

RMS® CCRA® Training Program

Geocoding & Hazard Retrieval Exercise: Northeast U.S. Windstorm Portfolio

Learning Objectives:

This exercise demonstrates the impact that address accuracy and geocoding resolution have on exposure data analysis applications.

The following files are available to you to answer the subsequent questions:

- GeoHaz Exercise Data.xls: Exposure data from a RiskLink®-DLM EDM. Includes data from EDM address table, EDM property table, EDM loccvg table, EDM hudet table, and RDM locstats table.
- NE_WIND_VALUExD2C.pdf: Distance to coast summary report
- GeoStan_StatusCodes v19.00.pdf: Centrus GeoStan geocoding status codes output into the RiskLink EDM to record address changes during the geocoding process and spatial accuracy.

PART I – Exposure Data Address Accuracy and Geocoding Resolution Review

You have been presented with a RiskLink-DLM EDM that has a portfolio containing 98 locations throughout the Northeast United States. It has 17 accounts and it includes a variety of commercial and residential properties. Prior to geocoding the portfolio, you review the address information in the address table of the EDM and note that there are a few questionable addresses shown in the table below.

LOCNUM	Address/Description	LOCNUM	Address/Description
24	New Heaven should be New Haven	1039	182222 is an unlikely street number
99	Universalblvd should be two words	985	No ZIP Code
332	Blk Horse Pike should be Black Horse Pike	1284	ZIP Code 67640 is from out-of-state
56	Nework should be Newark	272	East Lime should be East Lyme

Some geocoders, including those used in RiskLink, will automatically correct some address errors, so you geocode the portfolio in RiskLink. The resulting address table with the geocoding results is shown on the Geocoded Data sheet of the GeoHaz Exercise Data.xls spreadsheet.

1. What geocoding resolution(s) did you achieve for the locations in this portfolio?

2. Based on the geocoding level results, which of the previously observed errors were corrected? List the locations which were not corrected and provide a probable explanation as to why.

3. Unit 1 covers spatial uncertainty in address geocoding, and its influence on exposure data analysis and modeling results. Ideally, you want to send your locations into the catastrophe models having eliminated as much uncertainty in geocoding as you have control over. That said, you have achieved street level geocoding in almost all of your locations. Would you recommend fixing the remaining problems to improve these geocoding results before you perform the modeling process on this portfolio, and if so, how?

PART II – Assessing Geocoding and Hazard Uncertainty through Output Codes

Geocoding process changes are recorded in the EDM for some regions and geocoding levels. The next part of the exercise will use the data in the GeoHaz Exercise Data.xls spreadsheet and the GeoStan_StatusCodes v19.00.pdf document to assess geocoding uncertainty. The following data fields exist in the address table of the EDM:

GeoMatchCode – GeoStan match codes that indicate the portions of the address that matched or did not match to the GeoStan directory file. The code digits do not specifically refer to which address elements did not match, but rather to why the address did not match.

GeoLocationCode – This is the GeoStan location code and is an indicator of spatial accuracy of the assigned geocode.

GeoBuffer – Uncertainty code for a location geocoding to a high-resolution (in feet).

GeoResolutionCode – RMS geocoding match level (2 = high resolution [street, parcel or building]; 5 = postal code level)

4. Review the data in the GeoHaz Exercise Data.xls spreadsheet. Is there a relationship between the GeoMatchCode, GeoLocationCode, GeoBuffer, and GeoResolutionCode? If so, what is the relationship? Provide an explanation.

5. What do you notice about the relationship between the GeoMatchCode, GeoLocationCode, and GeoBuffer codes in the GeoHaz Exercise Data.xls spreadsheet?

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

6. Compare the DISTCOAST values with their associated DSCTMATCH values and GeoResolutionCode values in the GeoHaz Exercise Data.xls spreadsheet.

DISTCOAST – Distance to coast in miles

DSCSTMATCH – Hazard match resolution

How do the distance to coast resolutions (DISTCOAST), hazard match resolutions (DSCSTMATCH), and geocoding match (GeoResolutionCode) appear to relate to each other? Provide an explanation.

