1996 CAS Geo-Coding Survey
by the CAS Committee on Management Data and Information

EXECUTIVE SUMMARY

The purpose of the 1996 Geo-Coding Survey was to assess the current usage of geo-coded data in the casualty actuarial profession, and to foster development of new actuarial techniques using such data. A total of 152 CAS members returned a completed survey. The following are the key findings of the Geo-Coding Survey:

- Nearly four in ten (36.8%) respondents reported they were currently using geo-coded data for the monitoring of catastrophe exposures, while nearly one-third (30.9%) reported current use in the definition of rating territories.

- Over one in five (21.1%) respondents reported they were currently using geo-coded data for the determination of unexpected insurance costs for specific locations, and the same number reported use for marketing/underwriting.

- Close to half (48%) of all respondents reported they were not currently using geo-coded data for any purpose.

- Zip code data was named most frequently by respondents when asked the type of geographic data they were using for listed purposes. Zip code data was the most popular response for six of the seven listed purposes, such as the monitoring of catastrophe exposures or the definition of rating territories.

- Over nine in ten (90.8%) respondents report that they believe geo-coded data will become useful in the monitoring of catastrophe exposures, while over three-fourths (77%) believe geo-coded data will become useful in the definition of rating territories.

- A clear majority of respondents believe that geo-coded will become useful in the determination of unexpected insurance costs for specific locations (63.8%) or in marketing/underwriting (59.9%), while nearly one-half believe geo-coded data will become useful in competitive analysis (48%) or policy rating (47.4%).

- Of those using geo-coded data, nearly two-thirds (62%) indicate the source of latitude/longitude to be software that determines latitude/longitude from street address.

- When asked to describe successful applications of geo-coded data they believed would be of interest to the CAS membership, respondents mentioned catastrophe related applications most often. These applications included catastrophe modeling, catastrophe analysis, and catastrophe management.

- When asked to describe any significant problems in development of geo-coded applications they believe CAS members should be made aware of, respondents mentioned data quality

issues most often. These issues included inconsistency of data gathering, accuracy of geo-coded data software, and accuracy of street addresses and zip codes.

- When asked to provide references that they knew of that may be helpful in the development of geo-coded applications, Mapinfo software and Business Geographic magazine were mentioned most often by respondents.

- Nearly half (48%) of all respondents reported interest in participating as an attendee of a panel discussion focused on development of applications for geo-coding, while close to one-fourth (23.7%) reported interest in participating as a Limited Attendance Seminar attendee.

- A clear majority (57.9%) of respondents reported a designation of FCAS, while over one-fourth (26.9%) reported having over 21 years of actuarial experience.

- Close to six in ten (59.2%) respondents reported a property/casualty primary insurance company as their type of employer.

RESULTS

Applications

Item 1:
Indicate the areas of actuarial practice for which you are currently using geo-coded data:

Area of Actuarial Practice	Percentage of Respondents Using Geo-Coded Data
• Monitoring of catastrophe exposures	36.8
• Determination of unexpected insurance costs for specific locations	21.1
• Reserving	4.6
• Definition of rating territories	30.9
• Policy rating	15.8
• Competitive analysis	16.4
• Marketing/Underwriting	21.1
• Other	3.9
• Not currently using for any purpose	48.0

Please write in the type of geographic data that you are using for the listed purposes.

Area of Actuarial Practice	Type of data in use (number of responses)
• Monitoring of catastrophe exposures	zip code (37) street address (12) county (10) latitude/longitude (6) postal code (3) exposure information (2) various (2) location (1) geo-coded (1)
• Determination of unexpected insurance costs for specific locations	zip code (22) street address (7) county (5) latitude/longitude (4) rating territory (4) postal code (2) state and county (2) exposure information (1) location (1) census tract, block, and group (1)

Applications
(continued)

Please write in the type of geographic data that you are using for the listed purposes (continued).

Area of Actuarial Practice	**Type of data in use** (number of responses)
• Reserving	state (4) street address (1) county (1)
• Definition of rating territories	zip code (37) county (14) postal code (4) state and county (1) city (1) state (1) street (1) rating territory (1) river (1) bureau definitions (1) latitude/longitude (1) census tract, block, and group (1) creation of catastrophe zones to monitor exposure (1)
• Policy rating	zip code (16) county (4) postal code (2) rating territory (2) bureau territory (1) street address (1) town code (1) latitude/longitude (1) distance to work (1) distance to coastline (1) Rand McNally database (1) various (1)

Applications
(continued)

Please write in the type of geographic data that you are using for the listed purposes (continued).

