Sin embargo, si es una vía o camino, se recomienda que el punto no se ubique exactamente sobre la ruta, sino que se desplace a un punto a 30 metros a los lados de la vía<sup>9</sup> (véase Fig. 30).

<sup>9</sup> Según el decreto núm. 2770 de 1953, aún vigente, la anchura mínima de la zona utilizable para las carreteras nacionales de primera categoría, será de treinta (30) metros (Invías 1953).



**Figura 30.** Distancia recomendable para la georreferenciación en una vía. Las X rojas definen los puntos para georreferenciar a los lados de vía.

El siguiente ejemplo, muestra el sitio para la georreferenciación cuando se trata de un río, asignando el punto sobre la distancia media del río (véase Fig. 31A).

Localidad: *Río Combeima, entre la vereda Juntas y vereda Pico de Oro, departamento del Tolima.*



**Figura 31.** Entidades conectadas por una ruta. A) Asignación de coordenadas B). Extensión de la incertidumbre.

Para esta localidad se incluyen la incertidumbre por escala del mapa y extensión de la entidad geográfica de referencia. Este último parámetro corresponde a la distancia media del río hasta el límite más lejano (véase Fig. 31B).



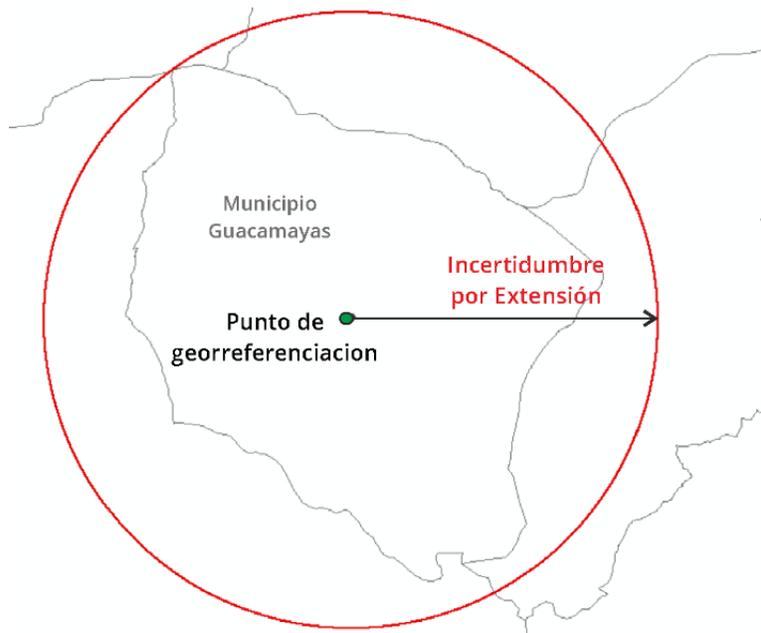
### **Sin datos de localidad, reporte Municipio**

Se ha propuesto que ante la falta de datos más específicos de localidad se use la división político administrativa más específica para georreferenciar, donde, la incertidumbre es el reflejo del estado de la localidad. La dificultad de hacer este tratamiento radica principalmente en zonas con relieve abrupto, donde en cortas distancias hay cambios de altura considerables, y por lo tanto, ambientes totalmente diferentes.

Alguna información como: la colección de referencia, fecha y ecología del espécimen puede ser de gran utilidad para mejorar la georreferenciación. Otro dato que puede ser de gran ayuda es la altura que por su relevancia se trata en un ítem aparte en este documento.

Cuando solo se cuenta con el dato de entidades de división político administrativa la asignación de las coordenadas se hace sobre el centroide del municipio, tal como lo muestra el ejemplo de la Figura 28.

Localidad: *Guacamayas-Sin datos de localidad*, departamento de Boyacá.



**Figura 33.** Georreferenciación y evaluación de la incertidumbre para un registro sin datos de localidad y con reporte del municipio.

La incertidumbre es considerable para este tipo de datos e incluye la incertidumbre por escala del mapa y extensión de entidad geográfica de referencia. La incertidumbre por extensión

equivale a la distancia desde el punto de georreferenciación hasta el límite más lejano del municipio.

### 4.3.3 Localidades del Nivel 4

Hace referencia a localidades que reportan áreas de gran extensión que pueden estar o no delimitadas en cartografía (áreas protegidas, resguardos, páramos, etc.) y que además contienen el dato de municipio o departamento. Con el fin de no generar sobrestimación del sitio colecta, se propone reducir al máximo estas localidades teniendo en cuenta la entidad político administrativa más específica (municipio o departamento). Los procedimientos generales para georreferenciación se han resumido en la figura 34.

