CNDDB Data Use Guidelines

Table of Contents

Limitations of the GIS dataset, and how to get more details.

The CNDDB GIS dataset is a balance of including enough of the key information to be useful, but holding back from including so much information that the file size would be overly large. The downside of this is that if you are trying to do detailed occurrence-level analysis, the GIS dataset does not contain all of the information that could be helpful.

To remedy this, you can utilize the dbf files that are part of the RareFind3 install to pull in additional details. In particular, the occ.dbf file contains some important comment fields that often contain critical additional details, such as threats (THRTCOM), ecology (ECOCOM), and general occurrence history (GENCOM).

This occ.dbf file and others are part of the "RareFind Data" package that is available from the CNDDB website under "Maps and Data" > "Monthly CNDDB Data":
http://www.dfg.ca.gov/biogeodata/cnddb/

To best use these files, you will also need the data dictionary that describes each field within these files. This documentation is part of the "RareFind and CNDDB Documentation" package in the same location as mentioned above.

Both of these packages should get installed on your PC under C:\cnddb3. The occ.dbf file will be under C:\cnddb3\rfdata, and the data dictionary document will be part of the full set of help files at C:\cnddb3\rfdocs\03data_dictionary.html.

You can view the entirety of the help files by starting with C:\cnddb3\rfdocs\index.html. Also of importance to review is the table relationship diagram C:\cnddb3\rfdocs\03relationships.html.

To use the occ.dbf or other dbfs, add the file to your ArcMap session and use the EONDX field to link (join) it to the GIS dataset. As a cautionary note, do not use Excel to utilize these dbf files. Excel has a record limit of 65536, and some of these files have more records than that. Bring them into ArcMap directly, or use Access.
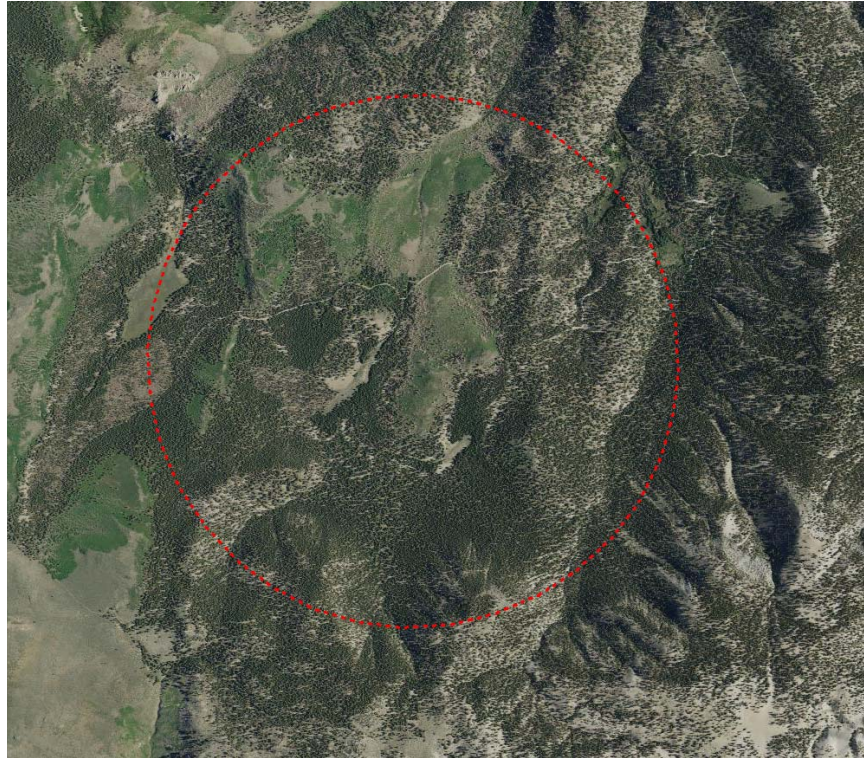
Simply double-click to open the data dictionary file or any of the other help files in a browser.

Why are there so many different shapes and sizes of CNDDB occurrences?

