



Minnesota Next Generation 9-1-1 Validation Overview

Overview

Minnesota's transition to Next Generation 9-1-1 (NG9-1-1) requires designated GIS data providers to produce accurate, standardized data for input into a statewide geospatial data repository. In a collaborative effort to assist data providers in meeting these requirements, the Emergency Communication Networks (ECN) division of the Department of Public Safety and the Minnesota Geospatial Information Office (MnGeo) have created a validation process to identify issues in these data.

This document clarifies the data validation process phases by providing a comprehensive list of the data checks conducted and results reported back for review. There are two types of returned results: (1) those that pertain to data requirements and therefore require error resolution, and (2) those that relate to data management best practices and error resolution is recommended but not required. These will be clearly distinguished and supplemental validations in each phase will only be made available once minimum requirements for that phase are met. Lastly, validations related to data management best practices will not be counted against a PSAP's level of completeness for any given validation phase.

Exceptions to validations will be considered and tracked where appropriate. Please contact the regional NG9-1-1 GIS representative regarding unresolvable exceptions. Data providers should recursively submit data for validation, correct returned errors, and re-submit the data for it to be re-run through the MnGeo validation process. Data will only be considered passing validation if it has been validated by MnGeo and meets a 100% compliance percentage, excluding allowable exceptions.

Please note, the current validation phases detailed in this document is not comprehensive to being ready for NG9-1-1 call routing. The phase that covers validations relating to the required polygon files and data maintenance workflows is currently being finalized. It is critical to note that data cannot be assessed for this final preparedness phase until the other validations have been remediated. This document is anticipated to be released prior to the end of the year. In addition, MnGeo will be deploying additional supplementary validations which will be detailed in this document as they become available—these supplementary validations serve to improve data quality but do not pertain to data requirements.

Data Pre-Processing

ECN directs PSAPs and GIS personnel to understand that there is currently not a one to one relationship between the MN GAC data standards and the [NENA NG9-1-1 Data Model](#). The last section of this document provides the cross-reference which establishes the correlation between the data field names and schema of the GAC RCL and ADP Standards and the NENA NG9-1-1 Data Model. ECN requests:

- 1) Only data elements that have a “crosswalk” relationship to the NENA NG9-1-1 Data Model from the MN GAC standard need to be propagated. These are identifiable by the presence of a NENA field name and NENA inclusion designation for that GAC data element.
 - a. Use the MN GAC field name, type and field width in these instances.
 - b. *Examples:* ADP: County Name, RCL: Left From Address
- 2) Data elements which are in the NENA standard but are not in the MN GAC standard can be propagated. These elements are listed toward the bottom of the crosswalk table below and are identifiable by the presence of NENA field names only – no GAC field name is supplied.
 - a. *Examples:* ADP: Neighborhood Community, RCL: Legacy Street Name
- 3) Data elements which are found in the GAC schema only and do not crosswalk to the NENA requirements do not have a NENA field name or inclusion designation. These data elements are ineligible for reimbursement under the federal grant.
 - a. *Examples:* ADP: CTU Code, RCL: Route ID

The following data requirements should be met *prior* to starting the validation process:

- 1) All fields mapped to the Minnesota NG9-1-1 schema must have the same field type as detailed in the corresponding GAC standard, and the field width must not exceed the GAC-defined field width (use NENA field type and width if missing from the GAC standard).
- 2) All mandatory elements must be present and contain non-null values. If a data element is conditional and the attribute value exists (even if not yet captured in the data), it must also be included and contain non-null values.
- 3) Fields with a standard domain must follow the GAC domain *or* agreed-upon values for MnGeo to translate to the domain values.
- 4) Fully parse out all street name elements in both the ADP and RCL datasets and propagated into the related attribute fields.
- 5) Fully spell out all abbreviated street name elements in both the ADP and RCL datasets.

Data providers will be required to verify field-mapping of the data from local format to the Minnesota NG9-1-1 data schema. Data providers will also be required to verify domain translation tables if abbreviated data elements will be fully spelled out by MnGeo.