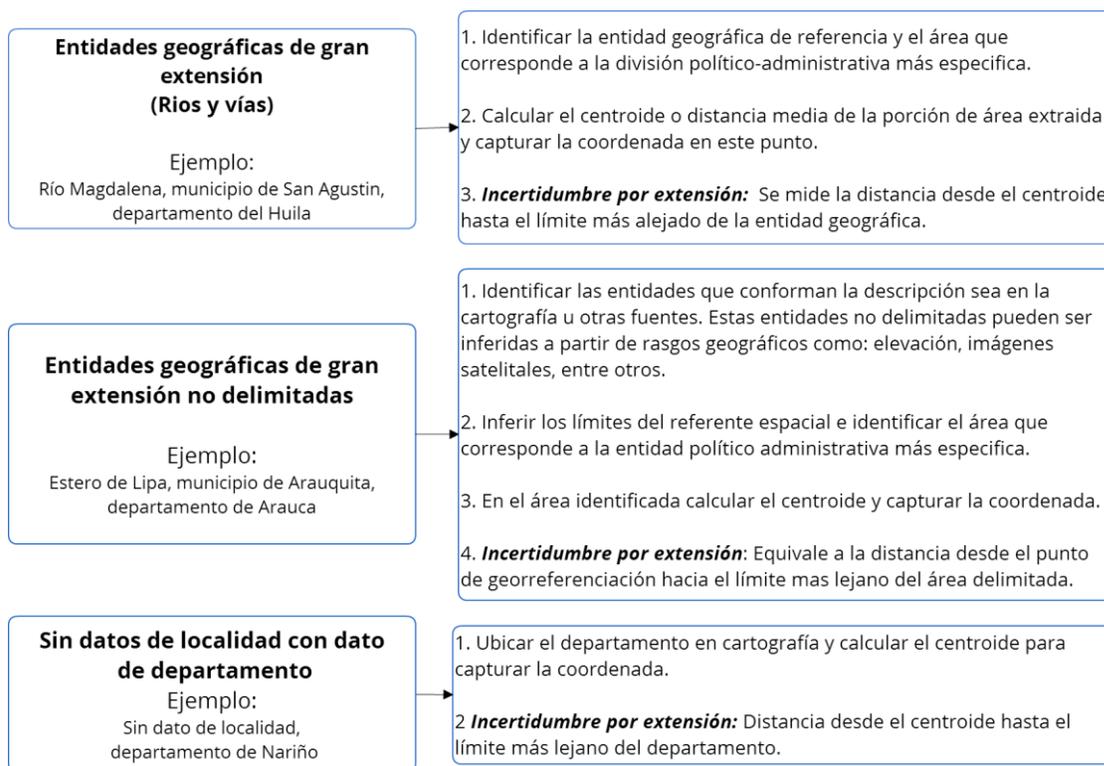


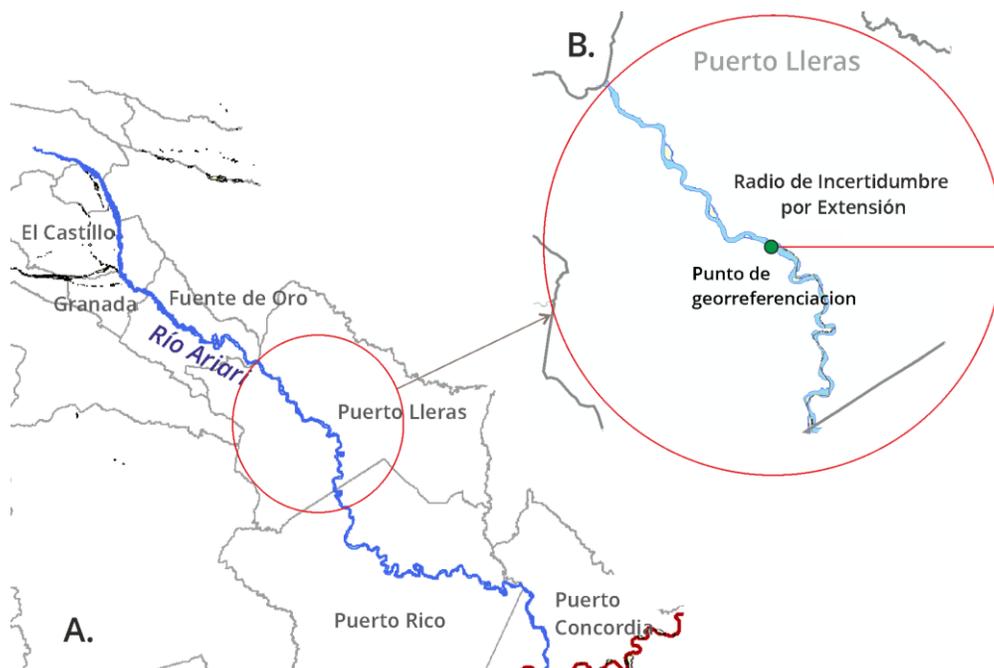
Figura 34. Procedimientos de georreferenciación de las localidades del nivel 4.

#### **Entidades geográficas que comparten varias entidades político administrativas**

La georreferenciación se hace extrayendo el área de la entidad geográfica reportada en la localidad que pertenezca a la entidad político administrativa más específica (municipio o

departamento), y sobre la cual se asigna la coordenada (véase Fig. 35). Este método puede aplicarse para otro tipo de áreas extensas como: parques naturales, humedales o áreas protegidas.

Localidad: *Río Ariari*, municipio Puerto Lleras, departamento del Meta.



**Figura 35.** Entidad geográfica que comparte varios municipios. A) Extensión total del río y división municipal B) Asignación de coordenadas y cálculo de incertidumbre.

Para esta localidad la asignación de coordenadas se realiza sobre el punto medio del segmento del río que hace parte del municipio de Puerto Lleras (véase Fig. 35B)

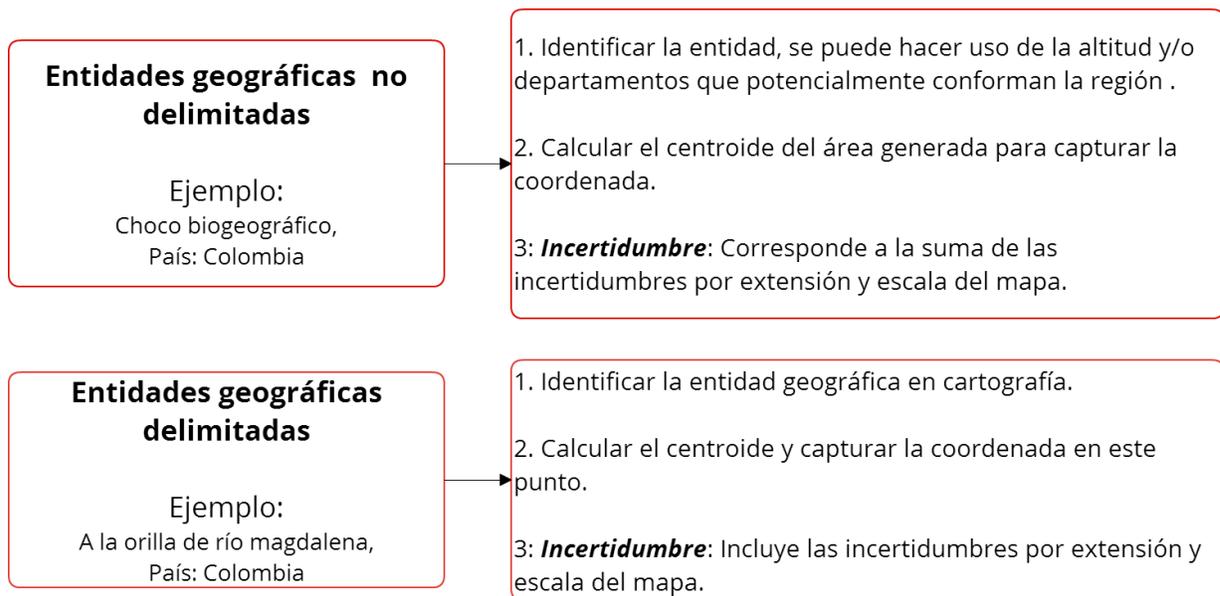
La incertidumbre corresponde a la suma de los parámetros de escala del mapa y extensión de la entidad geográfica de referencia. Este último se calcula midiendo la distancia en metros desde el punto de georreferenciación hasta el límite del río (véase Fig. 35B).

#### 4.3.4 Localidades del nivel 5

Son localidades con descripciones muy generales y no es posible afinar la georreferenciación porque solo reporta el país. Se puede presentar que la entidad geográfica en la cartografía está delimitada como las vías nacionales o ríos principales; o no delimitadas como regiones

naturales (Llanos orientales, Caribe, etc.), y cuyos procedimientos se han precisado para los dos casos (véase Fig. 36).

Al contar solo con referencias de entidades geográficas extensas, se parte del supuesto que en algún punto dentro del área de la entidad fue colectado el espécimen. Por lo tanto, la asignación de las coordenadas se hace sobre el centroide de la entidad y con una incertidumbre significativa, reflejo del estado de estas localidades (véase Fig. 36), y cuyo uso no se recomienda para análisis espaciales de gran detalle.

