

LOSS COSTS – FILED AND IMPLEMENTATION

NOVEMBER 3, 2022

COMMERCIAL PROPERTY

LI-CF-2022-146

NEW MEXICO COMMERCIAL PROPERTY EARTHQUAKE LOSS COSTS REVISION PROVIDED AND TO BE IMPLEMENTED

KEY MESSAGE

Loss Costs revision for filing CF-2022-REQLC in New Mexico is provided and being implemented.

Distribution Date: 8-23

BACKGROUND

In circular [LI-CF-2022-074](#), we advised that we would be submitting state-specific loss costs revisions in all ISO jurisdictions to reflect various enhancements pertaining to Earthquake coverage in companion rules filing CF-2022-REQRU.

ISO ACTION

We are providing and implementing the attached New Mexico loss costs revision in filing CF-2022-REQLC.

Refer to the attached explanatory material for complete details about the filing.

EFFECTIVE DATE

We do not establish an effective date for Commercial Property loss costs revisions in this state. Each insurer that elects to utilize this revision is responsible for determining its own effective date.

COMPANY ACTION

You must independently determine the final rates you will use and the effective date of any rate change. If you decide to use our prospective loss costs to revise your rates, you are NOT required to file anything with the Insurance Department.

You must document your files in case the Insurance Department wishes to review the information at a later date. In all correspondence on this revision, you should refer to ISO Filing Number [CF-2022-REQLC](#), NOT this circular number.

RATING SOFTWARE IMPACT

New attributes being introduced with this revision:

- Additional information will be required from the policyholder to complete a rating calculation.
- A new code is being introduced.
- Current loss costs are being withdrawn.
- Current factors are being withdrawn.

- A new calculation is being introduced.
- An existing rating formula is being rewritten.

POLICYHOLDER NOTIFICATION

If you decide to implement this revision, you should check all applicable laws for the state(s) to which this revision applies, to determine whether or not a specific policyholder notice requirement may apply. Please note that circular [LI-CL-2022-006](#) contains the ISO Guide To Renewals With Changed Conditions For Commercial Lines, which is available only as a guide to assist participating companies in complying with various conditional renewal statutes or regulations, for the major commercial lines of insurance serviced by ISO. The information in the Guide does not necessarily reflect all requirements or exceptions that may apply, and it is not intended as a substitute for your review of all applicable statutes and regulations concerning policyholder notification.

RELATED RULES FILING

In circular [LI-CF-2022-145](#), we are providing and implementing the corresponding rules supplement.

REVISION DISTRIBUTION

We will issue a Notice to Manualholders with an edition date of 8-23 (or the earliest possible subsequent date), along with any new and/or revised manual pages.

REFERENCE(S)

- [LI-CF-2022-145](#) (11/03/2022) New Mexico Supplement To The Commercial Property Multistate Earthquake Rules Revision Provided And To Be Implemented
- [LI-CF-2022-074](#) (08/10/2022) Commercial Property Multistate Rules Revision Being Submitted
- [LI-CL-2022-006](#) (02/22/2022) Revised Lead Time Requirements Listing

ATTACHMENT(S)

New Mexico Filing CF-2022-REQLC

FILES AVAILABLE FOR DOWNLOAD

To download all files associated with this circular, including attachments in the full circular PDF and/or any additional files not included in the PDF, search for the circular number on [ISOnet Circulars](#). Then click the Word/Excel link under the Full Circular column on the Search Results screen.

Please note that in some instances, not all files listed in the Attachment(s) block (if applicable) are included in the PDF.

DATA QUALITY

Statistical plan data reported to ISO is first processed through a system of rigorous automated data verification procedures so that only valid data would be used for ratemaking. Subsequent to this initial data submission review, additional analyses on the statistical plan data involving an even more customized data review for this line were performed by staff. During these processes, various data records were excluded from the review. The ISO staff responsible for this circular also reviewed the data for reasonableness.

ACKNOWLEDGMENT OF ACTUARIAL QUALIFICATIONS

The American Academy of Actuaries' "Qualifications Standards for Actuaries Issuing Statements of Actuarial Opinion in the United States" requires that an actuary issuing a Statement of Actuarial Opinion should include an acknowledgment with the opinion that he/she has met the qualification standards of the AAA. ISO considers this loss costs filing a Statement of Actuarial Opinion; therefore, we are including the following acknowledgment:

I, Bei Zhou, am an Actuarial Product Director for Commercial Property and Actuary for ISO. I am responsible for the content of this Statement of Actuarial Opinion. I am a member of the Casualty Actuarial Society and I meet the Qualification Standards of the American Academy of Actuaries to render the actuarial opinion contained herein.

COPYRIGHT EXPLANATION

The material distributed by Insurance Services Office, Inc. is copyrighted. All rights reserved. Possession of these pages does not confer the right to print, reprint, publish, copy, sell, file or use same in any manner without the written permission of the copyright owner. Permission is hereby granted to members, subscribers and service purchasers to reprint, copy or otherwise use the enclosed material for purposes of their own business use relating to that territory or line or kind of insurance, or subdivision thereof, for which they participate, provided that:

- (A) Where ISO copyrighted material is reprinted, copied, or otherwise used **as a whole**, it must reflect the copyright notice actually shown on such material.
- (B) Where ISO copyrighted material is reprinted, copied, or otherwise used **in part**, the following credit legend must appear at the bottom of each page so used:
Includes copyrighted material of Insurance Services Office, Inc., with its permission.

IMPORTANT NOTICE FOR USERS OF ISO PRODUCTS AND SERVICES

Please make sure that your company has authorized your use of this product and has complied with the requirements applicable in the jurisdiction where you plan to use it.

