

ISMS Data Clean-Up Checklist

Categories of Errors and Prioritization

Highest Priority from 1 to 3 (see PRIORITY column in table below)

- 1) Do the data follow basic ISMS rules or violate data integrity rules? (see Att. 2 for the ISMS Standards & Guides)
- 2) Have all the required data elements been entered based on the taxa-specific Survey Protocols?
- 3) Are there data entry errors?

ISMS Tabular Data QA/QC

| CHECKLIST | EXPLANATION | SUGGESTION | PRIORITY |
|---|--|---|----------|
| 1. Are any UTM fields null? | If UTM_E and/or UTM_N are null, then it's likely a polygon has not yet been linked to the tabular record. | Use the ISMS.mdb - qryISMS Null and Invalid UTM Coordinates by Data Source Code query to find problem records. | 1 |
| 2. Are there any Location* records that have no species observations entered? | The reason for entering a record in ISMS is that a species was being looked for in the case of a General Survey or a species was found in the case of a Species Location. It is a data entry error if there are any Location* records lacking a species observation. | Use the ISMS web query tools – Quality Assurance and Checks – No Species Related to a Location query to find problem records. | 1 |
| 3. Are all the data fields required by the S&M Survey Protocol entered into ISMS? | ISMS only “requires” a small number of data fields. The taxa-specific Survey Protocols have other REQUIRED data. Check to make sure all the required data has been entered. Check for null (blank) ISMS fields. | Use the ISMS.mdb to create queries based on the specific survey protocols and look for null values in ISMS fields required by the survey protocol. – The ISMS Team will soon have additional “canned” queries for this topic. | 2 |

*Locations refers to either General Survey record or Species Location (aka known sites) record locations.

More TABULAR errors...

| CHECKLIST | EXPLANATION | SUGGESTION | PRIORITY |
|--|---|---|----------|
| 4. Has the collection verification information been entered? | Mollusks, lichens, bryophytes, and fungi have had required species verification since winter 2002. Each species of those taxa entered since then should have a collection record in ISMS with information about either the verifier or the verification date. | Use the ISMS web query tools – Species Collection Null Validation query to find problem records. | 2 |
| 5. Odd species entered? | Check the Species Observation table species codes, scientific names, or common names to be certain that valid species have been entered. | Use the ISMS.mdb to query species in the ISMS_SURVEY_OBS_VU. Requires familiarity with species known to be found in your area. | 3 |
| 6. Odd minimum / maximum values in numeric fields? | Is there a zero or other low or high number where there should not be? Was an extra digit accidentally added? | Use the ISMS.mdb to query for minimum and maximum values. Check into anything abnormal. | 3 |
| 7. Typographic errors in alphanumeric fields? | Check the ISMS fields that don't have pick lists such as LOC_ID, PROJECT, or SURVEY_ID and notes fields. Spell project names, creeks, etc., the same for each record; follow the standard for your admin unit, usually set by the ISMS data steward. | Use the ISMS.mdb to query for unique values in text fields. Oracle is case sensitive so be consistent with capitalization; it will allow easier querying later. | 3 |
| 8. Odd values selected from fields with pick lists? | Check for “unique” values entered into pick list fields. For example, was a bedrock type of “limestone” entered, but there is no limestone on your admin unit? | Use the ISMS.mdb to query for unique (i.e., distinct) values in pick lists. | 3 |
| 9. Missing certain values from fields with pick lists? | Check for “unique” values entered into pick list fields again. Look for missing items such as if there are only “pre-disturbance” SURVEY_TYPES when you know that “Purposive” survey data has been entered. | Use the ISMS.mdb to query the ISMS fields with pick lists based on your knowledge of what the data should be. If something seems incorrect, double-check the field data collection forms. | 3 |