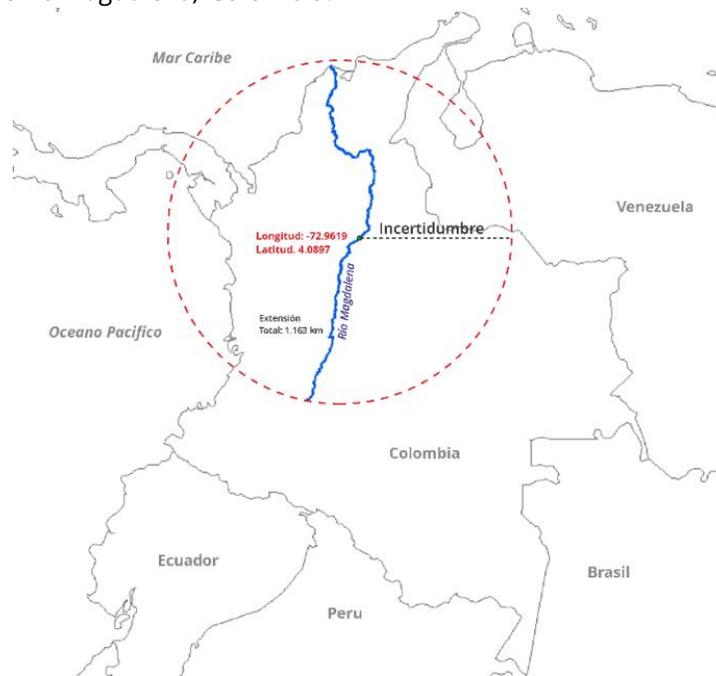


**Figura 36.** Procedimientos de georreferenciación de las localidades del nivel 5

En el ejemplo de la figura 37, el radio de incertidumbre se extiende desde el punto de georreferenciación hasta el límite más lejano del río Magdalena abarcando incluso áreas que no corresponden al mismo ecosistema.

Se puede explorar otro tipo de información para georreferenciar este tipo de localidades aunque puede tomar más tiempo del requerido, por ejemplo: la colección de referencia, ecología del espécimen e incluso el colector. No se recomienda hacer uso de la altura, ya que en áreas muy extensas se puede presentar el mismo dato en varios lugares, sin ningún parámetro que permita discriminar un dato de otro y que confiera mejor calidad a la georreferenciación.

Localidad: *Orilla de río Magdalena, Colombia.*

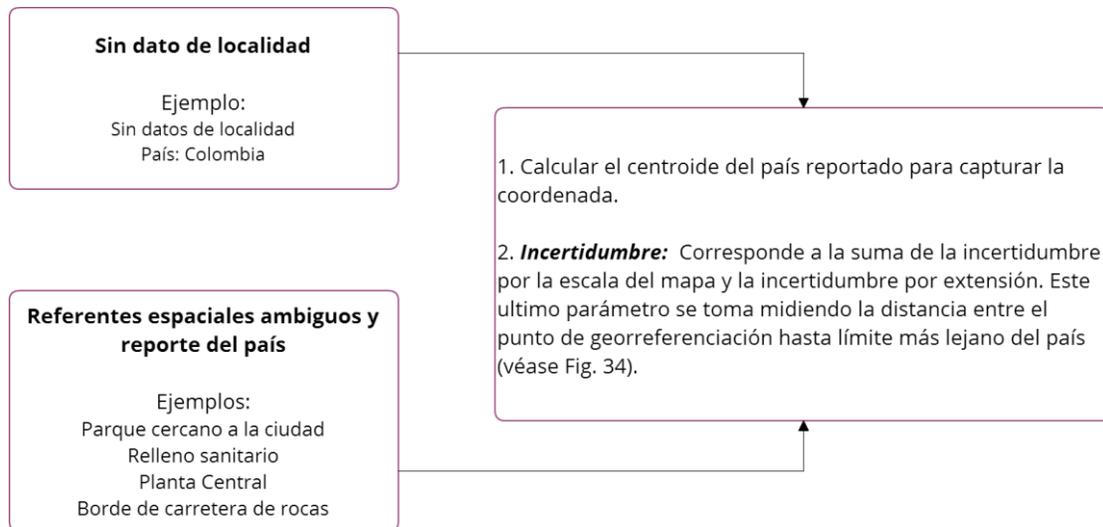


**Figura 37.** Georreferenciación e incertidumbre de una localidad de distribución nacional.

### 4.3.5 Localidades del nivel 6

Estas localidades presentan datos que no pueden ser rastreados en cartografía o no tienen datos de localidad y el único referente geográfico es el país. Se puede hacer uso del nombre del colector, tipo de colección y ecología de la especie para explorar otras posibilidades que contribuyan a mejorar la georreferenciación, sin embargo, como se ha precisado anteriormente la inversión de tiempo en su rastreo puede tomar mucho más tiempo y sin asegurar resultados óptimos.

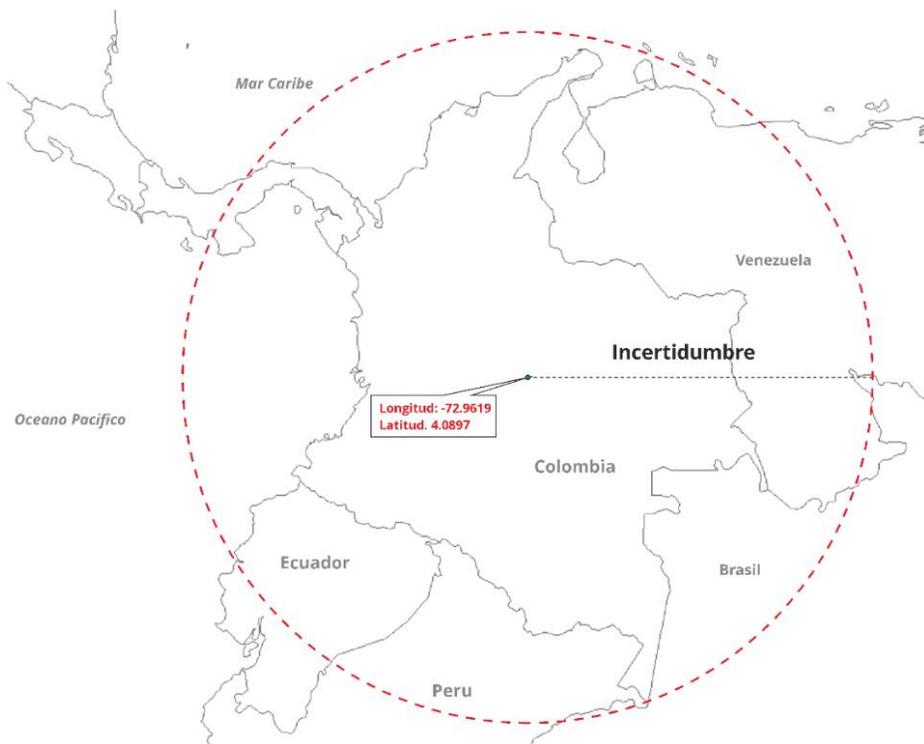
Todas las localidades de este nivel se georreferencian sobre el centroide del país y contiene una incertidumbre considerable reflejo del estado ambiguo de las localidades (véase Fig. 38). Aunque existe una alta incertidumbre, la georreferenciación se justifica por el uso que se les puede dar para análisis a nivel continental o mundial.



**Figura 38.** Procedimiento de georreferenciación de las localidades del Nivel 6.

El siguiente ejemplo, muestra que la captura de la coordenada se hace sobre el centroide del país y la incertidumbre incluso abarca áreas de otros países (véase Fig. 39).

Localidad: *Sin datos de localidad*, país Colombia.