We distribute both state-specific and multistate products and services. We do not distribute all the multistate products and services for use in every jurisdiction due to corporate policy, regulatory preference, or variations or lack of clarity in state laws.

We provide participating insurers with information concerning the jurisdictions for which our products and services are distributed. Even in those jurisdictions, each insurer must determine what filing requirements, if any, apply and whether those requirements have been satisfied.

Now, as in the past, all of our products and services are advisory, and are made available for optional use by participating insurers as a matter of individual choice. Your company must decide for itself which, if any, ISO products or services are needed or useful to its operation and how those selected for use should be applied. We urge that you be guided by the advice of your attorneys on the legal requirements.

CONTACT INFORMATION

If you have any questions concerning:

- The content of this circular and status of this filing, please contact:

Ron Punnoose
Commercial Lines Actuarial
201-469-2275
Ron.Punnoose@verisk.com
propertyactuarial@verisk.com

- Other issues for this circular, please contact Customer Support:

E-mail: info@verisk.com
Phone: 800-888-4476

Callers outside the United States, Canada, and the Caribbean may contact us using our global toll-free number (International Access Code + 800 48977489). For information on all ISO products, visit us at www.verisk.com/iso. To keep abreast of the latest Insurance Lines Services updates, view www.verisk.com/ils.

NEW MEXICO

COMMERCIAL EARTHQUAKE INSURANCE LOSS COST LEVEL REVISION EXECUTIVE SUMMARY

PURPOSE

This document:

- revises advisory prospective loss costs for Commercial Earthquake. ISO used an earthquake model developed by AIR to develop revised territories and loss costs. These loss costs represent a -40.7% statewide average change from the current ISO Earthquake loss costs.
- provides loss costs for revised Building Classes. The revised building classification definitions are provided in the companion rules filing CF-2022-REQRU.
- provides loss costs for revised Earthquake territories. The revised territory definitions are provided in the companion rules filing CF-2022-REQRU.
- revises Earthquake Personal Property rating procedure
- describes the procedure used to derive these prospective loss costs.

DEFINITION OF THE ISO PROSPECTIVE LOSS COSTS

Advisory prospective loss costs in this document are that portion of a rate that does not include provisions for expenses (other than loss adjustment expenses) or profit.

LOSS COST LEVEL CHANGES

The territory loss cost level changes are:

Territory	Indicated Earthquake Loss Cost Change*	Selected Earthquake Loss Cost Change*	Commercial Property Impact**
Territory 1	-85.5%	-85.5%	-2.3%
Territory 2	-31.4%	-31.4%	-1.5%
Territory 3	-14.7%	-14.7%	-1.0%
Territory 3A	247.2%	93.7%	1.6%
Total	-38.5%	-40.7%	-1.6%

* The indicated loss cost level changes are average changes from the current ISO loss costs for Earthquake coverage only. Changes vary by construction, coverage (building/personal property), and prior Earthquake territory.

** Each territory's commercial property impact is calculated as the change in total Commercial Property loss cost -- i.e., BG I, BG II, SCL and Earthquake loss costs combined. See page A-2 for more detail regarding this calculation.

PRIOR ISO
REVISIONS

The latest revisions in this state are:

<u>Reference Document or Filing</u>	CF-2008-REQ1	CF-98-REQ1
<u>Rates/ Loss Costs</u>	Loss Costs	Loss Costs
<u>Dates</u>		
Filed	01/08/2009	07/01/1998
Implemented	06/01/2009	02/01/1999
<u>Changes</u>		
Entire State	+67.5%	-23.0%

BASIS FOR THE
INDICATED LOSS
COSTS

The loss costs in this revision are based on the expected earthquake loss costs generated by Earthquake model, Touchstone Version 8.2.0, from AIR. AIR is a pioneer in developing catastrophe modeling techniques, which are now widely used to estimate potential catastrophe losses and is today a leading modeling and technology firm for risks associated with natural and man-made catastrophes, weather and climate.

See Section B for more information on the AIR model.

ADJUSTMENTS TO
MODEL OUTPUT

The model output was adjusted to reflect the following:

- 80% Coinsurance Provision
- Loss Adjustment Expenses
- Personal Property Loss Cost Relativities

COMPANY
DECISION

We encourage each insurer to decide independently whether the judgments made and the procedures or data used by ISO in developing the loss costs contained herein are appropriate for its use. We have included within this document the information upon which ISO relied in order to enable companies to make such independent judgements.

An individual company may benefit from comparison of its own experience or other models to the ISO loss costs, and may reach valid conclusions with respect to the manner in which its own costs can be expected to differ from ISO loss costs based on the model.

Some calculations included in this document involve areas of ISO staff judgment. Each company should carefully review and evaluate its own experience in order to determine whether the ISO selected loss costs are appropriate for its use.

The material has been developed exclusively by the staff of Insurance Services Office, Inc. ISO staff has relied on information, and unique knowledge and expertise, provided by AIR (a wholly-owned subsidiary of ISO, Inc.) for the derivation of the modeled loss costs used in this document.

COPYRIGHT
EXPLANATION

The material distributed by Insurance Services Office, Inc. is copyrighted. All rights reserved. Possession of these pages does not confer the right to print, reprint, publish, copy, sell, file or use same in any manner without the written permission of the copyright owner.