The CNDDB collects information from a wide variety of sources and makes this data available in a standardized text and graphic format. The different sizes of circles and polygons indicate the level of location detail provided in the source document(s). We currently use 10 graphic accuracy classes:

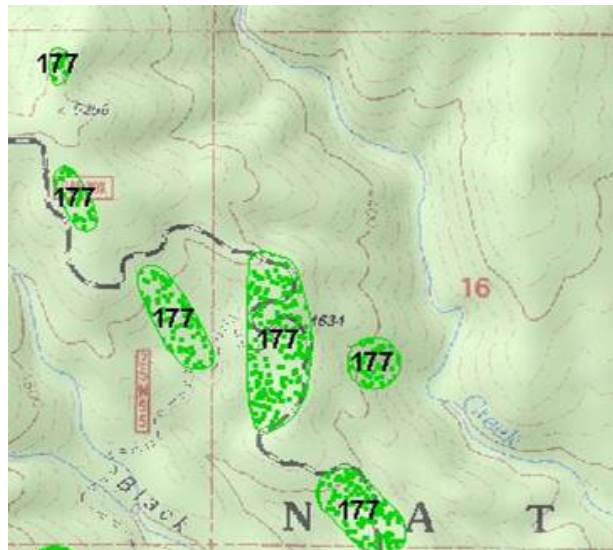| Accuracy Class	Description
1	Specific bounded area with an 80 meter radius
2	Specific, non-circular bounded area
3	Non-specific bounded area
4	Non-specific, circular feature with a 150 meter radius (1/10 mile)
5	Non-specific, circular feature with a 300 meter radius (1/5 mile)
6	Non-specific, circular feature with a 600 meter radius (2/5 mile)
7	Non-specific, circular feature with a 1000 meter radius (3/5 mile)
8	Non-specific, circular feature with a 1300 meter radius (4/5 mile)
9	Non-specific, circular feature with a 1600 meter radius (1 mile)
10	Non-specific, circular feature with a 8000 meter radius (5 miles)

The larger the circular feature, the more vague the location. The occurrence below is for a northern goshawk. The location was described only as "Sentinel Meadow." It is mapped as a 1 mile radius circle (Accuracy Class 9). If and when we receive updated information with a more precise location, the occurrence will be remapped more specifically.
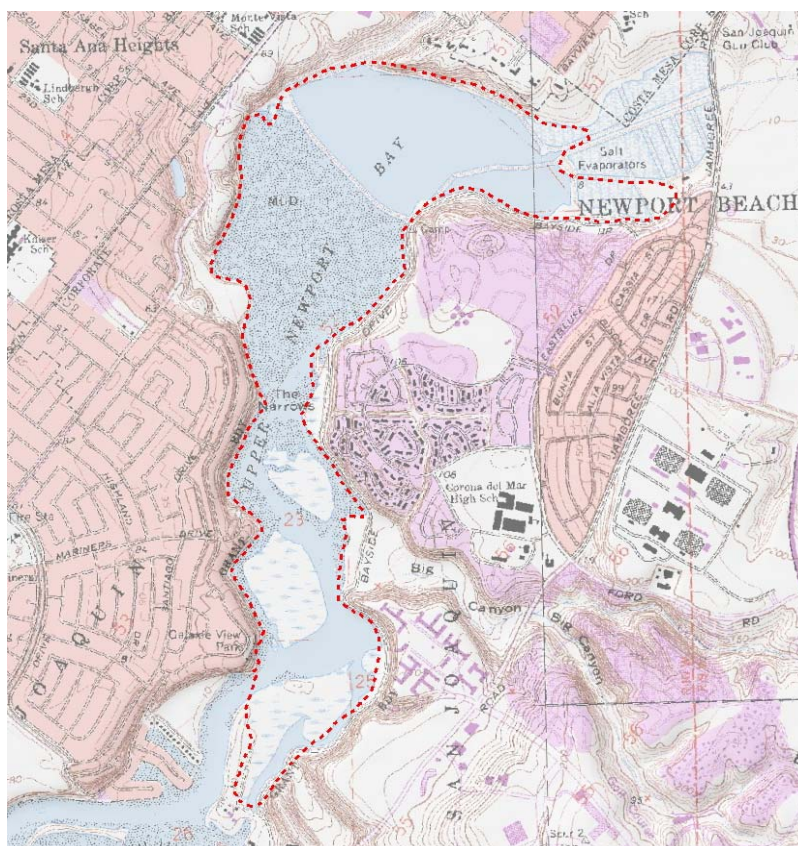

80 meter radius circles (Accuracy Class 1) represent very precise location data. For the occurrence below the mapped feature represents a specific pool where California tiger salamanders were found.