Validation Process

There are seven major validation phases:

- 1) Community Name Validation
- 2) Emergency Service Number Validation
- 3) Street Name Validation
- 4) Master Street Address Guide Validation
- 5) Address Validation
- 6) Database Integrity Validation (formerly called *Geospatial Validation*)
- 7) Edge-matching and Polygon-to-Location Data Validation (*not yet published; Est 1/2/20*)

Input Data

The following datasets are the minimum datasets required for currently published validations:

- 1) Address Points (ADP) – *geospatial point feature class*
- 2) Road Centerlines (RCL) – *geospatial linear feature class*
- 3) Master Street Address Guide (MSAG) – *standalone table*
- 4) Automatic Location Identification (ALI) – *standalone table*
- 5) Emergency Service Number (ESN) Lookup Table – *standalone table*
- 6) [Cities, Townships and Unorganized Territories \(CTU\)](#) – *standalone table*
- 7) [United States Postal Service \(USPS\) Preferred Postal Community Names](#) – *standalone table*

Major Validation Statistics and Results

The following is a list of the major statistics that come out of the validation phases. *For more detailed and comprehensive list of report statistics, proceed to the ‘Validation Tests’ section.*

- 1) Percentage of matching community names
- 2) Percentage of matching Emergency Service Numbers
- 3) Percentage of matching street names
- 4) Percentage of matching RCL to MSAG records
- 5) Percentage of RCL with no address range errors
- 6) Percentage of ALI records that geocode to GIS location data (ADP or RCL)
- 7) Percentage of ADP that geocode to the RCL
- 8) Percentage of RCL with no database integrity errors
- 9) Percentage of ADP with no database integrity errors

In addition to validation statistics, MnGeo will provide a table of individual geospatial features (e.g. RCL) or list of values (e.g. MSAG community names) that failed validation. These will serve as a guide in error resolution. Notice that each of the returned features will have a unique ID—if the feature was already assigned a unique ID locally, this will be used, otherwise MnGeo will assign a unique ID which (if desired) may be adopted locally.

Validation Tests

Validation is conducted on attributes and geometry via a series of validation checks pertaining to (and relating) each feature class (e.g. addresses). Below is a summary of each validation phase.

Community Name Validation

The purpose of this validation phase is to normalize community names across various disparate datasets. The MSAG and CTU community names should be synchronized across the MSAG, RCL, ADP, and CTU domain list, accordingly. Optionally, the postal community names should be synchronized across the RCL, ADP, and USPS preferred postal community name list. Verifying whether individual ADP and RCL segments have the correct MSAG community name will be conducted in later validation phases.

Data Requirements

- MSAG community names synchronized between the MSAG, RCL and ADP
 - a. Each MSAG community name in the RCL must be present in the MSAG
 - b. Each MSAG community name in the ADP must be present in the MSAG
- CTU names synchronized between the state CTU domain, RCL and ADP
 - o Each CTU name in the RCL must be in the state CTU domain, except for unorganized territories (UT)
 - o Each CTU name in the ADP must be in the state CTU domain, except for UTs

The following results will be provided. Errors highlighted in these results relate to data requirements and must be resolved to pass validation.

- 1) Percent compliant for the following:
 - a. RCL MSAG community names in the MSAG
 - b. ADP MSAG community names in the MSAG
 - c. RCL CTU names in the CTU domain
 - d. ADP CTU names in the CTU domain
 - e. Average total match rate of *a-d* above
- 2) Table of MSAG community names in the RCL, with 'PASS'/'FAIL' indicated in the following fields:
 - a. 'msag_presence' – indicates whether MSAG community name is in the MSAG
- 3) Table of MSAG community names in the ADP, with 'PASS'/'FAIL' indicated in the following fields:
 - a. 'msag_presence' – indicates whether MSAG community name is in the MSAG
- 4) Table of CTU names in the RCL, with 'PASS'/'FAIL' indicated in the following fields:
 - a. 'ctu_domain' – indicates whether CTU name is in the state CTU domain
- 5) Table of CTU names in the ADP, with 'PASS'/'FAIL' indicated in the following fields:
 - a. 'ctu_domain' – indicates whether CTU name is in the state CTU domain