**Figura 39.** Georreferenciación y cálculo de incertidumbre de registro donde el único referente espacial es el país.

### 4.3.6 Localidades del nivel 7

Son aquellas que tienen el nivel más bajo de calidad, cuentan con inconsistencias y ambigüedades en la información generando dudas en la veracidad de las descripciones. En este estado no se permite ubicar ningún referente espacial en la cartografía, y por lo tanto este tipo de localidades no se georreferencian.

En este nivel tienen lugar los registros producto de un decomiso donde la localidad reportada no corresponde al sitio de colecta del espécimen, también se encuentran localidades con información subjetiva con uso de términos como “posiblemente”, “quizá”, “proveniente de” en este estado la información de la descripción carece de veracidad, por lo tanto no se hace la georreferenciación (véase Fig. 40).

En estas localidades se recomienda hacer revisión de la etiqueta del espécimen para verificar si hubo información obviada.

Ejemplos:

- Andes?
- Localidad insegura posiblemente del Valle
- Sin dato de país - Pontevedra



**Figura 40.** Ejemplo de una localidad ambigua e inconsistente.

### 4.3 Tratamiento de la altura en la georreferenciación

El dato de altura es un atributo que puede mejorar considerablemente la georreferenciación aunque en los parámetros de calidad no está estipulado. Este dato se utiliza si es congruente con el área probable de colecta, ya que comúnmente se desconoce el método usado, ocasionando que la altura identificada en la cartografía difiera con el dato reportado por el colector.

La altura puede ir acompañada de cualquier tipo de localidad, sin embargo, si el área es muy grande, es decir, para localidades ambiguas o muy generales, la altura no es útil ya que el mismo dato se puede presentar en varios sitios y realmente su uso no genera mejor calidad de los datos.

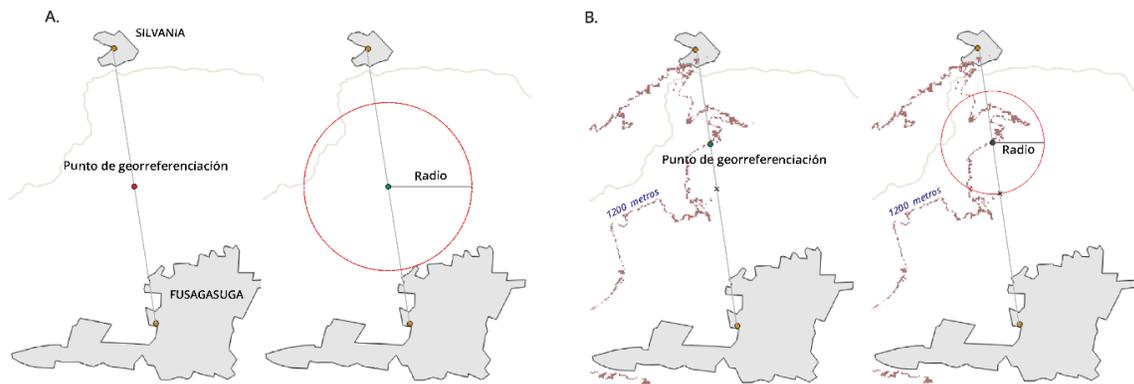
A continuación se muestra algunos tipos de localidades donde la altura permite una ubicación más precisa del sitio de colecta.

#### *Sitio de colecta ubicado entre dos entidades geográficas*

Si la localidad contiene solo dos entidades geográficas, tal como lo muestra la Figura 40A. La asignación de la coordenada se hace en el punto medio que separa las dos entidades, donde el sitio probable de colecta puede ser cualquier parte entre los límites de las entidades geográficas. Al incluir la altura se puede observar que el sitio para georreferenciar es más específico reduciendo la incertidumbre aun área más pequeña y que se expresa en una menor incertidumbre (véase Fig. 41).

El proceso metodológico para este tipo de localidades cuando hay un dato de altura consiste en asignar la coordenada sobre el pixel más cercano que intersecta la línea imaginaria y desde este punto se mide la distancia hasta el punto medio entre las dos entidades geográficas para calcular la incertidumbre (véase Fig. 41B).

Localidad: *Entre Silvania y Fusagasuga*, Departamento de Cundinamarca.

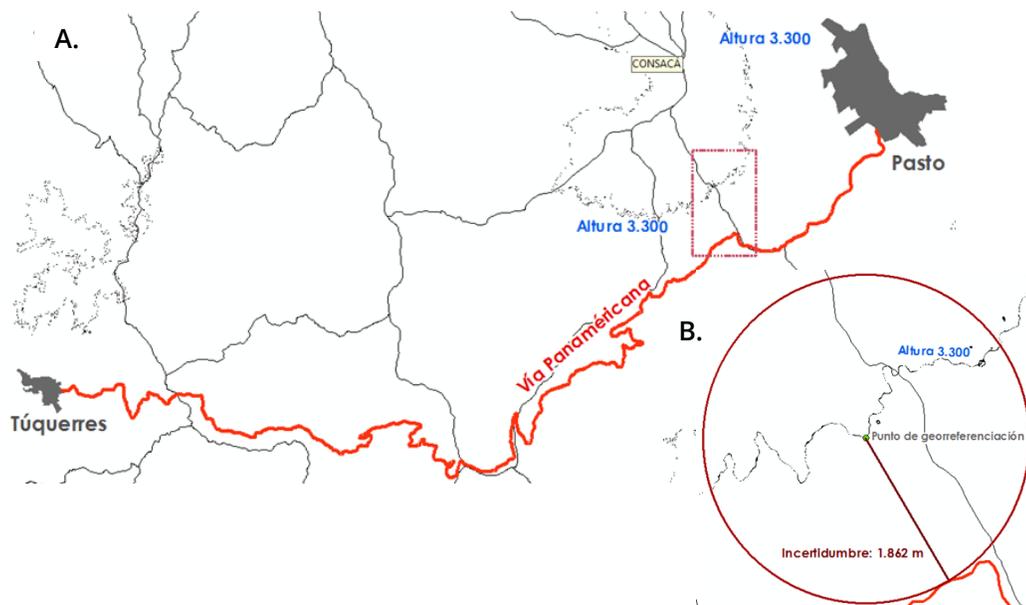


**Figura 41.** Comparación del proceso de georreferenciación de una misma localidad con ausencia y presencia de la altura. La línea recta corresponde a una línea imaginaria que conecta las dos entidades geográficas y que se utiliza para georreferenciar. A) Localidad sin datos de altura B) Localidad con una dato de altura de 1200 m.s.n.m.

**Entidades geográficas conectadas por una ruta**

En estas localidades, el sitio de colecta ocurre en alguna parte de la ruta que conecta dos entidades geográficas. Al incluir el dato de altura se identifica un área en particular, donde la vía es el medio de acceso para llegar al sitio (véase Fig. 42).

Localidad: *Vía entre Pasto y Túquerres*, departamento de Nariño, altura de 3300 m.s.n.m.



**Figura 42.** El sitio de colecta está en la ruta que conecta dos entidades. A) Vista general, el cuadro punteado define el sitio de interés B) Acercamiento del sitio de la colecta.