PART III – Geocoding Impact on Portfolio Conformance to Underwriting Guidelines

Your company-wide underwriting guidelines aim to restrict coastal business throughout the United States. These guidelines establish maximum limits for bands of territory starting at the coast and moving inland. In particular, the guidelines limit coastal exposure in the Northeast to the following limit amounts:

Maximum Coastal Limits by Distance to Coast Band for the Northeast U.S.

Distance to Coast Bands*	Maximum Limits (\$M)
0.0 – 0.3 miles	\$15,000,000
0.3 – 1.0 miles	\$75,000,000
1.0 – 5.0 miles	\$200,000,000
5.0 – 10.0 miles	\$500,000,000
10.0+	unlimited for wind

**Note: The top range is NOT inclusive of the maximum value. In other words, the 0.0 – 0.3 mile band considers all locations that have distance to coast matches from 0.0 miles up through 0.2999 miles. Those with distance to coast of 0.3 miles will fall into the second band.*

Your job is to figure out how to make your portfolio conform to these guidelines – if possible. To identify the current exposure, you refer to a recently run distance to coast summary report (NE_WIND_VALUExD2C.pdf). You also have the relevant exposure and results information available to you to as you start to determine if the portfolio is within the guidelines. (Note: Relevant information from the EDM and RDM is provided in the GeoHaz Exercise Data.xls spreadsheet)

- Does your northeast portfolio comply with the new guidelines? If not, which bands are over the maximum allowed limits?

8. You need to present this information to your Chief Underwriting Officer this afternoon. What additional information (if any) will you include with the exposure summary to help him understand the exposure in this region? Think about whether all geocoding uncertainty has been eliminated in your results, and what confidence level you have in the quality of the location data and geocoding results you have used in your analysis.

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

After reviewing the report, your CUO asks you to make recommendations about how to make this region comply with the new guidelines. Before you make your recommendation, you have decided to review the portfolio at a closer level to make sure that you are not making a decision based on inaccurate information. You decide that the greatest concern in using the exposure information is that there are two locations with postal code resolution geocoding.

9. List two reasons why using these locations with postal code geocoding resolution to make a critical company decision regarding your portfolio is of concern to you. Look not only at the quality of the address data but also at the geocoding results and the limits of each exposure.

ii.

In order to raise your confidence level in the data, you decide to improve the geocoding match for the postal code level locations and then re-run an AAL (average annual loss) analysis to observe whether an improved match would affect the overall risk profile for the portfolio. You correct the addresses that appeared incorrect and that resulted in postal code geocoding resolution matches, namely the “Universalblvd” of LOCNUM 99 and the “1822222” of LOCNUM 1039. The first is easy to correct – i.e., to change it to “Universal Blvd”; but the second may require more research. After some research, you find that “182 Calcutta St” is the valid address. Both corrections now yield high-resolution geocoding matches.

The distribution of limits by distance to coast band is shown in the table below for the original geocoded portfolio as well as the updated portfolio (i.e., the re-geocoded portfolio after correcting LOCNUMs 99 and 1039).

Distance to Coast Band	Maximum Limits (\$M)	Original Limits (\$M)	Updated Limits (\$M)
0.0 – 0.3 miles	\$15,000,000	\$13,992,248	\$13,992,248
0.3 – 1.0 miles	\$75,000,000	\$159,435,653	\$81,705,653
1.0 – 5.0 miles	\$200,000,000	\$451,197,039	\$528,675,919
5.0 – 10.0 miles	\$500,000,000	\$190,183,792	\$190,183,792
10.0+	unlimited for wind	unlimited for wind	unlimited for wind

10. Explain why the original limits and updated limits columns have different values.

11. Taking any other factors into account (for example, location level hazard or loss values within the prescribed distance to coast bands), make some side notes for your final report. Based on what you have learned today, would you recommend that your company continue with windstorm underwriting guidelines based solely on distance to coast? Justify your answer with a description.