NEW MEXICO

COMMERCIAL EARTHQUAKE INSURANCE

TABLE OF CONTENTS

SECTION A - Summary of Changes

Indicated and Selected Commercial Earthquake Loss Cost Changes by Territory and Estimated Impact on Commercial Property Loss Costs (Table 1).....	A2
Selected Commercial Earthquake Loss Cost Changes by Current and Revised Earthquake Territory (Table 2).....	A3
Map of the Current Territories (Table 3).....	A4
Map of the Revised Territories (Table 4).....	A5

SECTION B - Overview of the Earthquake Model and Adjustments to the Model

Description of the AIR Earthquake Model.....	B2-8
Selection of Territories	B9-10
Personal Property Loss Costs Revision.....	B11-14
Adjustments to the Model Output	B15
Loss Adjustment Expense Factors (Table 5).....	B16

SECTION C

Commercial Earthquake Loss Costs.....	C2-3
---------------------------------------	------

APPENDIX A

K-means Clustering.....	AX1-2
-------------------------	-------

NEW MEXICO

COMMERCIAL EARTHQUAKE INSURANCE

SECTION A - Summary of Changes

Indicated and Selected Commercial Earthquake Loss Cost Changes by Territory and Estimated Impact on Commercial Property Loss Costs (Table 1).....	A2
Selected Commercial Earthquake Loss Cost Changes by Current and Revised Earthquake Territory (Table 2).....	A3
Map of the Current Territories (Table 3).....	A4
Map of the Revised Territories (Table 4).....	A5

NEW MEXICO

TABLE 1

Indicated and Selected Commercial Earthquake Loss Cost Changes by Territory

<u>Territory</u>	<u>Indicated Earthquake Loss Cost Change</u>	<u>Selected Earthquake Loss Cost Change</u>	<u>Commercial Property Impact (a)</u>
Territory 1	-85.5%	-85.5%	-2.3%
Territory 2	-31.4%	-31.4%	-1.5%
Territory 3	-14.7%	-14.7%	-1.0%
Territory 3A	247.2%	93.7%	1.6%
Total	-38.5%	-40.7%	-1.6%

(a) The impacts of the selected Earthquake loss costs on average commercial property loss costs (including the current Earthquake loss costs) are calculated as (selected Earthquake loss cost + average commercial property loss cost) divided by (current Earthquake loss cost + average commercial property loss cost) - 1. Average commercial property loss costs are for Basic Group I, Basic Group II, and Special Causes of Loss combined, property damage coverages only, and reflect package modification and other rating factors, based on ratemaking experience through 09/30/2020. These impacts assume Earthquake coverage and, therefore, represent an average commercial property change for those who purchase Earthquake coverage. Since not all insureds purchase coverage for earthquake, the average impact on commercial property policies in New Mexico would be smaller.

NEW MEXICO

TABLE 2

Selected Commercial Earthquake Loss Cost Changes by Current and Revised Earthquake Territory

<u>Revised Territory</u>	<u>Current Territory</u>	<u>Selected Earthquake % Change (a)</u>	<u>Commercial Property Impact (b)</u>
1	22	-83.1%	-1.6%
1	23	-88.2%	-3.3%
1	24	-65.5%	-0.7%
2	21	-40.8%	-2.3%
2	22	-5.8%	-0.2%
2	23	-5.1%	-0.2%
2	24	11.4%	0.2%
3	22	-41.7%	-2.8%
3	23	-2.5%	-0.2%
3A	24	93.7%	1.6%

(a) The revision of ZIP code based territories can result in one old territory contributing to multiple new territories. The changes shown here capture the average impact for each of these old territory areas. There is further variation by construction and coverage.

(b) The impacts of the selected Earthquake loss costs on average commercial property loss costs (including the current Earthquake loss costs) are calculated as (selected Earthquake loss cost + average commercial property loss cost) divided by (current Earthquake loss cost + average commercial property loss cost) - 1. Average commercial property loss costs are for Basic Group I, Basic Group II, and Special Causes of Loss combined, property damage coverages only, and reflect package modification and other rating factors, based on ratemaking experience through 09/30/2020. These impacts assume Earthquake coverage and, therefore, represent an average commercial property change for those who purchase Earthquake coverage. Since not all insureds purchase coverage for earthquake, the average impact on commercial property policies in New Mexico would be smaller.