Data Recommendations & Supplemental Validations

- Follow the state-recommended UT naming convention:

'UT [LOCALLY USED NAME]' (e.g. 'UT Shovel Lake', 'UT T.157-R.30')
- As per recommended in the state standards, use the preferred postal community “whenever possible”; ensure synchronization between the RCL and ADP
 - o Each postal community name in the RCL is in the USPS Preferred Postal Community Names list for a specified county
 - o Each postal community name in the ADP is in the USPS Preferred Postal Community Names list for a specified county
 - o Each postal community name in the ADP is present in the RCL

The following results will also be provided. These results do not pertain to data requirements and do not affect the passing of this validation phase, thus records failing these tests will receive a 'WARNING' instead of 'FAIL'. However, addressing these results may improve data quality and/or help catch errors that may otherwise be missed.

- 1) Percent compliant for the following:
 - a. RCL postal community names in the USPS preferred postal community names list
 - b. ADP postal community names in the USPS preferred postal community names list
- 2) Table of postal community names in the RCL, with 'PASS'/'FAIL' indicated in the following fields:
 - b. 'postal_domain' – indicates whether postal community name is in the USPS preferred list
- 3) Table of postal community names in the ADP, with 'PASS'/'FAIL' indicated in the following fields:
 - c. 'postal_domain' – indicates whether postal community name is in the USPS preferred list

Emergency Service Number Validation

The purpose of this validation phase is to normalize emergency service numbers (ESNs) across various disparate datasets. ESNs should be synchronized across the MSAG, RCL, and ADP. Verifying whether individual ADPs and RCLs have the correct ESN will be conducted in later validation phases.

Data Requirements

- ESN synchronized between the MSAG, RCL and ADP
 - Each ESN in the RCL must be present in the MSAG
 - Each ESN in the ADP must be present in the MSAG

The following results will be provided. Flagged failures in these results relate to data requirements and must be resolved to pass validation.

- 1) Percent compliant for the following:
 - a. RCL ESN in the MSAG
 - b. ADP ESN in the MSAG
 - c. Average total match rate of *a-b* above
- 2) Table of ESNs in the RCL, with 'PASS'/'FAIL' indicated in the following fields:
 - a. 'msag_presence' – indicates whether RCL ESN is in the MSAG
- 3) Table of ESNs in the ADP, with 'PASS'/'FAIL' indicated in the following fields:
 - a. 'msag_presence' – indicates whether ADP ESN is in the MSAG

Street Name Validation

The purpose of this validation phase is to normalize street name elements across various disparate datasets. Street names should be synchronized across the MSAG, RCL, and ADP.

A street name is identified by concatenating the street name elements and MSAG community name fields. Recognizing there may be RCL segments without addresses, street names in the RCL that are neither in the ADP or the MSAG is allowed; however, these discrepancies are reported in case the data provider would like to add these to the MSAG, or in the rare event that both the MSAG and ADP are missing a street. Verifying the synchronization and accuracy of the combined street name, address range, parity, MSAG community and ESN will be conducted in later validation phases.

Data Requirements

- Street name synchronized between the MSAG, RCL and ADP
 - Each street in the MSAG must be present in the RCL

The following results will be provided. Flagged failures in these results relate to data requirements and must be resolved to pass validation.

- 1) Percent compliant for the following:
 - a. MSAG street names in the RCL
 - b. Average total percentage compliant for *a-b* above
- 2) Table of MSAG street names, with 'PASS'/'FAIL' indicated in the following fields:
 - a. 'rcl_presence' – indicates whether MSAG street name is in the RCL

MSAG Validation

The purpose of this validation phase is compare the RCL and MSAG datasets with a goal towards the RCL being able to provide the same answers the MSAG currently does. Each MSAG record consists of a street name that is assigned an MSAG community name, address range, ESN, and parity (a.k.a. EOB –“*Even, Odd, Both*”). MSAG records may capture theoretical address ranges, while RCL might represent actual address ranges, thus not all ranges in the MSAG are required to be in the RCL. Conversely, all ranges in the RCL *are* required to be in the MSAG. **NOTE: Parity in the MSAG indicates whether the ESN applies to even addresses, odd addresses or both even AND odd addresses. Parity indicated in the RCL is not considered in this validation since the RCL parity pertains to parity of addresses on either side of the RCL segment, while the parity in the MSAG pertains to parity as it relates to ESN.*

Data Requirements

- Each combination of street name, MSAG community, address range, and ESN in the RCL is represented in the MSAG.