Area of Actuarial Practice	**Type of data in use** (number of responses)
• Competitive analysis	zip code (15) postal code (4) rating territory (3) bureau territory (1) state (1) county (1) exposure information (1) latitude/longitude (1)
• Marketing/Underwriting	zip code (14) street address (4) county (4) latitude/longitude (3) postal code (3) census block group (2) census tract (1) overlay of census data (1) town code (1) bureau territory (1) exposure information (1) target marketing (1) risk selection (1) wind (1) drive distance (1) local tax (1) fire protection (1)
• Other	
Ratemaking - homeowner (2)	exposure information, zip code
Reinsurance	zip code
Agency management	street address, block group
Identifying policy holder in catastrophe areas	latitude/longitude

Applications
(continued)

Please provide any explanatory comments on the above applications that you believe will be helpful to the CAS in assessing the current state-of-the-art for applications of geo-coding.

- I plan to use data for expected losses by location, definition of rating territories, allocation of capital, but these studies aren't underway yet.
- Using IRAS (from RMS) to assist with catastrophe capacity management; considering additional uses, much interest in this area at this time.
- Street address used to match waste sites to publicly available listings.
- Systems development and statistical plans have a long way to go - at this point were getting started on a system to access zip code data. More detailed geo-coding would require substantial changes to statistical plans.
- As a regulator, I had to review the use of geo-coded data, or the proposed use of such data.
- Recently participated in the NAIC study of insurance availability in urban areas. The survey data included policy counts and premiums by homeowners policy form for zip code.
- As a reinsurer, we get limited data, particularly operating in the broker market. We may get for catastrophe exposures, total insured value by county.
- We expect to use expected profit by location to predict profits for given group property and casualty accounts.
- I don't believe our current practice should necessarily be considered "state-of-the-art." We are a small regional insurer, just beginning a major overhaul of our data reporting systems, partly to improve our access to more detailed location data (among many other issues).
- The ratemaking process performed by EQECAT for the California Earthquake Authority is a fine example of the issues I've checked above. (monitoring of catastrophe exposures, determination of expected insurance costs, competitive analysis, marketing/underwriting, other - ratemaking HO).
- A couple of areas that might deserve mention: Use of geo-coding in Business Planning and Strategy; Use of geo-coding in Dynamic Financial Analysis.
- We are moving toward policy-specific geo-coding due to the increasing need to asses risk at the policy level.
- Do not use lat/long for anything explicitly although our mapping software uses it internally.
- We are a reinsurance company, so detailed data is sometimes difficult to obtain and sometimes too time consuming to evaluate. We don't use detailed data for reserving and don't have territory rating.
- (Not currently using, but) Please note that with my previous employer, I was using 3 digits postal code for monitoring of catastrophe exposures and definition of rating territories.
- In all the above areas, zip code level data is used, but no geo-coded data.
- We are finding that zip code is too broad for many of the above applications.
- Data is easily available for US. Very expensive or not available for rest of world.
- I think the CAS should check with major personal lines companies to determine the more sophisticated programs that may be used to define rating territories. Also, check on use of geo-coding in Neural Network Analysis (Peter Wu?).

- There is a lack of geo-coded historical data that is available to set rates, etc.

Item 2:
Indicate the areas of actuarial practice for which you believe geo-coded data will become useful:

Area of Actuarial Practice	Percentage Who Believe Geo-Coded Data Will Become Useful
• Monitoring of catastrophe exposures	90.8
• Determination of unexpected insurance costs for specific locations	63.8
• Reserving	14.5
• Definition of rating territories	77.0
• Policy rating	47.4
• Competitive analysis	48.0
• Marketing/Underwriting	59.9
• Other	7.2

(reinsurance pricing (3), research, claims, business planning, international catastrophes, agency management, more innovative design of rating plans, risk management, allocation of capitol)

Please provide any explanatory comments on the above applications that you believe will be helpful to the CAS in fostering the future state-of-the-art for applications of geo-coding.

- Extent of future possibilities for geo-coding will depend on how precise the location/information provided is! For example, will geo-coding allow me to differentiate between two apartment buildings on the same block; one at 55 Main St, the other 60 Main St?
- Uniform (throughout the industry) use of geo-coding territories would simplify many things, starting with statistical reporting. Many things would flow from this, including simplified competitive analysis. Imagine...no more zip, city, county, town, convoluted territories.
- Anything is possible - it is a wonderful rating variable if the data is available.
- Most consumers/customers don't know their address in geo-coding form so wide spread use without an inexpensive translation mechanism will slow the use of location information in this format for insurance purposes.
- The categories above are redundant (rating territories = policy rating).
- Claims - Would be helpful to have an idea of how many and where to deploy claims adjusters post-event.
- I am especially intrigued with territorial ratemaking that uses continuous surfaces and using data at a street address level.
- Eventually, territory rating may be replaced by the detailed information associated with a given location: home values, crime statistics, etc. For hurricanes and tornadoes, we should consider looking at elevation.
- Ultimately, it will be the way to do business and will be useful for all of the above.