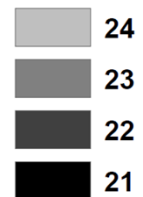
NEW MEXICO

TABLE 3
Current Territories



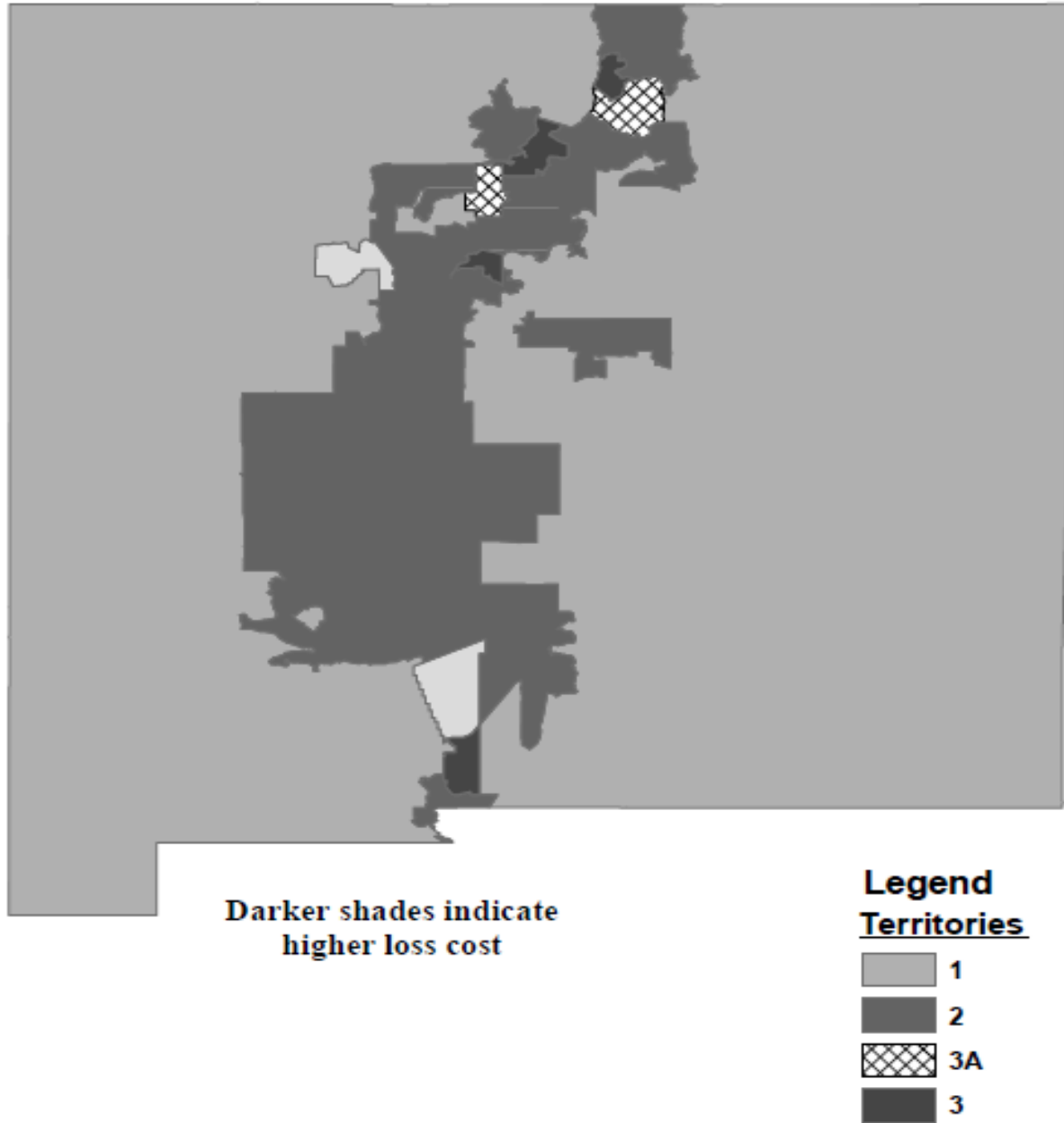
**Darker shades indicate
higher loss costs**

Legend
Territories



NEW MEXICO

TABLE 4
Revised Territories



NEW MEXICO

COMMERCIAL EARTHQUAKE INSURANCE

SECTION B - Overview of the Earthquake Model and Adjustments to the Model

Description of the AIR Earthquake Model	B2-8
Selection of Territories	B9-10
Personal Property Loss Costs Revision	B11-14
Adjustments to the Model Output.....	B15
Loss Adjustment Expense Factors (Table 5)	B16

DESCRIPTION OF THE AIR EARTHQUAKE MODEL

INTRODUCTION

An earthquake is caused by relative movement of the earth's crust and results in widespread shaking or trembling of the ground. This shaking can cause significant damage to buildings and contents. Since major earthquakes, causing widespread damage, occur very infrequently, a state-of-the-art earthquake computer model developed by AIR-Worldwide has been utilized in order to improve the accuracy and reliability of our advisory prospective loss costs.

OVERVIEW OF THE AIR EARTHQUAKE MODEL

The AIR earthquake model consists of the following components or modules:

- The Event Generation Module -- This module determines the frequency, magnitude and other characteristics of potential catastrophe events by geographic location. Scientific data from many sources are used to determine probability distributions for the many variables used to characterize an earthquake. These distributions are used to create a large catalog of simulated events and sampling from these distributions generates "years" of events. Many thousands of years are generated to produce a complete and stable range of potential annual experience of catastrophe event activity based on the latest available research, and to ensure full coverage of extreme events, as well as full spatial coverage.
- The Local Intensity Module -- Once the model probabilistically generates the characteristics of a simulated event, it propagates the event across the affected area. For each location within the affected area, local intensity is estimated. The intensity experienced at each site is a function of the magnitude of the event, distance from the source of the event, and a variety of local conditions (e.g., soil type).
- The Damage Module -- AIR scientists and engineers have developed mathematical functions called damageability relationships, which describe the interaction between buildings, both their structural and nonstructural components as well as their contents, and the local intensity to which they are exposed. These functions relate the mean damage level as well as the variability of damage to the measure of intensity at each location. Because different structural types will experience different degrees of damage, the damageability relationships vary according to construction type. Estimated losses are calculated by applying the appropriate damage function to the replacement value of the insured property.
- The Insured Loss Module -- This module determines the insured losses by applying policy conditions such as the appropriate percentage deductible to the total damage estimates resulting from the damage estimation module.

EVENT GENERATION MODULE

The event generation module generates frequency of earthquakes and key characteristics of each simulated earthquake including the epicenter, magnitude, rupture length and orientation, and depth.

Data Sources

The AIR earthquake modeling team collects information on historical earthquakes in the United States from a variety of sources, the most important of which are:

- United States Geological Survey (USGS)
- Electric Power Research Institute (EPRI)
- National Geophysical Data Center (NGDC)
- Southern California Earthquake Center (SCEC)
- Multidisciplinary Center for Earthquake Engineering Research (MCEER)
- Seismological Society of America (SSA)
- California Department of Mines and Geology (CDMG)

AIR's historical earthquake catalog represents a compilation and synthesis of several catalogs in the public domain, among them the NCEER-91 (Seeber and Armbruster, 1992), PDE 1985–present (Preliminary Determination of Epicenters, a monthly report), Seismicity of the United States (Stover and Coffman, USGS, 1993), and the United States Earthquake Data File (Stover, Reagor, and Algermissen, USGS, 1984), several USGS Open File Reports by Charles Mueller, and other reports.

