

Inventory & Monitoring Division

National Park Service



Standard Operating Procedure 3.11.14

Quality Assurance and Quality Control for Landscape Dynamics Metrics: Population Density and Total Population

Please cite this as:

Kineth, M. and L. Nelson. 2018. Quality assurance and quality control for landscape dynamics metrics: Population density and total population. Inventory & Monitoring Program Standard Operating Procedure NPS/NRSS/IMD/SOP—3.11.14. National Park Service, Fort Collins, Colorado.

Summary

This internal Landscape Dynamics SOP provides guidance on QAQC for population density and population total statistics.

Contents

	Page
Summary	1
Overview.....	3
Requirements	3
QA Processing & Procedures.....	3
Step 1: Quality Assurance	4
Step 2: Combine Statistics.....	5
Step 3: Replace Statistics Data – All AOAs.....	6
Step 4: Replace Statistics Data – Select AOAs	7
Step 5: Manage Data Views	9
Literature Cited	10
Appendix 1: Data Availability and Known issues	11
Data availability.....	11
Water Area.....	11
Additional issues	11

Overview

This SOP provides instruction on how to QA\QC final summary statistics for population density and population total metrics, derived from US Census data (SOP 2.11.12; Kinseth and Nelson 2018e).

Statistics can represent various areas of analysis (AOA). At a minimum, AOAs for statistics include a 30km buffer park boundaries (SOP 2.11.6; Kinseth and Nelson 2018c). Other AOAs can include custom AOAs submitted by Networks or Cooperators for analysis such as island boundaries, watershed boundaries, or other ecological boundaries of interest (SOP 2.11.8; Kinseth and Nelson 2018b).

Population statistics are derived from the decadal U.S. Census data beginning with 1990. Statistics for population density and population total, by AOA, include area-weighted density of people within an AOA (total people per km²), and area-weighted total population within an AOA. (SOP 3.11.7; Kinseth and Nelson 2018a).

It is assumed users of this SOP have prior knowledge of data paths, file locations, and repositories referenced in this SOP.

Requirements

- Knowledge and ability to use SSMS (SQL Server Management Studio)
- Read/write database security privileges for StatsQA and LD_Stage
- Run population density and total population summary statistics (SOP 3.11.7; Kinseth and Nelson 2018a).

QA Processing & Procedures

Steps 1 and 2 can be accomplished by any IMD-GIS staff. Steps 3 to 5 can only be accomplished by select IMD-GIS staff with the proper, elevated permissions within the database.

Before you start:

- Connect to SSMS and view the tables within the StatsQA database. There may be multiple population statistics tables, by Census year.
- Repeat steps 1, 2 for each population density/total table present in StatsQA.

Naming convention for statistics tables:

PDDPTT<YYYY>_<AOA Source>_<AOA type>_stats

Where:

YYYY = Census Data Year

AOA Source = CONUS, ALASKA, HAWAII, USVI, GUAM, SAIPAN, AMSAM, Other from Custom AOAs

AOA type = 30km, Other from Custom AOA

Step 1: Quality Assurance

1. View table and look for any NULL fields in DATE_CALC. If there are any, stop. Statistics may have to be rerun. If there are no NULLs, continue with QA.
2. Check all AOAs are represented in the table. The number of rows should match the number of features from the input AOA feature class list (if there is a discrepancy, stop immediately).

```
SELECT Count(  
    [DATA_YEAR] )  
    , [AOA_NAME]  
FROM [StatsQA].[dbo].[STATS TABLE NAME GOES HERE]  
group by [AOA_NAME]
```

There should be a Count of **1** for each row.