Non-circular features or irregular bounded areas (polygons) can be specific or non-specific in our terminology. The occurrence below, number 177 for Quincy lupine, is mapped as a series of specific bounded areas, indicating exactly where the plant was seen to occur at the time of the survey (Accuracy Class 2). This is still a single occurrence, made up of multiple parts; each part is less than 0.25 miles from the next nearest one. The CNDDB's default separation distance for occurrences is 0.25 miles.

Non-specific bounded areas (polygons) are used when we don't know exactly where the element was or may be found at that location (Accuracy Class 3). The occurrence below is for light-footed clapper rail in Upper Newport Bay. The entire bay is mapped since the exact location of the birds within the bay can vary from year to year and cannot be pinpointed.

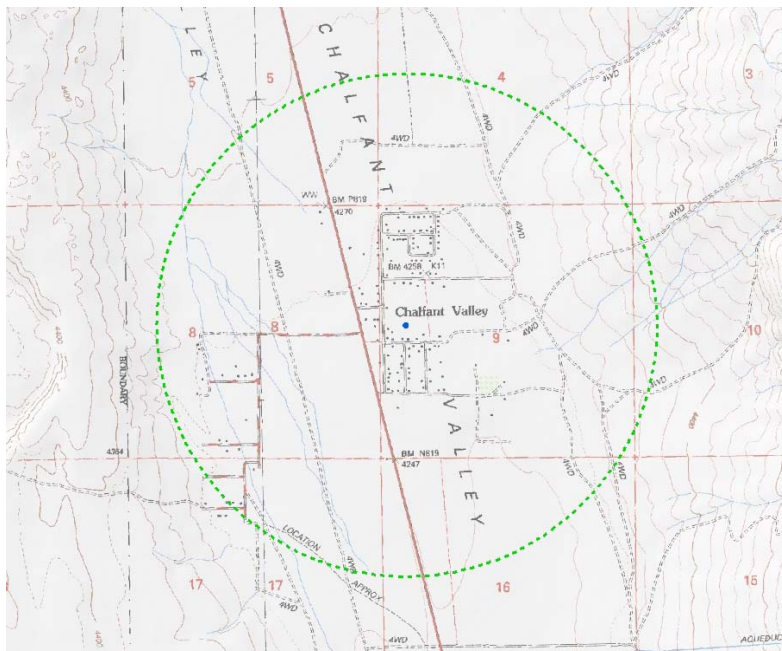## *What is the difference between the CNDDB polygon layer and the CNDDB point layer?*

The point layer should only be used for gross graphic representations of the CNDDB for large areas or areas which are densely populated with occurrences. The polygon layer houses more accurate occurrence location representations and information, and must be used when performing spatial analysis.

In the point layer, there is one point for each Element Occurrence. This point **is not** the point of the actual occurrence. Many CNDDB users report that they use the point layer because they think the point is the actual point of the "site" and that the polygon is generated from the point. In fact, the opposite is the case.
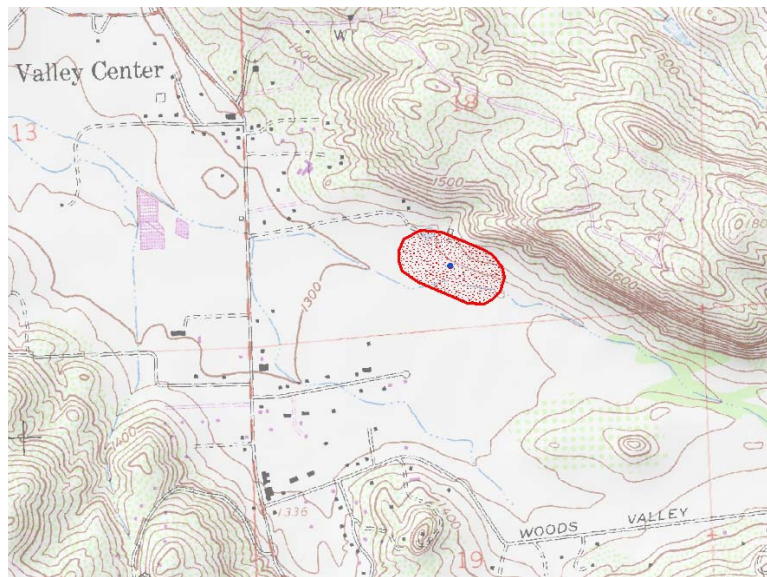
When a CNDDB biologist digitizes an Element Occurrence as a polygon or multiple polygons (Accuracy Class values 2 or 3), the point is the interpreted "center" of the occurrence as determined by the biologist.