Modeled Earthquake Characteristics

Frequency-Magnitude Distribution - Seismologists typically fit historical data on the frequency and magnitude of earthquakes to an exponential distribution called the Gutenberg-Richter (GR) relationship. The GR relationship applies globally and allows an extrapolation from the limited historical record to estimate a more complete picture of seismicity in an area.

The GR relationship holds over a wide range of magnitudes and can be described by two parameters: an occurrence rate of earthquakes of magnitude greater than or equal to some reference magnitude, characterized by the so-called “a-value”; and a “b-value” representing the rate at which the log of the cumulative annual frequency of earthquakes decreases as the magnitude increases (the slope). Scientists usually truncate this relationship at a limiting magnitude, m_x , above which the probability of occurrence is zero. Each of these three parameters depends upon the geology of the seismic zone under consideration.

EVENT
GENERATION
MODULE (cont'd)

While the GR relationship holds on a regional or global scale, it may not hold for individual faults. For some seismic zones, there exists evidence that earthquakes of a certain magnitude occur with a frequency that is not consistent with the rate predicted by the GR relationship. Scientists now believe that many faults tend to produce repeated earthquakes of a size that is “characteristic” of that particular fault or fault segment. It is from both the GR distribution and the estimated recurrence rate of characteristic earthquakes that the number of earthquakes that occur in each simulated year and their magnitudes are modeled.

Based on geological criteria, AIR seismologists have divided the country into two broad regions for purposes of modeling seismic risk in the United States. These two regions can be roughly categorized as corresponding to the plate boundary zone of the Western United States (WUS) and the intraplate zone of the Central and Eastern United States (CEUS). The earth's lithosphere consists of several large slabs of rock called plates.

These plates move upon an underlying region of hotter, less rigid rock. Most of the earth's seismic energy is released at plate boundaries where plates come in contact with each other. However, earthquakes also occur in the interior of plates. Geologists believe these intraplate earthquakes may be caused by ancient geological deformations or by variations in temperature and strength of the lithosphere.

To determine the frequency-magnitude distributions for earthquakes in different seismic zones, AIR scientists use all available information for each specific region. This includes historical earthquake catalogs and auxiliary geological data such as fault slip rates, paleoseismic data, geophysically derived moment rates, and tsunami records.

Focal Depth - In the AIR model, focal depth, or the depth beneath the surface of the earth where the initial rupture occurs, is drawn from a Gaussian distribution. Because seismic waves attenuate as they travel through the earth's crust, deeper earthquakes typically cause less damage. The parameters of the Gaussian distribution depend on the magnitude of the simulated earthquake and the actual depth of the seismogenic zone (the brittle upper crust within which earthquakes occur), which can vary from one seismotectonic region to another.

EVENT
GENERATION
MODULE (cont'd)

Rupture Length - Rupture length is a function of magnitude. Characteristic faults may be single or multisegmented. In the case of single-segmented faults, if a characteristic earthquake occurs, it is assumed to rupture the entire length of the fault. In the case of multisegmented faults like the San Andreas Fault, one or more contiguous segments may rupture simultaneously. In instances of such cascade events (the rupture of more than one segment), fault length is a function of the combined events.

LOCAL INTENSITY
MODULE

Local intensity in the AIR earthquake model for the United States consists of a mathematical description of ground motion (shaking intensity) and, for many seismic zones, settlement caused by liquefaction.

Shaking Intensity and Attenuation

After the model generates the source parameters of each simulated earthquake, it calculates the shaking intensity at each location affected by the event. This ground motion can range from barely perceptible trembling to violent shaking depending on the magnitude of the event, distance from the rupture, the geological characteristics of the region, and local site conditions. The reduction in intensity as the energy from an earthquake moves away from its source is referred to as attenuation. A series of recent studies have resulted in the development of attenuation functions for different ground motion parameters that are specifically designed for the various regions of the continental United States.

The parameters of the attenuation functions can vary significantly from one region to another, reflecting the underlying geology of the region. For example, attenuation in the Western United States is much higher than in the Central and Eastern United States. The tectonic stresses along the plate boundary region of the Western United States result in significant fracturing of the rock there. This fracturing results in a relatively rapid scattering of seismic energy. The denser rock of the Central and Eastern United States propagates seismic energy over a much larger area, since the rate by which the amplitude of seismic waves diminishes is much slower. An earthquake of a given magnitude will, accordingly, typically affect a smaller area in the western United States than will an equivalent earthquake in the Central or Eastern United States.

Calculations of local shaking intensity are modified to reflect local site conditions. These modifications are based on an analysis of high-resolution geological and soil maps, as well as characteristics of the soil at different depths obtained from bore log data.

DAMAGE ESTIMATION MODULE

Damage estimation for the United States regional earthquake model includes shake damage, which results from the lateral, vertical and torsional motions of the ground as well as damage resulting from deformation as buildings settle into liquefied soils.

Shake Damage - Advanced Component Method™

The AIR earthquake model for the United States incorporates a groundbreaking, state-of-the-art methodology for assessing and modeling building vulnerability to ground shaking. The Advanced Component Method (ACM™), developed by AIR engineers, largely replaces the subjective measures and expert opinion on which traditional methodologies have relied and is based instead on objective, rigorously scientific analytical procedures for measuring building deformation and damage.