The following results will be provided. Errors highlighted in these results relate to data requirements and must be resolved to pass validation.

- 1) RCL correspondence to MSAG
 - a. Percent match-rate of RCL to MSAG
 - b. Table of RCL features that are not represented in the MSAG (i.e. combination of street name, MSAG community, address range and ESN in the RCL is not captured in the MSAG)

Data Recommendations & Supplemental Validations

The following results will also be provided. These results do not pertain to data requirements and do not affect the passing of this validation phase. However, addressing these results may improve data quality and/or help catch errors that may otherwise be missed.

- 1) MSAG correspondence to RCL
 - a. Percent match-rate of MSAG to RCL
 - b. Table of MSAG records that are not represented in the RCL

Address Validation

The purpose of this validation phase is threefold: (1) ensure that address ranges within the RCL segments do not conflict with one another, (2) ensure all valid ALI civic locations can be geoverified, and (3) achieve addressing consistency between the ADP and RCL. Addresses consist of these data elements: address number, street name, MSAG community name, and ESN. All valid ALI records must be geocodable to a single location in either the ADP or RCL. Each address in the ADP must be geocodable to a single location on the RCL.

Data Requirements

- 1) RCL
 - a. *NULL address range* – features must have numeric left and right FROM/TO address ranges
 - b. *Parity fields* – parity left and right must match the corresponding left and right parity in the address range fields, except for instances when parity is ‘BOTH’
 - c. *Overlapping address ranges* – features with the same full street name, MSAG community name, and ESN must not have overlapping address ranges
- 2) ALI geoverification to GIS location data (ADP or RCL) – 100% match*
- 3) ADP geoverification to RCL – 100% match

**Not all features in the ALI database can be geoverified. Please contact your PSAP representative and/or the 9-1-1 service provider to verify these exceptions. Once identified please communicate them to the regional NG9-1-1 GIS representative.*

The following results will be provided. Flagged errors (including ‘FAIL’s and ‘TIE’s) relate to data requirements and must be resolved. Allowable exceptions must be negotiated with and marked by the regional NG9-1-1 GIS representative.

Address Ranges

- 1) Percentage of RCL without address range errors
- 2) Table of RCL features (‘results_rcl’), with a field indicating ‘PASS’/‘FAIL’ for the following:
 - a. ‘null_adr_range’ – indicates whether there is a NULL value in the address ranges
 - b. ‘parity_fields’ – indicates whether the parity indicated in the parity fields matches the parity of the numbers in the address range fields; ‘BOTH’ is considered an exception
 - c. ‘overlapping_range’ – indicates whether the address ranges in the RCL overlap with the address ranges of another RCL along the same street (using the same full street name, MSAG community name, and ESN as the street identifier); *NOTE: please also see result item 3 below, which contains single features identifying each RCL pair with overlapping ranges*
- 3) Polyline feature class (‘overlapping_errors’) containing features that identify two RCL segments with overlapping address ranges—each feature is a merge of two RCL segments from the original RCL dataset that have overlapping address ranges; this feature class serves as a guide to identify

the original RCL segments that contain overlapping ranges and is a complement to the 'overlapping_range' validation field above.