For occurrences requiring circles with varying radii (Accuracy Class values 1, 4-10), the point is generated at the center of the circle. But the circle represents a level of vagueness in the occurrence, and the center point is not a reflection of the true occurrence location.
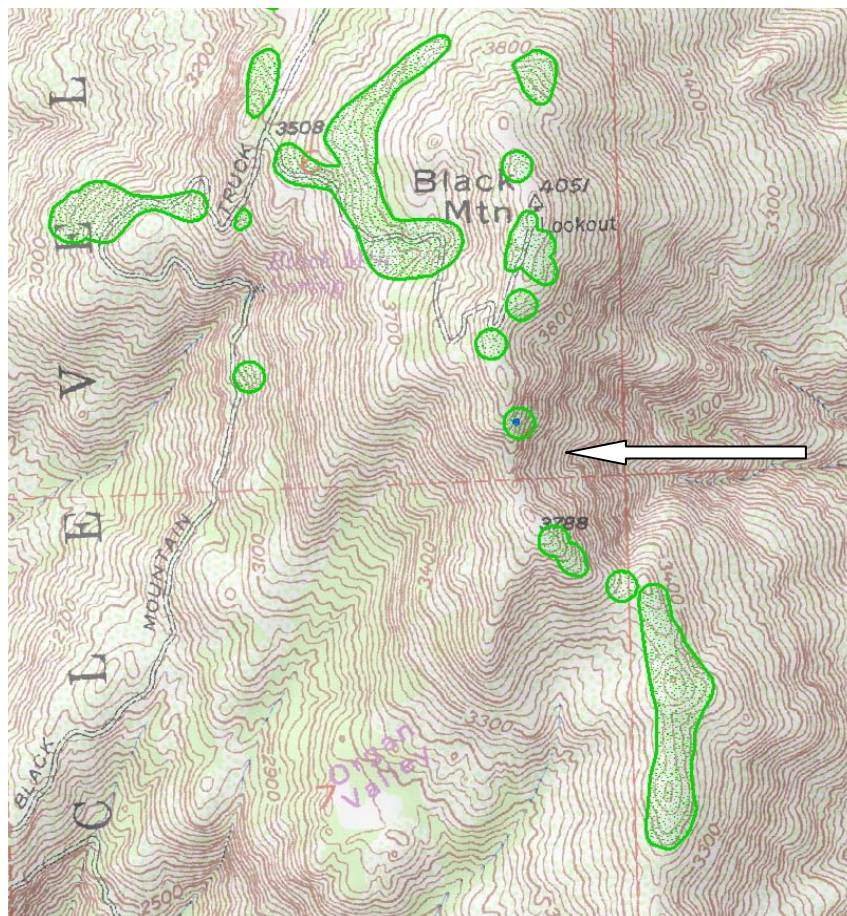
For the circular features the point is always the centroid.

For single bounded area (polygon) features, the point is approximately in the center of the polygon.