Building Damage - Buildings are damaged when they are exposed to intense relative deformations that result from ground shaking. The response of individual buildings to ground shaking varies dramatically, however, depending on their configuration and natural period, which is defined as the time it takes for a building to complete a single cycle of oscillation, starting from an initial position and returning back to its initial position. The variation in the response of buildings to ground motion becomes immediately apparent when observing actual damage patterns after a major earthquake. In the course of post-disaster surveys conducted by AIR engineers, completely collapsed buildings could be seen next to buildings that remained entirely intact.

Factors determining the natural periods of buildings include stiffness, mass, and “fixity” condition, or lateral force resisting system, which relate to the connecting elements that tie column to foundation, beam to column, etc. within a building. In general, as stiffness and fixity increase, the natural period of the building decreases. Similarly, the less slender a building is, the lower will be its natural period, in general. The movement of very stiff, or rigid, structures tends to mirror the movement of the ground. While there may be little internal structural deformation, the magnitude of the lateral forces that must be absorbed is significant. Flexible structures, on the other hand, deform.

Such deformation results in damage to both structural and non-structural components, though such deformation is less likely to lead to complete failure. The philosophy behind present design codes is to maximize the ductility (i.e., the degree to which buildings can deform beyond the initial point at which damage occurs but before total collapse) of buildings, and particularly of a building’s lateral force resistance system.

DAMAGE
ESTIMATION
MODULE (cont'd)

Traditionally, models have estimated building vulnerability to earthquakes based either on some subjective measure of intensity, such as the Modified Mercalli Intensity (MMI), or on a single parameter of ground motion, such as Peak Ground Acceleration (PGA). But PGA cannot differentiate between the structural characteristics of individual buildings. Indeed, using PGA as the measure of intensity implicitly assumes that, for all buildings, the top of the building moves exactly in unison with the bottom, that is, like the ground itself. While this may be a reasonable approximation for rigid, low-rise buildings, it provides a misleading picture of how other building types and portfolios of buildings will move and deform in response to the ground motion they experience.

ACM uses spectral displacement as an objective measure of intensity. Spectral displacement is the maximum horizontal displacement experienced by a building during an earthquake. When displacement occurs, the building and its component parts are deformed; deformation causes damage. In ACM, damage depends not on how the ground is shaking, but rather on how the building is shaking; specifically, it depends on the deformation response of individual structural and nonstructural components to ground motion. Because each building has different mechanical characteristics and a different natural period, each will be subjected to a different seismic intensity (i.e., spectral displacement) and, hence, a different damage state.

Contents Damage - In the AIR earthquake model for the United States, contents damageability is a function of occupancy class. Occupancy class provides insight into the kinds of contents contained in the building and hence their relative vulnerability. The primary determinant of contents damage in ACM is not spectral displacement, but rather spectral acceleration. Contents, unlike beams and columns, do not detect interstory drift. The lateral forces they experience are proportional to their mass multiplied by the floor acceleration.

While spectral acceleration is the primary determinant of contents damage, building damage as determined by Spectral displacement also results in contents damage.

Contents will be damaged when suspended ceilings collapse or, at very high levels of building damage, when beams and columns begin to fall on them. In the AIR model, then, contents damageability is a function of both occupancy class and the building damage ratio.

**INSURED LOSS
MODULE**

In the last step of the model, insured losses are calculated by applying the policy conditions including coverage limits and deductibles for the various lines of business and classes of coverage to the total damage estimates resulting from the damage estimation module. A range of policy conditions can also be used.

**MODELED LOSS
COST OUTPUT**

The model calculates a loss cost (average annual loss per \$100 of value) by summing the expected annual losses from all modeled earthquakes that impact specific locations. For each type of coverage (buildings, contents), the model produces output in ZIP code/construction/deductible detail.

SELECTION OF TERRITORIES

1. INTRODUCTION

The purpose of varying loss costs by territory is to reflect the different risk characteristics of different geographic areas. Risks located in areas with higher seismic activity or weak soil types should be charged more than risks located in less seismic areas.

Since mean damage ratios (MDRs) are produced by the model in ZIP code detail, it is possible to vary loss costs by ZIP code. ZIP codes were aggregated into territories for the following reasons:

- The earthquake model produces an estimate of the MDR, but there is variation around the mean. Combining ZIP codes with similar MDRs into territories is one way of smoothing.
- Ease of implementation by insurers. It is easier to have only a few loss costs for a state as opposed to one loss cost for each ZIP code.

Since seismicity and soil type can be very local phenomena, the earthquake risk can vary greatly between neighboring ZIP codes. Therefore, it was decided that territories need not consist of contiguous ZIP codes. Rather, each territory was created as a collection of ZIP codes that have a similar loss potential.

The following sections describe:

- Clustering analysis that group ZIP codes into territories.
- Determination of final territories.

2. CLUSTERING ZIP CODES INTO TERRITORIES BY STATE

ZIP codes in each state were combined into territories using a k-means clustering method with the most recent year's ISO Commercial Property exposure by ZIP code, which represents the amount of exposure in each ZIP code, as weights (see appendix A for k-means clustering). In order to cluster the ZIP codes into territories, it was necessary to calculate exactly one clustering MDR for each ZIP code to represent the ZIP code's loss potential. The Commercial Lines clustering MDRs for each ZIP code were based on the construction class that was found in the greatest number of ZIP codes.

3. SELECTING THE NUMBER OF TERRITORIES

In order to select the number of territories for a state, consideration was given to the following:

- Homogeneity index (see appendix A.3 for homogeneity index).
- The number of ZIP codes in each territory - Any territory that had too few ZIP codes was combined into another territory with a similar MDR.
- Differences in clustering MDRs between territories - Territories with small absolute differences in MDRs were combined into a single territory.