Applications
(continued)

Please provide any explanatory comments on the above applications that you believe will be helpful to the CAS in fostering the future state-of-the-art for applications of geo-coding (continued).

- Given the complexity of working with such data, I envision only those specializing in monitoring catastrophe exposures to be using such data on a regular basis.
- Presentation at the CAS Ratemaking Seminar in Las Vegas provided the state-of-the-art ideas which were excellent.
- Knowing the location and mapping the neighboring exposures to the geo-coded risk you are writing is an invaluable resource.
- Location of insured and of losses could change rating concepts. For instance, automobile rating could depend on where you drive (to work, to grandma's house) besides just where your garage is.
- Location is an important determinant of the concentration and exposure to many perils. It will be more common to use location analysis as data and GIS software becomes widely available.
- Policy holders will not/do not know their geo-code; rate structures need to be simple enough to file and explain. Zip code is simple enough, while latitude/longitude is not.
- Geo-coded data could be particularly helpful in monitoring earthquake exposure.
- For personal auto, could be used for rating purposes to refine traditional zip code rated territories.
- Catastrophe exposures vary over short distances that are not captured any other way.

Sources

Item 3:
If you are using geo-coded data, indicate the sources of latitude/longitude that you are using:

Source	Percentage of Respondents Using Source
• Software that determines latitude/longitude from street address	62.0
• Designation of map location on computer screen	15.2
• On-site radio signals to satellites	6.3
• Other (post office publication of zip codes, all business assigned to geographic center for each zip code, zip codes from location of insured property)	6.3

Please provide comments on the above alternatives that you believe will be helpful to the CAS membership in assessing practical alternatives for development of geo-coded data.

- We use RMS's IRAS model.
- Satellites should be used in the future; e.g. a satellite photo in a hurricane aftermath could be overlaid with a map of company policyholders. Claim services could be routed and estimates could be made prior to claims being filed.
- Relationships of site to known landmarks can be of great value to underwriters, actuaries, and claims. Information such as known pollution sites, or distance to coastline or fault line.
- For determining territorial definition, we use a Rand/McNally database which is keyed by FIPS - place code, state, zip.
- IRAS earthquake model by RMS uses lat/long.

Success and Problems

Item 4:
Please describe successful applications of geo-coded data that you believe will be of interest to the CAS membership:

- Catastrophe modeling.
- Catastrophe exposure.
- Catastrophe analysis.
- Catastrophe management.
- Catastrophe modeling software by EQECAT and RMS.
- Determining catastrophe rates.
- Pricing alternate contract terms.
- Auto theft analysis from trackers.
- NY, NJ hurricane deductible zones.
- FL (FWUA) voluntary credits.
- Territory redefinition using census data and geo-coding.
- Earthquake modeling and predictive modeling.
- I think all the areas in #2 above, except reserving, will provide successful applications.
- Not quite finished, but we're putting together a system to group various postal codes together to analyze appropriate territories and changes in territories.
- Risk location/concentration and exposure determination.
- We examine loss data by postal code in order to define territories and rate policies for homeowners insurance; we review every year and move postal codes from one territory to another as dictated by either loss experience or our competitive position in the marketplace. We once examined loss data by postal code to refine our territory definitions for auto insurance; we will periodically review to make sure out territory definitions still make sense.
- Result is only as good as the initial address entry. Misspellings, etc., can have a large impact.
- Using 3-digits postal code to refine large urban territories (both property and automobile).
- IRAS earthquake loss model.
- Marketing, underwriting, rating.
- Entire book of biz geo-coded in US. GPS units in use in Asia and Latin America.
- Assessing hazard exposure is very successful using geo-coded data.
- Obviously, the area of greatest payback would be in monitoring catastrophe exposures. This has helped our company tremendously in negotiations with our catastrophe reinsurers. Also, extremely useful internally for business planning.
- Currently assigning rating territories from keyed in address; agents/customer service reps no longer need to learn territory definitions.
- I feel that zip code rating is superior.
- Coastal exposures can be problematic when trying to view exposures. Some geo-coding software are not consistent so data that may be geo-coded by one software but later used in another may appear to be out in the "water".
- Software to display physical locations of zip codes to check for contiguousness of territories. Also to view patterns in loss costs.
- Density of exposures related to hurricane or earthquake risk.