However, for Element Occurrences that consist of multiple bounded area (polygon) features, there is still only one point because it is still one Element Occurrence. The point is arbitrarily placed in one of the bounded area parts by the staff mapper. This example shows why the point layer is only useful at a very small (zoomed out) map scale. Too much information is lost at the larger (zoomed in) map scales.
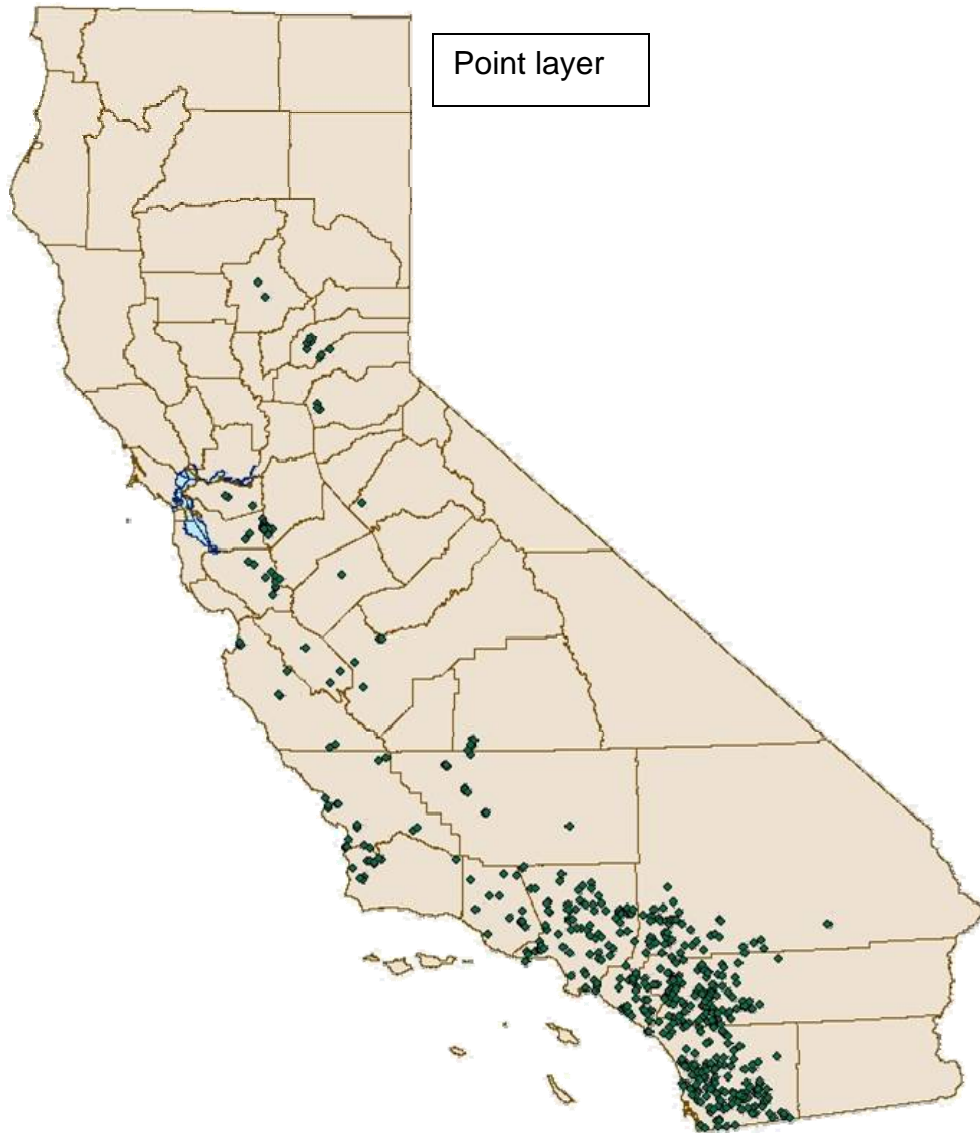
When should you use the point layer instead of the polygon layer?

The only time the point layer should be used is when the map scale is so small (zoomed out) that you can't see all of the polygon features. The point layer should never be used for any type of analysis.

The two maps below show the statewide distribution of coast horned lizard. The point layer is appropriate to use in this case because when the polygon layer is used, many of the smaller occurrences don't show up very well.

Polygon layer

Point layer

Can CNDDB information be displayed on maps that will be publicly available?