4. INTRODUCTION OF SUB-TERRITORIES

By itself, the revision of ZIP code-based territories will cause the loss costs of most insureds to either decrease or increase. To mitigate the impact of this revision where such loss cost swings, in combination

with the overall loss cost changes by territory, are large and positive, this filing introduces intermediate sub-territories.

The sub-territories were created based on an evaluation of the combined effect of the territory definition revisions and the loss cost level changes in this filing. If, after tempering, the estimated commercial property impact of the Earthquake loss cost change for a revised territory/current territory combination would be greater than +100%, the ZIP codes in that territory combination were assigned to a sub-territory. The sub-territories are distinguished by an "A" following the territory number.

Personal Property Loss Cost Revision

INTRODUCTION

ISO provides Commercial Earthquake personal property loss costs for four damageability grades. The four damageability grades are:

- Personal property grade 1 with exceptionally high susceptibility
- Personal property grade 2 with high susceptibility
- Personal property grade 3 with moderate susceptibility
- Personal property grade 4 with slight susceptibility

Personal property loss costs vary by territory and building class.

LOSS COST DEVELOPMENT

Occupancy class is used to presume the types of personal property (for example, carpets at a carpets manufacturer, or bottles and glasses at a tavern) at a given location and is therefore used to calculate the damageability of personal property for a given insured.

Loss costs of forty AIR occupancy classes, including the base occupancy of General Commercial, are examined to analyze magnitude and volatility of loss cost for each occupancy class. Occupancy ratios, which is the personal property loss cost of each occupancy class over the personal property loss cost of the base occupancy class, were calculated for each predetermined territory (see B-9 for Selection of Territories) and building class. Thirty occupancy classes showed stable and consistent occupancy ratios across building classes and territories with marginal differences in the ratios ranging from .85 to 1.15. These occupancy classes were selected to develop loss costs for personal property grade 2 and 4. The occupancy ratios of nine other occupancy classes are high and vary significantly by building class and territory. These nine occupancy classes, which are characterized as having personal property with exceptionally high susceptibility, are used to develop loss costs for personal property grade 1.

Development of Personal Property Grade 3 Loss Costs

Loss costs for personal property grade 3 are developed as the exposure-weighted average of personal property loss costs across all ZIP codes in a given territory for each building class, with occupancy class of General Commercial that is the basis of building loss cost and territorial selection (see B-9 for Selection of Territories).

Personal property grade 3 loss costs are calculated by the following formula:

i = subscript for territory

j = subscript for ZIP code

b = subscript for building class

w_{ij} = exposure for *j* ZIP code in *i* territory

$LC_{jb}^{(3)}$ = personal property grade 3 loss cost for b building class in j ZIP code

$LC_{ib}^{(3)}$ = personal property 3 loss cost for b building class in i territory

$$LC_{ib}^{(3)} = \frac{\sum w_{ij} \times LC_{jb}^{(3)}}{\sum w_{ij}}$$

Development of Personal Property Grade 2 and 4 Loss Costs

Personal property loss costs for grade 2 and 4 are based on the AIR model output for thirty occupancy classes that exhibit stable occupancy ratios. Since these occupancy ratios are consistent across all building classes and territories, building class D1 was selected to run the analysis. Average occupancy ratios of building class D1 over all territories were calculated for each of the thirty occupancy classes. Using K-means clustering methodology (see appendix A for k-means clustering), the thirty occupancy classes were divided into three groups by average occupancy ratios. Three occupancy relativities were calculated by averaging occupancy ratios within each group. Since the middle occupancy relativity is close to 1, the occupancy relativities for personal property grade 2 and 4 are those for the higher group and the lower groups, respectively. Personal property loss costs for personal property grades 2 and 4 were calculated by multiplying the personal property loss costs of personal property grade 3 by the selected relativities for personal property grade 2 and 4.

Below are the detailed steps taken to develop personal property grade 2 and 4 loss costs:

Step 1. Calculate exposure weighted personal property loss costs with building class of D1 for each territory(i) and occupancy class(k).

k = superscript for high-moderate-low susceptible occupancy class

$LC_{jD1}^{(k)}$ = loss cost of k occupancy class for D1 building class in j ZIP code

$LC_{iD1}^{(k)}$ = loss cost of k occupancy class for D1 building class in i territory

$$LC_{iD1}^{(k)} = \frac{\sum w_{ij} \times LC_{jD1}^{(k)}}{\sum w_{ij}}$$

Step 2. Calculate occupancy ratios by territory and occupancy class.

$R_{iD1}^{(k)}$ = occupancy ratio of occupancy class k for D1 building class in i territory

$$R_{iD1}^{(k)} = \frac{LC_{iD1}^{(k)}}{LC_{iD1}^{(3)}}$$

Step 3. Calculate average occupancy ratios.

$R_{D1}^{(k)}$ = average occupancy ratio of occupancy class k for D1 building class

$$R_{D1}^{(k)} = \text{Average}(R_{iD1}^{(k)})$$

Step 4. Run K-mean clustering analysis. The output of the analysis are occupancy groups with similar average occupancy ratios.

$R_{OG1}^{(k)}$ = occupancy ratios of occupancies that are assigned to occupancy group 1 (OG1)

$R_{OG2}^{(k)}$ = occupancy ratios of occupancies that are assigned to occupancy group 2 (OG2)

$R_{OG3}^{(k)}$ = occupancy ratios of occupancies that are assigned to occupancy group 3 (OG3)

Step 5. Calculate occupancy relativities for personal property grade 2 and 4.