ALI to GIS Location Data

- 1) Percent match-rates for the following comparisons:
 - a. ALI to GIS location data (RCL *or* ADP)
- 2) Table of ALI records, with a 'ali_to_gis' field indicating the following:
 - a. 'PASS' – indicates ALI address geocoded to a single ADP or RCL
 - b. 'TIE' – indicates ALI address geocoded to multiple ADPs or RCLs (*this is not allowed*)
 - c. 'FAIL' – indicates ALI address was not found in the ADP or RCL

ADP to RCL

- 1) Percent match-rates for the following comparisons:
 - a. ADP to RCL
- 2) Table of ADP features, with a 'adp_to_rcl' field indicating the following:
 - a. 'PASS' – indicates ADP address geocoded to a single location on the RCL
 - b. 'FAIL' – indicates ADP address was not found in the RCL

Data Recommendations & Supplemental Validations

- Review ALI records that geocode to the RCL but not the ADP, as these may indicate a possible missing address point
- Flipped address ranges – address range FROM should not be larger than TO since RCLs are to be digitized in the direction of increasing addresses

The following results will also be provided. These results do not pertain to data requirements and do not affect the passing of this validation phase. However, addressing these results may improve data quality and/or help catch errors that may otherwise be missed.

- 1) Table of ALI records, 'ali_to_rcl_only' – these are ALI records that geocode to the RCL, but not the ADP, indicating a possible missing address point
- 2) RCL Supplemental Checks – the following results will be flagged for review but do not necessarily indicate true errors:
 - a. 'flip_ranges' – indicates whether the FROM/TO address ranges seem to be opposite of intent (i.e. flipped; FROM address is higher than the TO address)

Database Integrity Validation

The purpose of this validation phase is to ensure data and database integrity. Geometry of features will also be assessed. Requirement 1a below pertains to unique IDs, which (as stated above) will be populated by MnGeo if they are missing. Exceptions to requirements 2a and 2b below may be negotiated with and marked by the regional NG9-1-1 GIS representative.

Data Requirements

- 1) RCL and ADP (*the following three tests will be conducted on PSAP, ESB and provisioning boundaries in the final validation phase*)
 - a. *Duplicate IDs* – features must have a unique ID unlike any other feature’s unique ID
 - b. *Duplicate records* – features must not have identical attribute information (this considers all NENA attributes except unique ID)
 - c. *Empty geometry* – all records must have geometry

- 2) RCL only
 - a. *Street intersection splits* – RCLs must be split at intersections
 - b. *Connections and dangles* – RCLs must be snapped to other RCLs, leaving no dangle nodes; exceptions may include dead ends, roundabouts, and island roads

The following results will be provided. Flagged errors in these results relate to data requirements and must be resolved to pass validation. Allowable exceptions must be negotiated with and marked by the regional NG9-1-1 GIS representative.

- 1) Percentage of features with no database integrity validation errors for each of the following (must be 100% error-free to pass, excluding exceptions):
 - a. RCL
 - b. ADP

- 2) Table of RCL features (‘results_rcl’), with a field indicating ‘PASS’/‘FAIL’ for the following:
 - a. ‘duplicate_ids’ – indicates whether unique ID is identical to another feature’s unique ID
 - b. ‘duplicate_records’ – indicates whether the attributes (discluding unique ID) are identical to another feature’s attributes
 - c. ‘empty_geometry’ – indicates whether a record is missing geometry

- 3) Table of ADP features (‘results_adp’), with a field indicating ‘PASS’/‘FAIL’ for the following:
 - a. ‘duplicate_ids’ – indicates whether unique ID is identical to another feature’s unique ID
 - b. ‘duplicate_records’ – indicates whether the attributes (discluding unique ID) are identical to another feature’s attributes
 - c. ‘empty_geometry’ – indicates whether a record is missing geometry

- 4) Point feature class (‘unsplit_intersections’) containing points where RCL segments are not split at intersections with other RCL segments and/or a RCL contains a dangle node.

Edge-matching and Polygon-to-Location Data Validation

The purpose of this validation phase is to assess polygon topology, attribution, and edge matching in the Emergency Service Boundary and PSAP polygons especially as they relate spatially to the RCL and ADP. MnGeo is currently finalizing this process in the ECN NG9-1-1 GIS Implementation Roadmap. Details about the validation requirements and process of this phase will be added to this document upon completion of the roadmap.