Success and Problems
(continued)

Item 4 (continued):
Please describe successful applications of geo-coded data that you believe will be of interest to the CAS membership:

- As a reinsurer, we have used it only to monitor our catastrophe exposures. We rely on the reinsurance intermediaries and/or clients to collect data and run models.
- Hurricane modeling, reinsurance exposure analysis, rating territories.
- Map of geological fault lines and location of exposure by geo-coded location with reference to fault lines.
- I use geo-coded data to monitor earthquake exposure in a region.
- Geo-coding makes it much easier for a direct writer phone operation to properly determine the rating territory since the insurer representative may not know the geographic area where the risk is located.

Item 5:
Please describe any significant problems in development of geo-coding applications you believe CAS members should be made aware of:

- Blanket rated covers enough information; may not currently be captured.
- Constant updating of new addresses.
- Addressing the large variability around the expected results.
- Risks with hundreds/thousands of locations - difficult to charge enough to cover costs of capturing all the specific location data.
- Consistency in measuring devices (accuracy).
- Several, separate policies for the same coverage written on risks at the same geo-code, e.g. earthquake coverage for renters in a high-rise.
- Major problem would be credibility associated with finely divided data.
- Data quality problems have surfaced frequently.
- Multi-location policies are especially difficult to handle.
- Distribution of geo-coded applications to agents.
- Data quality, model parameter estimating, regulatory concerns.
- Capturing accurate raw data.
- The major problems I see are dealing with what are sure to be large quantity of the data, (i.e. cut so fine that many observations are required before anything useful can be gained) and relating the data to other, known, data for verification.
- Computer software still (at least what we use) can not interpret similar spellings.
- Accuracy of TIGER, census, street information in the software.
- Quality of internal data - for example capturing billing address vs. site address.
- Inconsistency of data gathering, e.g. bad zip code in valid city.
- Difference between loss location and premium location.
- Zip codes change constantly, keeping up with them is costly.
- Lack of data quality of street address leading to low hit rates for geo software.

Success and Problems
(continued)

Item 5 (continued):
Please describe any significant problems in development of geo-coding applications you believe CAS members should be made aware of:

- We have challenges with "split zips" - a zip code which has more than one rate. Our rating system can't use the person's address to properly place them in the section of the zip; we rely on a person to look up information on a map or a descriptive table - sometimes the person just assigns "the first one" of the two or more choices.
- Unavailability of latitude/longitude coordinates for most locations. Do you know the coordinate for your home??
- Data quality - We have problems where zip codes sometimes don't map on the correct state. We find mailing address coded on all risks on one policy, rather than the actual location.
- Because of the ever changing numbers of zip codes, it is important to utilize software that is regularly updated. Data quality of the addresses being utilized is also a watch out.
- Accuracy concerns on street segments.
- Accuracy of geo-coding/underlying data.
- To match the longitude/latitude, you need exact address. Many addresses are not exactly correct, i.e. street instead of drive, north or south not included with street, etc.
- Coding of varying limits, classification information, etc. by location on a single policy.
- It is difficult to determine whether county data is "good enough" or if we need zip code or street address data.
- Credibility issue as the volume of data decreases rapidly.
- Redlining issue.
- Software needs to be continually updated for new zip code definitions.
- Regulatory acceptance; it shouldn't be a problem, but I fear that it will; change is difficult.
- Need to be careful of over-refinement of territories.
- Need to be careful when using for auto insurance because cars are mobile (unlike houses).
- Level of data needed to make geo-coded location data valuable, i.e. coverage, limits, type of construction, contents, etc.
- Credibility procedures! / Lack of credibility. / Credibility issues and techniques.
- The actual coding by street address is complicated by numerous factors. The software's street index is incomplete. Finding the location as opposed to billing address can be difficult and multiple location policies are present.
- Blanket risk coding on commercial accounts.
- Annual or more frequent shifts in zip code boundaries.
- A Canada conversion of rural addresses or postal codes to geo-code may not be accurate enough for certain applications (i.e. catastrophe analysis).
- Perils that are not discrete enough for high resolution analysis: hail, brush fires, mud slide.
- Willingness to believe model output without appreciation for inherent uncertainty.
- Errors in geo-coding and address changes.
- What is the source of this information for a property? If the insured, it can be falsified. Systems development costs, cost of capturing data, lack of interest by senior management.