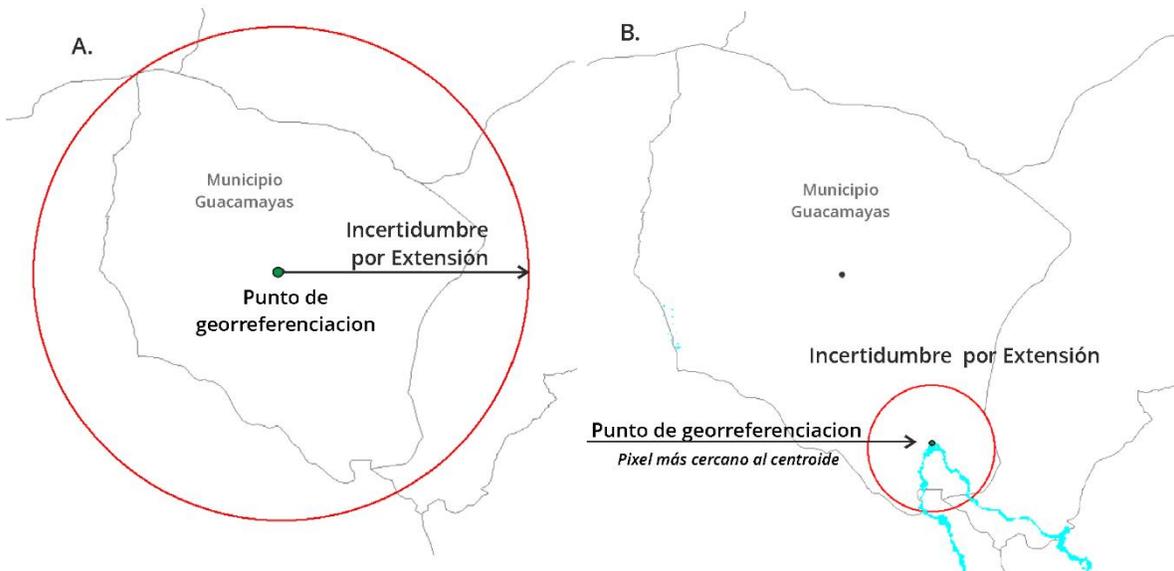
En este caso se tiene que la altura es próxima a la ruta, en términos metodológicos se puede escoger como punto de georreferenciación el pixel más cercano a la vía, y desde este hasta la vía se considera la extensión de la incertidumbre (véase Fig. 42).

### Registros sin datos de localidad

Según los parámetros expuestos, si no hay la localidad específica del sitio de colecta se procede a georreferenciar a partir de la información de la división político administrativa. También aplica para entidades geográficas de gran extensión como ríos, vías principales o áreas protegidas, etc. En estos casos, el dato de altura se puede ver limitado ya que en una vasta área es muy probable que exista el mismo dato de altura en distintas áreas. Sin embargo, a nivel de municipio el uso de la altura podría ser aplicable, solo sí, la altura se presenta en un área específica.

En el siguiente ejemplo se puede mostrar como la altura identifica un área particular donde es más probable que haya ocurrido el sitio de colecta (véase Fig. 43B) que sitio dado por el centroide del municipio (véase Fig. 43A).

Localidad: *Guacamayas-Sin datos de localidad*, departamento de Boyacá, 3400 m.s.n.m.



**Figura 43.** Georreferenciación y evaluación de la incertidumbre para un registro sin datos de localidad y con reporte del municipio. A. Sin dato de altura B. Con una altura de 3400 m.s.n.m.

## **CAPÍTULO 5.**

### **HERRAMIENTAS ÚTILES PARA LA GEORREFERENCIACIÓN**

## 5.1 Cálculo automático del centroide

La herramienta “**v.type**” ejecutada desde GRASS (Software que funciona con el lenguaje de programación Python) en Quantum gis (Qgis) ha demostrado ser efectiva para el cálculo del centroide. Esta herramienta realiza una relación de las coordenadas de cada uno de los vértices con el centroide, para luego determinar las distancias con respecto a los nodos más lejanos<sup>10</sup> y buscar un equilibrio geométrico que permita calcular el centroide del polígono.

El cálculo automático puede ser útil para ubicar el centroide de una capa completa de municipios, centros poblados o cualquier entidad representada por un polígono. De esta manera optimiza el tiempo ya que no es necesario calcular el centroide para un solo polígono sino para un conjunto de polígonos. Este mismo método también aplica para líneas que en la cartografía pueden representar un río o una vía, donde el centroide se ubica de manera automática en la distancia media de la entidad.

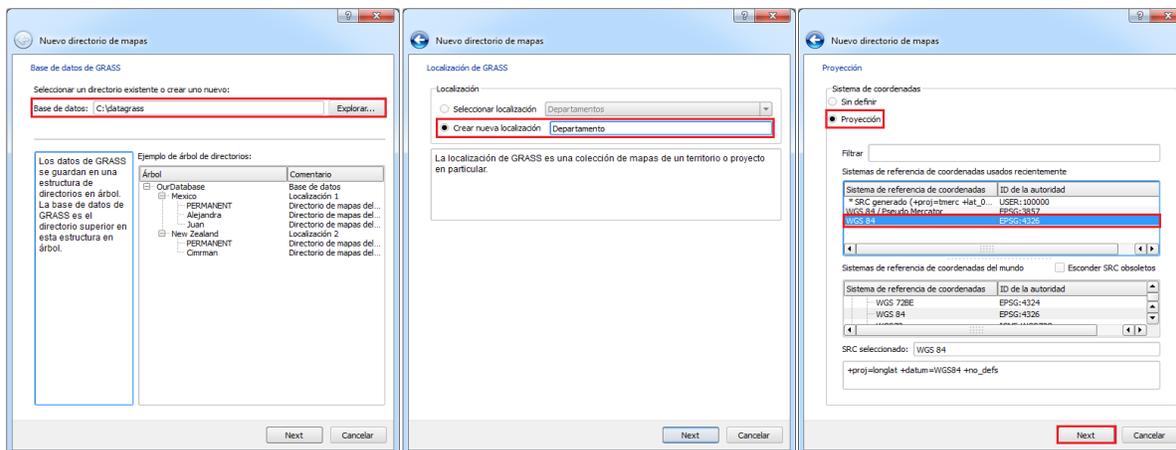
### 5.1.1 Elaboración del Workspace

Inicialmente se debe elaborar el espacio de trabajo en carpetas específicas para guardar la información de los procesos y resultados, en la dirección que defina para lo cual debe crear una carpeta con el nombre *Datagrass*, donde se guarda el archivo (Ejemplo: *Centros\_poblados.shp*). Recuerde que el Workspace solo se crea una vez.