Depending on the scale of the map, CNDDB data can be displayed or summarized in some form. The concern is that, while it is important that the CNDDB information is available to those whose job it is to conserve species, there is the very real possibility that some people will use the detailed location information to do harm to a species or its habitat. Because of the sensitivity of the data, we try to limit the level of location detail that is made readily available to the public.
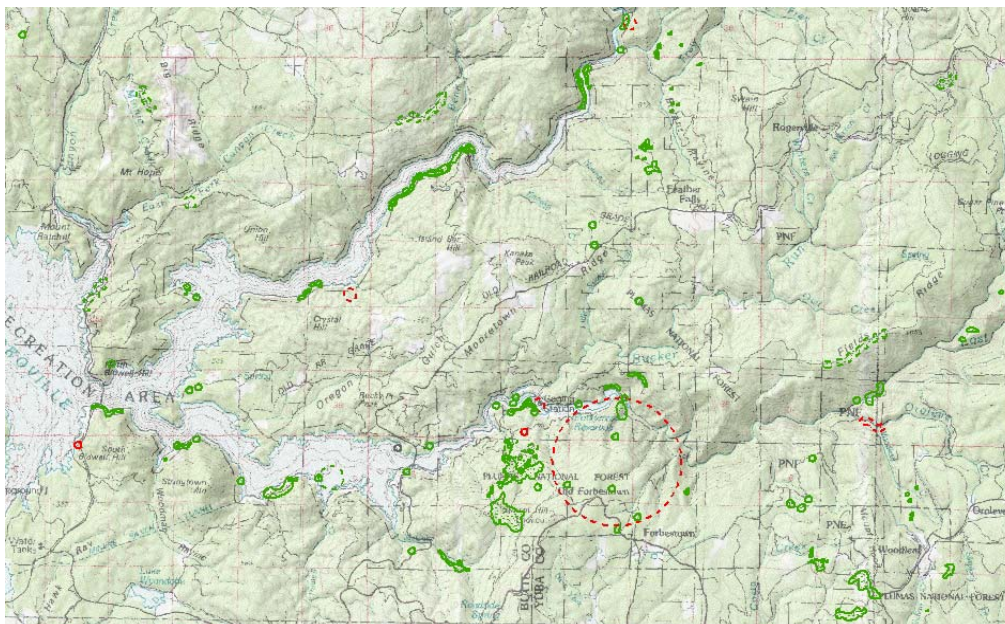
Please abide by the following guidelines when displaying CNDDB data:

Symbology: Always use the standard CNDDB symbology. The CNDDB avl files can be found in the CNDDB3\gis folder that you download with the Data Updates from the DFG Data Portal https://nrmsecure.dfg.ca.gov

Disclaimer/Date: Always include the CNDDB disclaimer on your map, along with the month/year of the dataset you are using. The disclaimer text is:

"CNDDB version MM/YYYY. Please Note: The occurrences shown on this map represent the known locations of the species listed here as of the date of this version. There may be additional occurrences or additional species within this area which have not yet been surveyed and/or mapped. Lack of information in the CNDDB about a species or an area can never be used as proof that no special status species occur in an area."

For maps at a scale larger than 1:350,000: At any scale larger (more zoomed in) than 1:350,000 the polygon layer should not be shown on a public map. This is because at scales larger than 1:350,000, there is enough detail for a user to fairly easily determine exactly where a species is located and that is what we are trying to prevent. The map below is at a scale of 1:100,000. Section lines are visible and it would be easy to find these locations. Therefore, this is too detailed for use as a publicly displayed map.
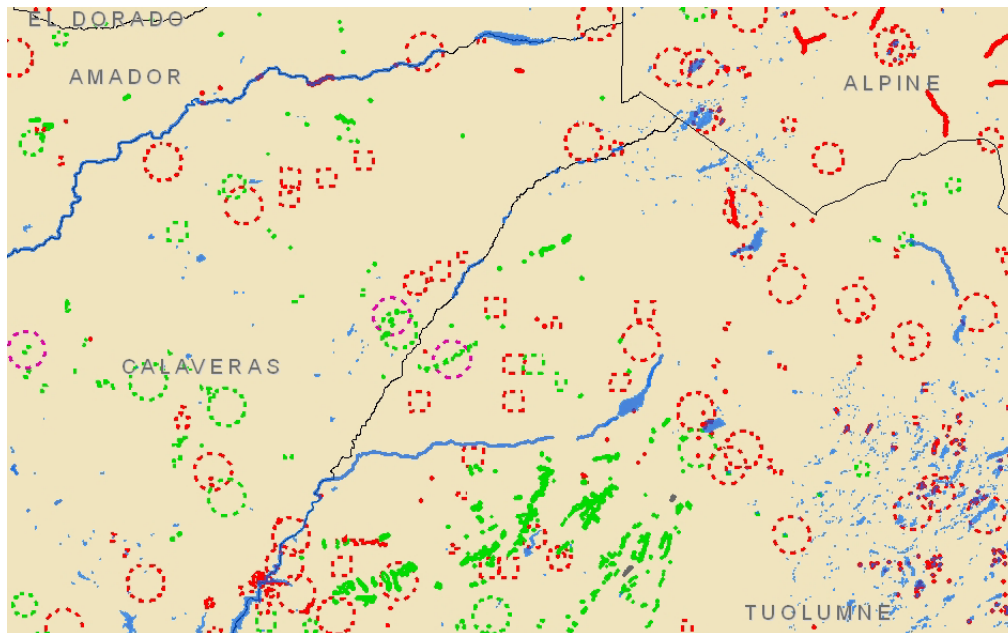
At map scales larger than 1:350,000, the area of interest should be shown without any CNDDB Element Occurrences. A table (see below) can be included stating "The following species are known to occur within this area."