References

Item 6:
Please provide any references that you know of that may be helpful in the development of geo-coding applications. Such references may be printed materials, vendor organizations, professional organizations, or any other source of information:

Multiple responses are indicated by the number in parentheses

- RMS (2)
- Strategic Mapping Inc. (2)
- ETAK (2)
- Advance Technology Corporation, Atlanta, GA, Mark Fouraker, 770-399-4343
- Toprate, Insurquote (rating services)
- Maplinx (software)
- Mike Miller (actuarial consulting)
- American Demographic magazine (2)
- Business Geographic magazine (4)
- ESRI (3)
- Mapinfo software (4)
- EQECAT
- Workers Compensation Insurance Rating Bureau (California) - data by zip code
- GIS World Magazine
- Geographic Data Technology (GDT) (2)
- ISO
- Vista Information Services
- The software we use is called IRAS from the vendor Risk Management Solutions.
- Tactician Corporation (software)
- USGS
- Compu search organization operating in Toronto

Future CAS Programs

Item 7:
Please indicate below the type of future activity focused on development of applications for geo-coding in which you would be interested in participating:

Type of Activity	Percentage of Respondents with Interest
Panel Discussion - attendee	48.0
Panel Discussion - panel member	3.9
Discussion Paper Program - author	0.0
Discussion Paper Program - reviewer	5.3
Limited Attendance Seminar - attendee	23.7
Limited Attendance Seminar - discussion leader	1.3
Other	.7

Other Comments

Item 8:
Please provide any additional comments that you believe would be helpful to the CAS in facilitating development of geo-coding applications among CAS members:

- As a regulator, I see future concerns about territorial definitions. How will they look? Will they exist? How will the laws of various states need modification to adapt to the changing technology?
- Must recognize difference in personal lines vs. commercial lines risks with hundreds of locations.
- Workshops, data source description, model building papers, etc.
- Coin the term "geode" to mean the smallest geographical unit under consideration. Geodes would be defined by regulators with industry assistance. Geodes should be along easily identifiable physical or political boundaries. Territories would be aggregations of geodes. Most Importantly: In order for computerized hurricane models to become accepted, they must utilize real exposure and loss data to calibrate the model. Exposure and loss data should be reported by geode and each geode should be small enough that we would expect that wind damage would be uniform throughout the geode.
- I'm not confident that expending resources in this direction would be fruitful for the CAS.
- ISO geog u/w system.
- Geo-coding in ratemaking will require a great deal of information, so only the largest companies will have sufficient data. To make geo-coding of interest (of practical interest to many actuaries' employers), a large database would need to be available. Perhaps the CAS would work with statistical agencies to gather this information so that it is reliable and available. This might expand the interest level among actuaries and help the small companies from being the victims of large companies that can use this information to exploit the current rating territory definitions.
- There must be other more important issues to be spending time on.
- How about an article in the Actuarial Review?
- What about the impact on companies' systems departments?
- Geo-coding has not been used by me because my job function does not.
- Not much to offer!
- Everyone benefits from a universal adoption of risk location identification. Reinsurers, regulators, statistic reporters, primary companies, etc. all seeing risk the same way would be ideal. Geo-coding (at least on the surface) would seem to offer this.
- Not to get too carried away, but will the CAS be seeking the NAIC's input on the potential for geo-coding to become mandatory?
- Wouldn't it be nice to have a regulation promulgated that actually helped an entire industry?
- Any discussion should consider practical applications.

Member Profile

Item 9:
Please include some information about yourself:

Actuarial Designation	Frequency	Percent
FCAS	88	57.9
ACAS	57	37.5
No response	7	4.6
Total	152	100.0

Years of Actuarial Experience	Frequency	Percent
0-5	10	6.6
6-10	36	23.7
11-15	31	20.4
16-20	29	19.1
21+	41	26.9
No response	5	3.3
Total	152	100.0

College Degree	Frequency	Percent
None	1	.7
BA or equivalent	109	71.7
MA or equivalent	34	22.3
Ph.D.	3	2.0
No response	5	3.3
Total	152	100.0

Type of Employer	Frequency	Percent
Property/Casualty Primary Insurance Company	90	59.2
Reinsurance Company	11	7.2
Consulting Firm	26	17.1
Insurance Broker	8	5.3
State Insurance Department	5	3.3
Other Government Entity	0	0.0
Organization serving the insurance business	4	2.6
University or college	1	.7
Other	2	1.3
No response	5	3.3
Total	152	100.0