3. Check for AOAs with '0' population total ([SUM_AWP_F_AOA]) or '0' population density ([AWPD_AOA]). If any AOAs show '0' for these fields, check the spatial extent of the AOA. Only AOAs that are entirely protected (as defined in PADUS; i.e. parks/AOAs that encompass an entire island like DRTO for example) should show this result:

```
SELECT [OBJECTID]  
    , [FREQUENCY]  
    ,etc. remaining fields  
FROM [StatsQA].[dbo].[STATS TABLE NAME GOES HERE]  
order by AWPD_AOA
```

4. Check the density value (AWPD_AOA) against (AWPD_CHECK). AWPD_DIFF should fall between -0.005 and 0.005.

```
SELECT [AOA_NAME], ([SUM_AWP_F_AOA]/[SUM_CCBAF_AOA]) as [AWPD_CHECK],  
[AWPD_AOA], ([SUM_AWP_F_AOA]/[SUM_CCBAF_AOA])-[AWPD_AOA] as [AWPD_DIFF]  
From [StatsQA].[dbo].[STATS TABLE NAME GOES HERE]  
where [FREQUENCY] != 0  
order by [AWPD_DIFF]
```

Step 2: Combine Statistics

Where applicable, combine data with matching year/AOA to create one table of population density and total.

Go to Step 3 if there are no tables to combine.

1. Create a new table for combined Census statistics (by Census year, AOA type). Add the following to the select statement, on the line above 'From':

```
INTO [StatsQA].[dbo].[NEW/DESTINATION TABLE HERE]
```

Naming convention for new tables should match table names in LD_Stage, for the same metric (if any):

Example New Table = [StatsQA].[dbo].[pddptt1990_stats_30km]

2. Refresh StatsQA and look for the new table.
3. Insert additional tables to the new table as needed:

```
INSERT INTO [StatsQA].[dbo].[NEW TABLE]
```

```
( [FREQUENCY]  
  , [SUM_AWP_F_AOA]  
  , [SUM_CCBAF_AOA]  
  , [SUM_CBA_W_AOA]  
  , etc. for remaining fields)
```

```
SELECT [FREQUENCY]  
  , [SUM_AWP_F_AOA]  
  , [SUM_CCBAF_AOA]  
  , [SUM_CBA_W_AOA]  
  , etc. for remaining fields
```

```
FROM [StatsQA].[dbo].[ADDITIONAL SOURCE TABLE(S)]
```

4. When done, check combined table.

```
SELECT Count(  
  [DATA_YEAR])  
  , [AOA_NAME]  
FROM [StatsQA].[dbo].[STATS TABLE NAME GOES HERE]  
group by [AOA_NAME]
```

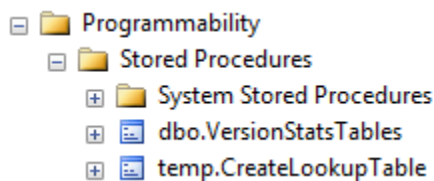
The number of rows should equal the total of all AOAs from the individual tables that were merged.

Step 3: Replace Statistics Data – All AOAs

Skip to Step 4 if updating/adding a partial list of AOA statistics

Move the final tables from StatsQA to the LD_Stage database when all QA checks are done and all data values are as they should be (requires elevated privileges). If adding new YYYY data/statistics, go to #2.

1. Version previous population tables (if necessary):
 - a. In LD_Stage expand Programmability → Stored Procedures:

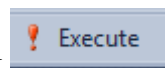


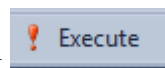
- b. Right click on dbo.VersionStatsTables → “Script Stored Procedure As” → EXECUTE TO → New Query Window.
 - c. Replace @fromTable with the name of the table to be versioned, in single quotes (example below in red):

```
DECLARE @RC int
DECLARE @fromTable varchar(60)

-- TODO: Set parameter values here.

EXECUTE @RC = [dbo].[VersionStatsTables]
    'pddptt2010_30km_stats'
GO
```



- d. Click  to date-stamp existing table
 - e. Refresh to see table with date stamp appended:

 zArchive.pddptt1990_30km_stats_2015-11-25

2. Copy the new statistics table into LD_Stage, repeat for all new tables.
 - a. Right click on table → Script Table As → Select To → New Query Editor Window
 - b. Run:

```

SELECT **all table fields here **
INTO [LD_Stage].[dbo].[STATS TABLE NAME GOES HERE]
FROM [StatsQA].[dbo].[STATS TABLE NAME GOES HERE]