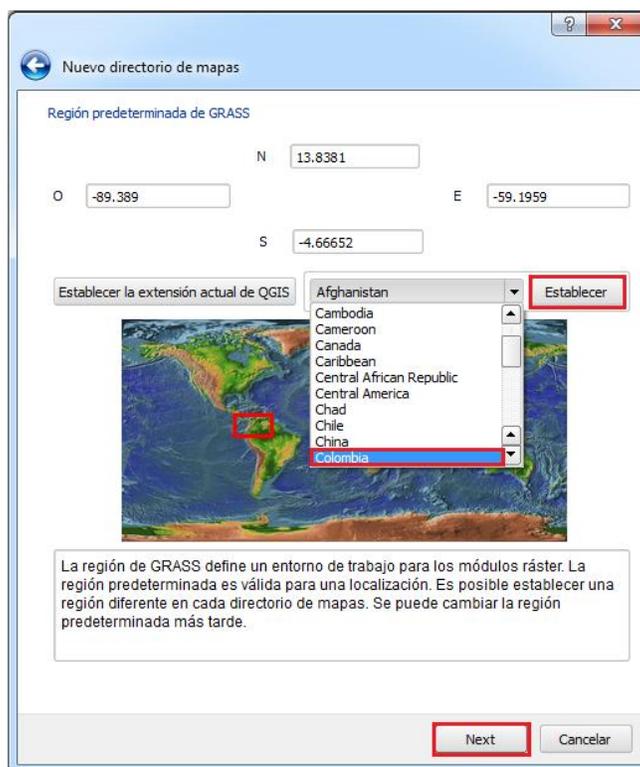
1. Siga los siguiente pasos:
  - a) Abra Q-gis y en la *Barra de menú*, seleccione *Complementos* y se despliega una pestaña, seleccione GRASS y luego *Nuevo directorio de mapas*. En la nueva ventana que aparece en la casilla *Base de datos*, de clic en *Explorar* y busque el directorio donde guardo los shapefile y haga clic en *Next*.
  - b) En el siguiente paso, haga clic en *Crear nueva localización* y escriba el nombre del nuevo archivo.
  - c) Posteriormente escoja la proyección WGS 84 y haga clic en *Next*.

---

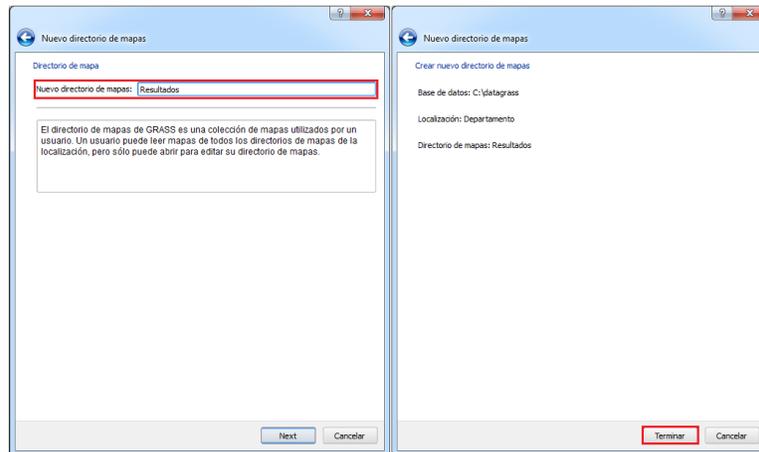
<sup>10</sup> <http://grass.osgeo.org/grass65/manuals/v.type.html>



- d. En la siguiente ventana se visualiza el globo terráqueo, donde puede constatar que el espacio geográfico sea el correcto, de lo contrario seleccione el país en *Establecer la extensión actual del QGIS* y haga clic en *Establecer* y *Next*.

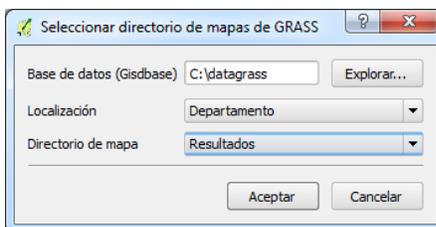


- e. En la ventana que aparece debe definir el nombre del archivo donde va a depositar los resultados. Finalmente, haga clic en *Next*, *Terminar* y *Aceptar*.

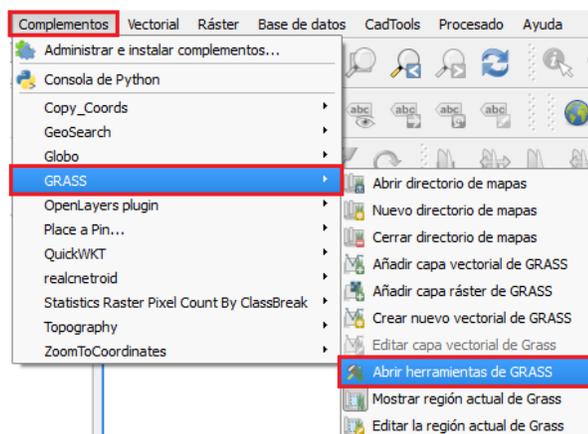


### 5.1.2 Creación de centroides

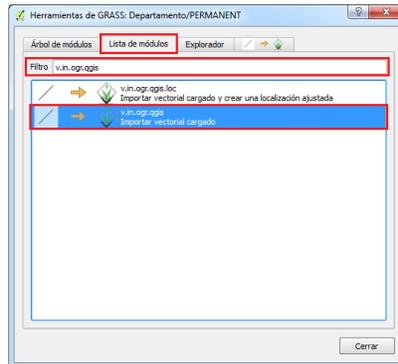
- Cargue y active la capa a la cual va calcular el centroide de los polígonos.
- Vaya a la barra de herramientas y seleccione la opción *Complementos* y le aparece una ventana que le permite seleccionar el directorio de mapas de GRASS, haga clic y seleccione el archivo.



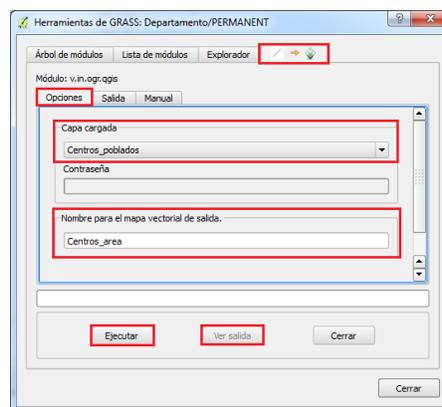
- Ahora diríjase la barra de herramientas, elija *Complementos*, seleccione *GRASS* y haga clic en *Abrir herramientas de GRASS*.



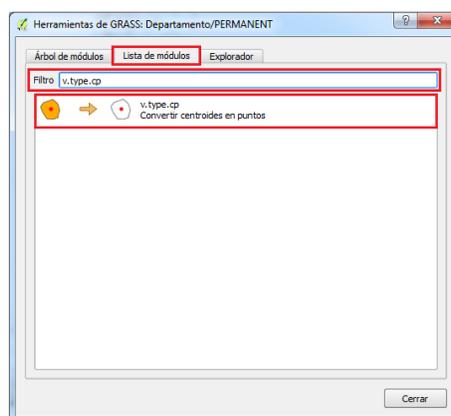
- d. En la ventana *Herramientas de GRASS* busque la pestaña *Lista de módulos* y busque la herramienta *v.in.ogr.qgis* y haga clic.



- e. En la siguiente pestaña clic en la casilla *Opciones*, escoja *Capa cargada* y haga Clic *Centros\_poblados*. Después defina el nombre de mapa vectorial de salida, asigne el nombre de *Centros\_area* y haga clic en *Ejecutar* y *Ver salida*.