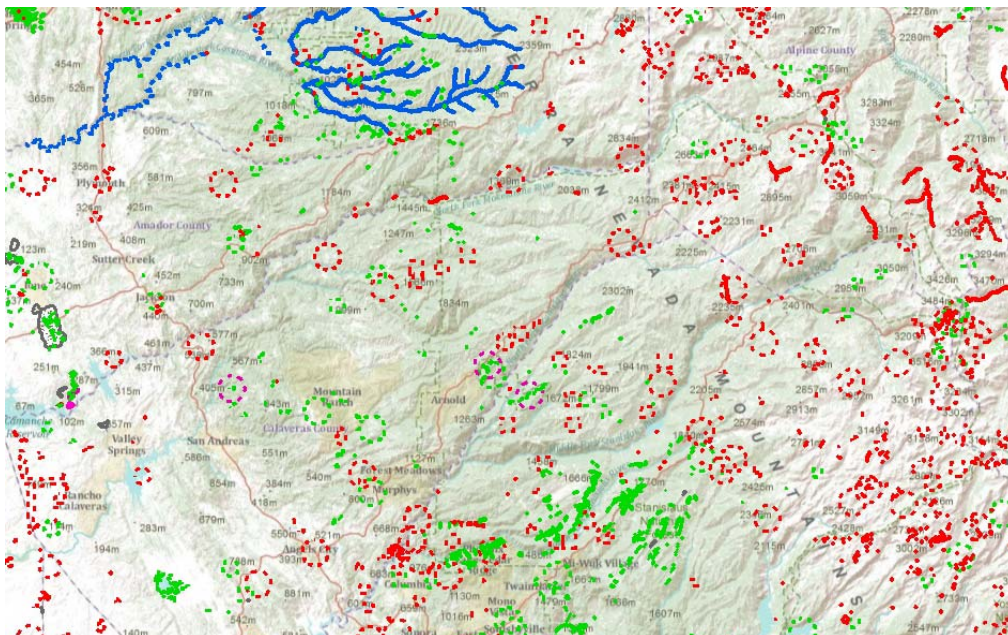| Scientific Name	Common Name
Rana boylii	foothill yellow-legged frog
Rana sierrae	Sierra Nevada yellow-legged frog
Pandion haliaetus	osprey
Haliaeetus leucocephalus	bald eagle
Accipiter gentiles	northern goshawk
Myotis yumanensis	Yuma myotis
Myotis evotis	long-eared myotis
Myotis thysanodes	fringed myotis
Myotis volans	long-legged myotis
Lasionycteris noctivagans	silver-haired bat
Lasiurus blossevillii	western red bat
Antrozous pallidus	pallid bat
Buxbaumia viridis	buxbaumia moss
Fissidens pauperculus	minute pocket moss
Peltigera hydrothyria	aquatic felt lichen
Sanicula tracyi	Tracy's sanicle
Packera eurycephala var. lewisrosei	Lewis Rose's ragwort
Packera layneae	Layne's ragwort
Wyethia reticulate	El Dorado County mule ears
Pyrrocoma lucida	sticky pyrrocoma
Lupinus dalesiae	Quincy lupine
Clarkia biloba ssp. brandegeeae	Brandegee's clarkia
Clarkia gracilis ssp. albicaulis	white-stemmed clarkia
Clarkia mosquinii	Mosquin's clarkia
Carex vulpinoidea	brown fox sedge
Rhynchospora capitellata	brownish beaked-rush
Fritillaria eastwoodiae	Butte County fritillary

For maps at a scale between 1:350,000 and 1:500,000: within this range one may display the CNDDB polygon layer information on a public map, but without labeling the individual features. Reference data should be very limited. Reference data can include county boundaries and water features (streams/lakes), and not much else. Much more in the way of reference features would again allow a viewer to fairly easily determine exactly where a species is located. Once again, one may provide a list of the species found within the map extent, but may not identify which graphic features belong to which species.