ISMS Spatial Data QA/QC

| CHECKLIST | EXPLANATION | SUGGESTION | PRIORITY |
|---|---|---|----------|
| 1. Are the polygons in the correct geographic location? | Check the accuracy of the UTM's. For example, do all the shapes for your Data Source Code display within your administrative unit on land administered by your agency? | Use GIS. First look for obvious outliers. Then, look more closely to see if they are in the correct place (e.g., watershed, timber sale, township/range) in your administrative unit. | 1 |
| 2. Are any of the Location* records duplicates? | Did someone enter the exact same record twice? | Use the ISMS ArcView Query extension – Duplicate Record Finder tool to locate problem records. Data edits must be made in the ISMS database by deleting the duplicate record. | 1 |
| 3. Are there shapes that share the same LOC_CN or LOC_ID in the shapefile attribute table? | The ISMS Rule is that a tabular location* record in ISMS has only one polygon corresponding to one record in gISMS. Each location can have one or more surveys. Each location can have one or more species. But if you have more than one polygon for a location, then you'll need to determine which is the correct polygon and remove the extra(s). | Use the ISMS ArcView Query Extension – QA/QC – Duplicate Record tool to locate problem records. | 1 |
| 4. Are the distances between the Species Location (aka Known Sites) polygons correct for the specific taxa? | See the ISMS Standards and Guides for the distances for each taxa group. Some for example, say that occurrences within 300 feet (100m) of one another should be considered a single location in ISMS. | Use the ISMS ArcView Query Extension – Nearest Neighbor Analysis tool to locate problem records. | 1 |

*Locations refers to either General Survey record or Species Location (aka known sites) record locations.

More SPATIAL errors...

| CHECKLIST | EXPLANATION | SUGGESTION | PRIORITY |
|--|--|---|----------|
| 5. Was another Location record entered for a revisit survey when actually an additional Survey record should have been added to the original Location? | Once a General Survey or Species Location (aka known site) has been entered once, you can enter a revisit to the same survey area or known site as an additional “survey” record under the same parent location. (see ISMS Users Guide section 5.3.4). Adding more location records for each revisit is incorrect. | Data edits must be made in the ISMS database, but look for errors spatially. | 1 |
| 6. Are there null values in the gISMS attribute shapefile tables for LOC_ID or LOC_CN? | Null LOC_ID or LOC_CN in the gISMS shapefile attribute tables indicate that the polygon is not properly linked to the ISMS tabular data. | Use the ISMS ArcView Query Extension – QA/QC – Null LOC_CN/ID tool to locate problem records. | 1 |
| 7. Does a polygon have a LOC_ID and LOC_CN, but you cannot link or join it to any ISMS tabular record based on LOC_ID? | This indicates that a LOC_ID has been changed in the ISMS tabular record. Whenever that happens, the gISMS polygon must be re-linked to correct the LOC_ID in the gISMS shapefile attribute table. | Use the ISMS ArcView Query Extension – QA/QC – Mismatched LOC_Ids tool to locate problem records. | 1 |
| 8. Is there a General Survey associated with each Species Location? | Except for “Incidental” Species Locations (aka known sites), there should be a General Survey record in ISMS for each Species Location. Focus on current S&M species. | This can be checked most easily in ArcView by overlaying the Species Location polygon layer such as flora.shp on top of the General Survey polygon layer. | 2 |

Checking for coherent polygons and digitizing SPATIAL errors...

| CHECKLIST | EXPLANATION | SUGGESTION | PRIORITY |
|---|--|---|----------|
| 1. Do the polygons have slivers? | Slivered polygons are a sign of messy digitizing. They occur when you don't zoom in close enough in digitizing and accidentally create "spikes" that bisect the polygon boundary. | Use the ISMS ArcView Query Extension – QA/QC – Multi-part polygons and polygon slivers tool to locate problem records. | 1 |
| 2. Are there any "multi-part" polygons? | A "multi-part" polygon is two or more polygons that share the same tabular record in ArcView. Remember that one of the ISMS Rules is that there can only be one polygon per ISMS Location record. If you have multi-part polygons, it's likely that additional polygons should be treated as separate ISMS Location records. | Use the ISMS ArcView Query Extension – QA/QC – Multi-part polygons and polygon slivers tool to locate problem records. | 1 |
| 3. Are there overlapping or "island" polygons? | Overlapping and "island" polygons occur when one shape is completely contained within or intersected with a second shape. | The ISMS shapefile model allows these, but they can cause difficulty when querying so it's best to be aware of where they occur. Try to avoid creating these types of polygons. | 3 |
| 4. Are there any General Survey or Species Location polygons that are less than 9-m radius? | The ISMS Rule is that all polygons must be a minimum of 9-meters radius. | Use the ISMS ArcView Query Extension – QA/QC – Polygon size errors tool to locate problem records. (note: Low priority because this rule may change in ISMS version 2) | 3 |