- f. Regrese a la pestaña *Lista de módulos* y busque la herramienta *v.type.cp* que le permite convertir área a centroide y cierre la ventana.



- j. Para finalizar de clic derecho sobre la capa y seleccione *Guardar como* y revise que se encuentre creado.

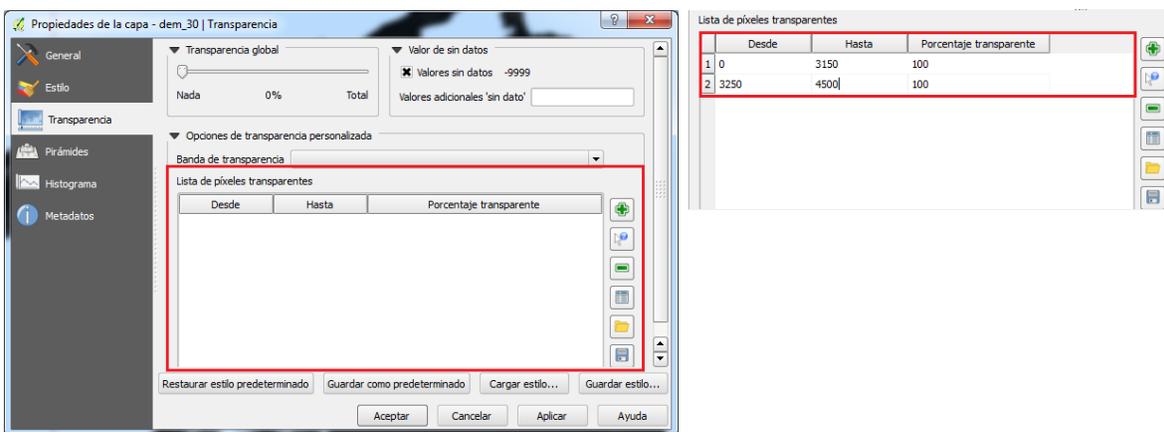
## 5.2 Manejo de archivos raster en QGIS

El formato ráster es uno de los formatos usados en los sistemas de información geográfico, donde se maneja el concepto de resolución, definido como el detalle de la información con la que se trabaja, entre más grande sea el pixel o malla, más general es la información del ráster.

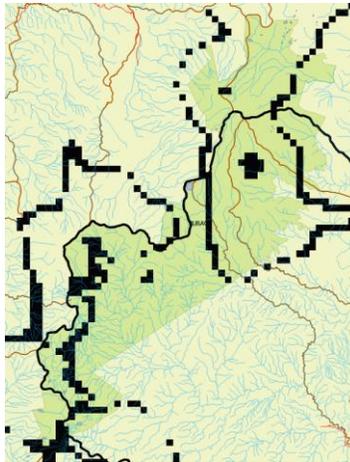
En Q-gis existe una herramienta muy fácil de usar que permite discriminar datos dando transparencia a datos específicos que no se quieren visualizar, generando una selección por rangos de información.

El uso de la transparencia en los archivos raster es útil para visualizar los píxeles de la altitud reportados en la localidad permitiendo realizar procedimientos de georreferenciación más detallados, siempre y cuando la altitud coincida con el área donde se ubica la localidad.

El acceso a esta herramienta se realiza haciendo clic derecho sobre la capa y seleccionando la opción *propiedades de la capa*, posteriormente haga clic en la opción de *Transparencia*. En la Banda de transparencia puede seleccionar los píxeles de altitud que NO se desea visualizar haciendo clic en el símbolo (+). Enseguida le aparecen las casillas para establecer los rangos a los cuales debe dar un 100% de transparencia, tal como se muestra en las siguientes imágenes.



Después de realizar este procedimiento podrá visualizar en cartografía los píxeles de altitud que reporta la localidad ya que será el rango al que no se aplica transparencia, con esto podrá obtener una localización más precisa del sitio de colecta.



### 5.3 Herramienta para la validación de localidades georreferenciadas

Esta herramienta consiste en un *script* que permite la verificación geográfica para Colombia y países vecinos, además de la taxonómica de registros biológicos. Este *script* desarrollado en R fue elaborado por el Laboratorio de Biogeografía y Bio-acustica (LABB) del Instituto Humboldt y cuya ejecución cuenta con un manual de procedimientos<sup>11</sup>. Se recomienda su uso ya que puede validar un amplio conjunto de datos y arroja resultados en archivos fácilmente migrables a Excel.

Es importante aclarar que esta herramienta trabaja a partir del registro biológico, donde, cada localidad corresponde a un espécimen colectado en un sitio específico con una relación uno a uno. En el proceso de validación que se quiere ejecutar para la georreferenciación no se cumple, ya que una localidad puede tener asociado varios registros biológicos con una relación de uno a varios. Por tal razón, se requiere hacer algunas modificaciones en el formato de la base de datos de ingreso para que se pueda ejecutar el *script*.

En la siguiente tabla, se especifican los campos, el tipo de formato y la descripción del contenido que debe llevar el archivo para ser reconocido por el *script*.

---

<sup>11</sup> Verificación de registro biológicos: <https://code.google.com/p/verificacion-registros-biologicos/>

**Tabla 12.** Campos y formato de validación para ejecutar el script de verificación geográfica de registros biológicos.

Campo BD Localidades georreferenciadas IAvH <sup>12</sup>	Nombre del campo BD Validación <sup>13</sup>	Tipo	Descripción	Ejemplo
LocalityID	ID	Numérico	Identificador único de la localidad	5445
Colección	Nombre	Texto	Nombre de la colección que contiene la localidad. Si no es estrictamente necesaria esta información se recomienda reemplazar el nombre por NA_NA	NA_NA
País	pais	Texto	Nombre del país en siglas. Para Colombia usar CO	CO
Departamento	departamento	Texto	Nombre del departamento de la localidad, según IGAC.  Se recomienda no usar tildes	Cesar
Municipio	municipio	Texto	Nombre del municipio donde se ubica la localidad, según IGAC. Se recomienda no usar tildes	Manaure Balcon del Cesar
Latitud_georref	latitud	Numérico	Coordenada latitud de la georreferenciación en grados decimales	10.366439
Longitud_georref	longitud	Numérico	Coordenada longitud de la georreferenciación en grados decimales.	-72.903228
Fecha georreferenciación	Fecha_inicial	Texto	Equivale a la fecha de la georreferenciación, en este caso no es informativo, se mantiene el campo pero se reemplaza la fecha por NA en todos los registros	NA
Localidad_georref	Localidad	Texto	Nombre de la localidad depurada, sin tildes.	Corregimiento de Manaure carretera a la frontera más o menos cerca de Sabana Rubia

<sup>12</sup> Estos campos corresponden al formato de base de datos de las colecciones IAvH y no se deben registrar en la base de datos de validación. Se muestra en esta tabla con fines ilustrativos de equivalencia con el formato exigido para usar el código de verificación en R.

<sup>13</sup> El nombre de los campos en la base de datos usada para la validación deben ser idénticos al mencionado en la tabla. En los campos que no contengan información (vacíos) se debe usar la sigla NA.