The following map is at a scale of 1:350,000. County lines and water features are shown along with the CNDDB polygon layer. No roads are displayed. A table like the one on page 10 could be added.
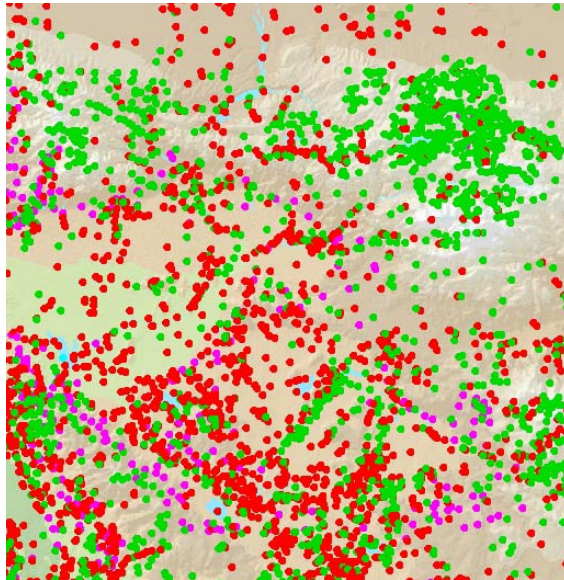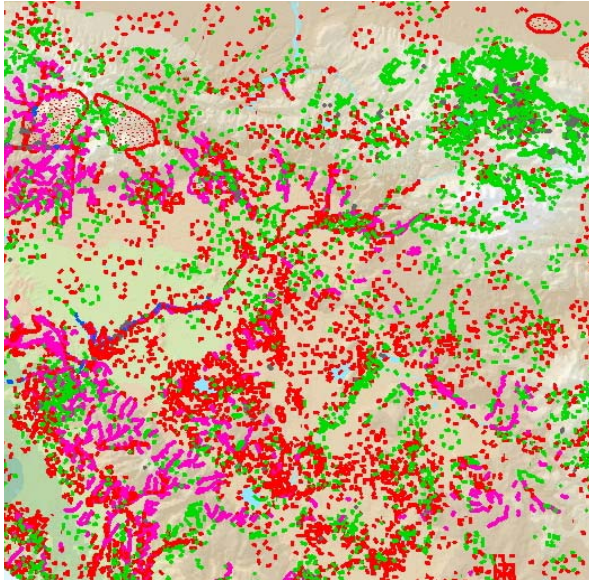

For maps at a scale of 1:500,000 and above: For maps at this scale and smaller (zoomed out), one may show CNDDB polygon layer data on a public map with whatever reference information is desired (roads, topography, imagery, etc).

The map below is at a scale of 1:500,000. Even though the base map contains a lot of detail, it is zoomed out enough to prevent pinpointing exact locations.

Consider switching to the point layer at scales smaller (more zoomed out) than 1:500,000 for better map clarity.

The maps below are at a scale of 1:750,000. The map on the left shows the polygon layer and the map on the right shows the point layer.

My project is only on one USGS quad, but I was told I should do a 9-quad search and look at all of the CNDDB information on my quad and the surrounding 8 quads. Why do I need to do this?

The CNDDB is a positive sighting database. It does not predict where something may be found. We map occurrences only where we have documentation that the species was found at the site. There are many areas of the state where no surveys have been conducted and therefore there is nothing on the map. That does not mean that there are no special status species present. By looking at what has been documented on your quad of interest and on the eight surrounding quads, you can estimate what might be found in similar habitats to those within your area of interest. The next step is to conduct surveys to document what is present and submit the information on special status species to the CNDDB.

9-quad searches can be easily done in desktop GIS, or by using the free CNDDB Quick Viewer (http://www.dfg.ca.gov/biogeodata/cnddb/mapsanddata.asp).

In the CNDDB Quick Viewer, navigate to your quad of interest and click on the "waffle" tool at

the top of the page with the CNDDB logo to see what Elements CNDDB has mapped for that quad and for the surrounding 8 quads.

To see unprocessed, unmapped Element information at CNDDB, click on the waffle tool with the

file cabinet logo . This tool and its companion single quad "unprocessed data search" icon

 provide the only access the public has to our "backlog" of data yet to be processed.