

# Section 3. Locality

The screenshot shows a web form titled "Locality" with several input fields and a "Tools" button. The fields are: Country, State/Province, County, Municipality (with an "IGNORE" button), Locality, Locality Security (checkbox), Latitude, Longitude, Uncertainty, Datum, Verbatim Coordinates, Elevation in Meters, and Verbatim Elevation. A "Tools" button is located between the Verbatim Coordinates and Elevation in Meters fields. Below the Tools button, a yellow box contains a section titled "Fields below revealed by clicking the tools button". This section includes: Lat: [ ] [ ] [ ] [ ] N [ ] W [ ], Long: [ ] [ ] [ ] [ ] W [ ], Zone: [ ], East: [ ], North: [ ], Hemisphere: North [ ], Insert Lat/Long Values, Insert UTM Values, and a section titled "IGNORE ALL THESE" with fields for T, N, R, E, Sec, and Details, along with a Meridian Selection dropdown and an Insert TRS Values button. Below the yellow box are fields for Georeferenced By, Georeference Sources, Georeference Remarks, Georeference Protocol, Georef Verification Status, and footprint (polygon), with "IGNORE" buttons for the last two.

Fig. 3.1. Basic Data Entry Form, Section 3 Expanded.

Providing the best possible information in this box is very important because it is used to link the specimen record to checklists and environmental layers.

**Country:** This is the first level geography field records in Symbiota. The field will autocomplete but, as with any autocomplete field, care is essential to make sure that the correct country name is added. Once a recognized country is entered in the country field, the choices offered for the stateprovince field are restricted to the applicable names. Entering "Pakistan" in the country field will ensure that "Utah" will not be offered for the StateProvince field.

**StateProvince:** This is the second geographic level in Symbiota. IN addition to the standard provinces, FATA is treated as a single province.

**County:** This is really just the third geographic level used. For Pakistan, it seems best to treat this as the DISTRICT. At present, there are no district names in the table. It can be changed by compiling a csv file showing the province in one column, the districts in the second column.

**Municipality:** This is a fourth geographic level. *Do not use it.* It results in more confusion than it is worth.

**Locality:** A verbal description of where the specimen was collected – information that someone with access to a map could locate, for example: x miles or kilometers northwest of a permanent landmark (an oasis, well-known cave, or the center of a nearby village/town/city). When making your own field notes, remember that it may be read by someone from another country or in another century. Do not use abbreviations and

**Locality Security:** If this is checked, only people with special permission will be able to see the locality, latitude and longitude fields. To get permission, ask the person in charge of the herbarium you are associated with. That person needs to ask me to give the necessary level of permission. The reason for this two step process is that I

do not know most of the people with accounts in OpenHerbarium.

One reason for checking the box would be that the location is sensitive, perhaps because it has lots of fascinating plants that are in danger of being eliminated by collectors or because the landowner does not want people to know what is present. To block access to detailed locality data about particular taxa requires creating a list of such taxa. Taxa can be sensitive in one state/province but not another.

## GEOREFERENCE INFORMATION

A georeference has three components: coordinates, uncertainty, and a datum. All three items need to be recorded. The most frequently used coordinates today are latitude and longitude. eds

**Latitude and Longitude:** The information should be in decimal degrees. If you are using a GPS, you can set it to provide data in this format.

Do not use more than 5 decimal places. In Pakistan,  $0.00001^\circ$  corresponds to a difference of 1.1 m of latitude, 0.9 m of longitude, more precise than most GPS units used by biologists. Towards the poles, the distance for a degree of longitude is much less than for a degree of latitude. To find out how different, go to <http://www.csgnetwork.com/degreenllavcalc.html>.

If your data are in degrees, minutes, and seconds, click the "Tools" button. This will reveal the fields shown in the box that has a pale yellow background (if you are looking at an electronic version of this document) or grey (if you are looking at a black and white print version of this document). You can enter the data as degrees and decimal minutes or degrees, minutes, and seconds and Symbiota will calculate the corresponding decimal number.

Latitudes north of the equator are positive; those south of the equator are negative. Longitudes for places east of the prime meridian are positive if they are east of the prime meridian, negative if they are to the west of it. The prime meridian used to be defined by a feature on the earth's surface, a particular location at the Royal Observatory in Greenwich, England. That changed with use of Global Positioning Systems (GPS). The position of satellites is determined by the earth's center of mass. For that reason, the prime meridian today is defined by reference to the earth's center of mass. It goes over the surface of the earth about 102 m east of the Greenwich meridian. All GPS units use the new meridian, as does the WGS84 datum.

There are two cautions when using the tool to calculate decimal latitude and longitude values:

1. You will have to change the directions from North and West to North and East.
2. You must click "Insert Lat/Long values" to have the program calculate and enter the decimal degree values. When you do, the data you entered will be transferred to "Verbatim Coordinates". "Verbatim" means "exactly what was entered".