## BIBLIOGRAFÍA

- Baker, R.J. et al., 1998. Bioinformatics, museums and society: Integrating biological data for knowledge-based decisions. *Museum of Texas Tech University*, 187, p.4.
- Beaman, R., 2003. Biogemancer: Automated georeferencing for natural history collections. Available at: <http://www.biogemancer.org>.
- Beaman, R., Wieczorek, J. & Blum, S., 2004. Determining Space from Place for Natural History Collections: In a distributed digital library environment. *D-lib Magazine*, 10(5), p.9.
- BOLFOR & ETSFOR, 1999. *Cartografía y uso de la tecnología GPS* Ramiro Duc., Santa Cruz, Bolivia.
- Cato, P.S., 1991. The value of natural history collections in Latin American conservation. In *Conservation Education*. Conservation Education, pp. 416–430.
- Chapman, A.D., 2005. *Principles and methods of data cleaning: Primary species and species-Occurrence data* G. B. I. Facility, ed., Copenhagen.
- Chapman, A.D. & Wieczorek, J., 2006. *Guide to best practices for georeferencing BioGeomanc.*, Copenhagen.
- CONABIO, 2008. Georreferenciación de localidades de colecciones biológica. Manual de procedimientos. In México: Comisión Nacional para el Conocimiento y Uso de Biodiversidad (CONABIO), p. 177.
- DANE, 2001. *Evolución histórica de la división político administrativa-DIVIOLA-de Colombia*, Santafé de Bogotá.
- Fernández-Coppel, I.A., 2001. *Localizaciones geográficas: La Proyección UTM(Universal Transversa Mercator)*,
- GBIF, 2008. Taller calidad de datos. In GBIF-ES, ed. *Dónde y Cuando?* España.
- Gómez, F.J., 1992. Los sistemas de información geográfica, su importancia y su utilidad en los estudios medioambientales. In Historia, ed. *Cuadernos de Sección*. pp. 455–465.
- Graham, C.H. et al., 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in ecology and evolution*, 19(9), pp.497–503.
- Guisan, A. & Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, pp.993–1009.
- Guralnick, R. & Neufeld, D., 2005. Challenges building online GIS services to support Global Biodiversity mapping and analysis: lessons from the mountain and plains database and informatics project. *Biodiversity Informatics*, 2, pp.56–69.

- IGAC, 1996. *Diccionario geográfico de Colombia, tomo I*, Bogotá, Colombia.
- IGAC, 1994. *Resolución 64*, Bogotá.
- IGAC, 2004. *Tipos de coordenadas manejados en Colombia*,
- Invias, 1953. *Normas sobre las vías públicas nacionales\_ decreto 2770 de 1953 Vías*, Bogotá, Colombia.
- Munt, D.D., 2013. Taller de GEOLOCATE. , p.33. Available at: [http://www.gbif.es/ficheros/TallerSIGGeolocate\\_10/2010\\_01\\_Intro\\_georreferenciacion.pdf](http://www.gbif.es/ficheros/TallerSIGGeolocate_10/2010_01_Intro_georreferenciacion.pdf): 20/ 11/013. [Accessed September 18, 2014].
- Muñoz, M., Hernández, J.C. & Colín, J., 2004. Georreferenciación de las localidades de colectas biológicas. *Biodiversitas*, 54, pp.8–15.
- Murphey, P.C. et al., 2004. Georeferencing of museum collections: A review of problems and automated tools , and Informatics Initiative ( Mapstedi ). *Phyloinformatics*, 3, pp.1–29.
- Overton, J.M. et al., 2002. Information pyramids for informed biodiversity conservation. *Biodiversity and Conservation*, 11, pp.2093–2116.
- Proctor, E.J., 2004. *Reducing variation in georeferenced locality descriptions*. San Francisco State University.
- Proctor, E.J., Blum, S. & Chaplin, G., 2004. *A software tool for retrospectively georeferencing specimen localities using ArcView. Georeferencing natural history collectio localities*,
- Rhoads, A.F. & Thompson, L., 1992. Integrating herbarium data into a geographic information system: requirements for spatial analysis. *Taxon*, 41, pp.43–49.
- Rojas, O.R. et al., 2002. La colecta científica en el neotrópico: el caso de aves de México. *Ornitología Neotropical*, 13, pp.209–214.
- Soberón, J. & Peterson, T., 2004. Biodiversity informatics: managing and applying primary biodiversity data. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 359(1444), pp.689–98.
- Sua, S., Mateus, R.D. & Vargas, J.C., 2004. *Georreferenciación de registros biológicos y gacetero digital de localidades I. A. von Humboldt*, ed., Bogotá, Colombia.
- TDWG, 2011. *Elementos Darwin Core : Una guía de referencia rápida\_versión 2.0*, Bogotá.
- Tobar, D., 2002. *Informe de la curaduría de la colección de mariposas "Ernesto Wolfgang Schmidt-Mumm" del IAvH*, Bogotá, Colombia.

USGS, 1999. Standards for digital line graphs. In *National Mapping Program Technical Instructions*. United States, p. 65.

Wieczorek, J., 2001. MaNIS/HerpNet/ORNIS Georeferencing Guidelines. *University of California, Berkeley*. Available at: <http://manisnet.org/GeorefGuide.html> [Accessed March 22, 2014].

Wieczorek, J., Guo, Q. & Hijmans, R., 2004. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science*, 18(8), pp.745–767.

## OTROS RECURSOS

Elementos de Darwin Core: Una guía de referencia rápida.

[http://www.sibcolombia.net/c/document\\_library/get\\_file?p\\_l\\_id=33134&folderId=33501&name=DLFE-301.pdf](http://www.sibcolombia.net/c/document_library/get_file?p_l_id=33134&folderId=33501&name=DLFE-301.pdf)

Instituto Alexander von Humboldt-IAVH, 2004. Manual de descarga del modelo digital de elevación. Documento de apoyo al taller georreferenciación de localidades en colecciones biológicas. Versión 0.0. 11 p. Responsables: Díaz, S. R & Ramírez, C.

Instituto Alexander von Humboldt-IAVH. 2004. Uso de la calculadora MaNIS para evaluar la incertidumbre por coordenadas. Documento de apoyo al taller georreferenciación de localidades en colecciones biológicas. 7 p. Responsables: Jojoa L. M & Ramirez, C.

Nasa. EOSDIS: NASA's Earth Observing System Data and Information System, Reverb-ECHO. Modelo de elevación digital. Disponible en:

[http://reverb.echo.nasa.gov/reverb/#utf8=%E2%9C%93&spatial\\_map=satellite&spatial\\_type=rectangle](http://reverb.echo.nasa.gov/reverb/#utf8=%E2%9C%93&spatial_map=satellite&spatial_type=rectangle)

MaNIS/HerpNet/ORNIS Georeferencing Guidelines. Disponible en:

<http://manisnet.org/GeorefGuide.html>

Verificación de registros biológicos. Disponible en: <https://code.google.com/p/verificacion-registros-biologicos/>

Tulane University Biodiversity Research Institute. 2005. GEOlocate. Disponible en:

<http://www.museum.tulane.edu/geolocate/support.html>

TDWG-Subgroup on Biological Collection Data. Disponible en:

<http://www.bgbm.org/tdwg/acc/Software.htm>. Fecha 04/08/2014.