```

LD_Stage tables should read format:

[MetricYYYY_stats_AOA] → [pddptt2010_stats_30km]

(see existing tables in LD_Stage for additional examples)

- c. Refresh LD_Stage and look for the table. It should be adjacent to the date-stamped table (if any).
3. In LD_Stage, change schema of date-stamped table(s) from dbo to zArchive:

```
ALTER SCHEMA zArchive TRANSFER dbo.[DATE STAMP TABLE NAME HERE] ;
```

- a. Refresh LD_Stage and check all date-stamp tables are moved from dbo to zArchive:

Tables, example, zArchive:

```

+ zArchive.pddptt2010_30km_stats_2015-11-25
+ zArchive.pddptt2010_30km_stats_2017-01-19

```

4. Delete the moved tables from StatsQA.

5. **Go to Step 5 – Manage Data Views**

Step 4: Replace Statistics Data – Select AOAs

Move stats table of interest from LD_Stage to StatsQA and replace data for AOAs that have been updated, or add data for new AOAs to the table.

1. To append updated AOA statistics, or to add new AOA statistics to existing statistics tables:

- a. COPY the table to be updated FROM LD_Stage TO StatsQA:
Right click LD_Stage table → Script Table As → Select To → New Query Window

```

INTO [StatsQA].[dbo].[STATS TABLE NAME_copy]
FROM [LD_Stage].[dbo].[STATS TABLE NAME]

```

- b. Get initial count of AOAs:

```

SELECT Count(
    [DATA_YEAR])
    , [AOA_NAME]
FROM [StatsQA].[dbo].[STATS TABLE NAME_copy]

```

group by [AOA_NAME]

2. Delete from the “_copy” table all AOAs/rows to be updated/replaced

a. Right click → Edit Top 200 Rows → click white sql button and edit query as needed

i. Select and delete AOAs/rows that are updated.

Edits are immediate and cannot be undone – make sure you are editing the appropriate table.

b. Or, edit sql statement as needed. Run.

```
DELETE FROM [StatsQA].[dbo].[STATS TABLE NAME_copy]
```

```
where AOA_NAME like 'APIS%'
```

```
OR AOA_NAME like 'AZRU%'
```

```
OR AOA_NAME like 'BIBE%'
```

```
OR AOA_NAME like 'BITH%'
```

```
OR AOA_NAME like 'BLCA%'
```

```
OR AOA_NAME like 'CARL%'
```

```
,...add/edit as needed
```

i. Deleted rows cannot be undone. Make sure to delete from the appropriate table.

3. Query for count on edited table:

```
SELECT Count(  
    [DATA_YEAR])  
    , [AOA_NAME]  
  
FROM [StatsQA].[dbo].[STATS TABLE NAME_copy]  
  
group by [AOA_NAME]
```

4. Insert new/updated AOA statistics from the newly created table into the ‘_copy’ of the table (make sure to match fields based on table being updated; exclude field ‘OBJECTID’):

```
INSERT INTO [StatsQA].[dbo].[STATS TABLE NAME_copy]  
  
    ([FIELD LIST HERE]  
  
SELECT  
  
    FIELD LIST HERE  
  
FROM [StatsQA].[dbo].[NEW TABLE HERE]
```

5. Check Count. The number of rows should have increased by the number of AOAs added in the previous step and should now reflect required total for AOA type.

```
SELECT Count(  
    [DATA_YEAR])  
    , [AOA_NAME]  
FROM [StatsQA].[dbo].[STATS TABLE NAME_copy]  
group by [AOA_NAME]
```

6. Visually check table for the applied updates.
 - a. Right click → “Select Top 1000 rows”
 - i. Order by AOA_Name – look for updated AOA data
 - ii. Order by OBJECTID
 - iii. Other types of queries
7. **Go to STEP 3 – Replace Statistics Data – ALL AOAs** and follow steps to replace/update table in LD_Stage

Step 5: Manage Data Views

If required, go to Inventory and monitoring database documentation: Environmental Setting Protocol SOP (SOP 7.1.1; Kinseth et al. 2018) and check the Landscape Dynamics portion.