For UTM coordinates, click the tools button and enter them into the second set of boxes in the shaded area. "UTM" stands for "Universal Transverse Mercator" and it is a system that is used on printed maps. For more information, check the article in Wikipedia. Note that you MUST include the zone as part of the information you provide. There is an interactive map at <http://whatutmzoneamiin.blogspot.com/p/map.html> for locating your zone. The Easting and Northing can be read from the map you are using or, if you have set your GPS to give readings in UTM, you can obtain them there.

When you click "Enter UTM Values", Symbiota will calculate the decimal latitude and longitude values and transfer the UTM data to the "Verbatim Coordinates Field".

**Uncertainty:** This is NOT the uncertainty that is recorded by your GPS unit or camera. It is an estimate, in meters of the radius of a circle around the point where you took the GPS reading within which you collected the plant. Even if you recorded the GPS data right by the plant, the uncertainty will be at least 55 m unless you have a high end GPS unit. If you took one reading for an area and collected several plants in the area, the uncertainty is the

distance from the GPS reading to the most distant plant.

The uncertainty reported by a GPS unit is its statement of how accurately it can interpret that satellite data that it is using. This can be affected by the terrain (uncertainty is higher in mountainous terrain) and whether the area is densely forested or not and various other factors. It will be lower than the uncertainty in your measurement.

The reason uncertainty is important is that, if someone wants to use herbarium records to estimate the ecological niche of a species, they need to know how accurate the geographic coordinates are. For instance, the Flora of Pakistan states where species occur within Pakistan in terms of a 2 degree grid so each record is just the center of a square 2 degrees by 2 degrees. This makes the uncertainty of 293,295m (yes, 293 km). Records from many other floras vary in their uncertainty because the regions used in reporting distribution differ in shape.

Maps with dots marking localities with differing levels of uncertainty can be very misleading. They may have many dots with low uncertainty in one area but a few, with high uncertainty, located far away. For this reason, Symbiota, by default, does not show dots for records with uncertainty greater than 1000 m. To see all dots, use the Spatial Module (an option on the Search tab).

**Datum:** A datum is a model of the earth that is used to when measuring latitude and longitude. Before satellites, countries used different datums, choosing one based on measurements in their own country or a nearby country. Reference points were defined through the country and marked by more or less permanent markers. The problem is that the lat/long values differ depending on the datum used. It can be compared to measuring the distance over a bumpy surface using rulers of differing flexibility. A very flexible ruler will enable you to fit the ruler closely to the surface and will give you a longer measurement than one that is so stiff you have to lay it on top of the bumps. As noted above, satellites have also led to a change in the location of the prime meridian.

By default, most GPS units use the World Geodetic System defined in 1984, which is usually abbreviated to WGS84, as the datum but they can be set to use other datums. It uses the “new” prime meridian. Google Earth (and web sites that use it or Google Maps use WGS84.

**Georeferenced By:** The name of the person who determines the georeference values (lat/lon, uncertainty, and datum). It is usually the collector but might be someone else on the collecting trip, or if working with existing specimens, someone else all together.

**Georeference Source:** Ideally, this is a GPS unit. If you do not have one, there are various web resources:

- <http://itouchmap.com/latlong.html>;
- <http://mynasadata.larc.nasa.gov/latitudelongitude-finder/>;
- <http://mygeoposition.com/>.

The first one uses satellite imagery which may make it easier to identify a location.

If georeferencing existing specimens with information on the label, use “Assumed to be collector”.

**Georeference Remarks:** If you had to move away from your plant to obtain a GPS reading, that information should go here (“Lat/lon from 50 m south of plant; trees prevented closer reading”). It is often left blank.

**Georeference Protocol:** Often left blank. It might be simply “GPS unit” or it might be miles along road as measured by the odometer in your car. It might be if you collected a plant near the road and had no GPS unit. Then you might have to record the distance to the nearest town from your car’s odometer. There are a set of recommended practices in a downloadable booklet at <http://herpnet.org/herpnet/documents/biogeomancerguide.pdf>.

**Elevation:** Elevation should be entered in meters. A topographic map or one of the web sites listed below will give you a better reading than most GPS units.

- <http://www.altitude.nu/> Provides best satellite image which helps in identifying location)
- <http://veloroutes.org/elevation/>
- <http://www.freemaptools.com/elevation-finder.htm>

- <http://www.whatismyelevation.com>

If you have just one elevation estimate, put it in the left hand box and leave the one to its right empty. If collections were made at and between two locations that were at different elevations, the highest elevation goes in the box to the right. If you have a measurement in feet, put it in “Verbatim Elevation” in the format “2024 ft” or “2024-3600 ft”. Symbiota will fill in the appropriate values in meters.