Literature Cited

- Kineth, M. and L. Nelson. 2018a. Calculating statistics for national landscape dynamics metrics: Population density and total population. Inventory & Monitoring Program Standard Operating Procedure NPS/NRSS/IMD/SOP—3.11.7. National Park Service, Fort Collins, Colorado.
- Kineth, M. and L. Nelson. 2018b. Creating custom areas of analysis for use as a part of the IMD Environmental Settings Protocol. Inventory & Monitoring Program Standard Operating Procedure NPS/NRSS/IMD/SOP—2.11.8. National Park Service, Fort Collins, Colorado.
- Kineth, M. and L. Nelson. 2018c. NPS unit boundary processing: Boundary-derived areas of analysis for use as a Part of the IMD Environmental Settings Protocol Inventory & Monitoring Division Standard Operating Procedure NPS/IMD/SOP—2.11.6. National Park Service, Fort Collins, Colorado.
- Kineth, M. and L. Nelson. 2018d. Processing land cover source data to derive landscape dynamics metrics: Protected area data. Inventory & Monitoring Program Standard Operating Procedure NPS/NRSS/IMD/SOP—2.11.19. National Park Service, Fort Collins, Colorado.
- Kineth, M. and L. Nelson. 2018e. Processing source data to derive landscape dynamics metrics: Population source data. Inventory & Monitoring Program Standard Operating Procedure NPS/NRSS/IMD/SOP—2.11.12. National Park Service, Fort Collins, Colorado.
- Kineth, M., L. Nelson, and S. Schupbach. 2018. Inventory and monitoring database documentation: Environmental Setting Protocol. Inventory & Monitoring Division Standard Operating Procedure NPS/NRSS/IMD/SOP—7.1.1. National Park Service, Fort Collins, Colorado.

Appendix 1: Data Availability and Known issues

Data availability

United States Census Bureau data is released every 10 years. Population Total and Density summary statistics will only change between source data releases if the AOA changes. These updates for new or altered AOAs are calculated annually. Please note the DATE_UPDATED field in the statistics services – this date indicates when the posted statistics were last updated/calculated.

Water Area

Population total and population density summary statistics are area weighted calculations based on the sum of amount of area, by census block group, remaining within an AOA after protected areas have been erased and census reported water area has been subtracted.

Because census reported water area is only subtracted from the total area for a census block group, and not spatially represented by an erase, for some census block groups this can result in a negative area. For census block groups where this occurs, water area is added back to the total area and calculations for population total and population density are based on AOA area with ONLY protected areas erased.

All population totals and population densities should be evaluated as generalized values, with the above scenario in mind.

Protected Areas Erase

Protected Areas designated GAP Status 1, 2 (PADUS) are erased from census block group areas within an AOA. These areas are, by designation, protected from development and therefore are excluded from population density calculations. Exclusion of these areas provides a better estimated population density for an AOA by removing area where people cannot legally reside. No manipulation or changes to the source protected area data (PADUS) occurs, and the erase depends explicitly on the GAP Status attribute. See Protected Area Processing SOP (SOP 2.11.19; Kinseth and Nelson 2018d) for additional details about PADUS data and the known issues related to the database and its associated data.

Census Block Group Area change over time

Census Block Groups can (and do) change with each Census. Bear this in mind if conducting a comparison across years. Statistics calculations are done in such a way as to normalize the data, but there could be inherited bias from a change in census block group polygon size and shape, decade by decade.

Additional issues

PARK_CODE is valid for AOAs derived specifically from individual parks. Some Custom AOAs are not park-based and therefore will not have a PARK_CODE value.