R_{OG1} = occupancy relativity for occupancy group 1 (OG1)

R_{OG2} = occupancy relativity for occupancy group 2 (OG2)

R_{OG3} = occupancy relativity for occupancy group 3 (OG3)

$$R_{OG1} = \text{Average}(R_{OG1}^{(k)})$$

$$R_{OG3} = \text{Average}(R_{OG3}^{(k)})$$

$$\text{where } R_{OG1} > R_{OG2} \approx 1.00 > R_{OG3}$$

Step 6. Develop loss costs for personal property grade 2 and 4 by territory and building class.

$$LC_{ib}^{(2)} = R_{OG1} \times LC_{ib}^{(3)}$$

$$LC_{ib}^{(4)} = R_{OG3} \times LC_{ib}^{(3)}$$

Development of Personal Property Grade 1 Loss Costs

Personal property loss costs for personal property grade 1 are based on the AIR model output for nine occupancy classes whose personal property are highly susceptible to earthquake such as Wineries and Stone/Clay/Glass/Ceramics products. Since these personal properties are known to be more fragile than any of those in personal property grades 2 – 4, and the personal property loss costs of these occupancy classes showed significant volatility, a selection was made so that in all instances these classes are considered more susceptible to earthquake than the general commercial class. Average occupancy ratios for the nine occupancy classes were calculated for each territory and building class. Because the patterns of average occupancy ratios for building classes are consistent directionally across territories, but with varying magnitude, a multi-dimensional K-means (see appendix A for k-means clustering) clustering method based on the calculated average occupancy ratios by territory for all building classes was performed to group the territories (personal property territory group). Six personal property territory groups were selected to achieve over 96% explanation of the variance based on the homogeneity index (see appendix A.3 for an explanation of the homogeneity index). Within each personal property territory group, the occupancy relativities were computed by averaging occupancy ratios for each building class. For any given territory, personal property grade 1 loss costs were calculated by multiplying personal property loss costs of personal property grade 3 by the occupancy relativities in the corresponding personal property territory group and building class.

Below are the detailed steps taken to develop Personal property grade 1 loss costs:

Step 1. Calculate exposure weighted personal property loss costs for each territory(*i*), occupancy class(*m*) and building class(*b*).

m = superscript for exceptionally high susceptible occupancy class

t = superscript for territory group

$LC_{jb}^{(m)}$ = loss cost of *m* occupancy class for *b* building class in *j* ZIP code

$LC_{ib}^{(m)}$ = loss cost of *m* occupancy class for *b* building class in *i* territory

$$LC_{ib}^{(m)} = \frac{\sum w_{ij} \times LC_{jb}^{(m)}}{\sum w_{ij}}$$

Step 2. Calculate occupancy ratios by territory, occupancy class and building class.

$R_{ib}^{(m)}$ = occupancy ratio for *m* occupancy class and *b* building class in *i* territory

$$R_{ib}^{(m)} = \begin{cases} \frac{LC_{ib}^{(m)}}{LC_{ib}^{(3)}} & \text{for } LC_{ib}^{(m)} > LC_{ib}^{(3)} \\ 0 & \text{for } LC_{ib}^{(m)} \leq LC_{ib}^{(3)} \end{cases}$$

Step 3. Calculate average occupancy ratios by territory and building class.

R_{ib} = average occupancy ratio for *b* building class in *i* territory

$$R_{ib} = \text{Average}(R_{ib}^{(m)})$$

Step 4. Perform multi-dimensional K-mean clustering analysis. The output of the analysis are personal property territory groups with similar occupancy relativities.

Input : RB_i = a vector of occupancy ratios for all building classes in *i* territory

$$= \{RB_{iA1}, RB_{iB1}, RB_{iC1}, RB_{iD1}, RB_{iD2}, RB_{iD3}, RB_{iE1}, RB_{iE2}, RB_{iE3}\}$$

Output : $RB_i^{(t)}$ = a vector of occupancy ratios for all building classes in *i* territory that is assigned to personal property territory group *t*

Step 5. Calculate occupancy relativity for each building class for personal property territory group *t*

RB^t = a vector of occupancy ratios for each building classes in personal property territory group *t*

R_b^t = a occupancy relativity for *b* building class in personal property territory group *t*

$$RB^t = \text{Average}(RB_i^{(t)}) = \{R_b^t\} = \{R_{A1}^t, R_{B1}^t, R_{C1}^t, \dots, R_{E3}^t\}$$

Step 6. Develop loss costs for personal property grade 1 by territory and building class.

$$LC_{ib}^{(1)} = R_b^t \times LC_{ib}^{(3)}$$

FURTHER ADJUSTMENTS TO THE EARTHQUAKE MODEL OUTPUT

INTRODUCTION	The AIR model provides loss costs (expected earthquake losses per \$100 of replacement cost) by construction class at different deductible levels for each ZIP code. Some additional adjustments have been applied to the model output to calculate revised earthquake loss costs.
PERSONAL PROPERTY LOSS COSTS	ISO provides Commercial Earthquake loss costs for four personal property damageability grades. The grades are assigned based on occupancy and/or type of personal property with grade 4 representing a slight susceptibility, grade 3 a moderate susceptibility, grade 2 a high susceptibility, and grade 1 an exceptionally high susceptibility to earthquake damage. See B10-B13 for Personal Property Loss Cost Revision.
COINSURANCE PROVISION	<p>Since the modeled loss costs relate average loss to the total property value, an adjustment for the 80% coinsurance provision has to be made. The filed loss costs are for 80% of the property value, so the base deductible is calculated based on that 80%. For example, if the base deductible is 5%, then the deductible is in fact 4% ($= 5\% * 80\%$) of the total property value. From that point, losses up to 80% of the property value are covered. So for the 5% base deductible example, the layer of coverage is between 4% and 84% of the property value.</p> <p>Loss Cost (LC) at the base deductible and 80% coinsurance = $(\text{LC at the } (0.8 * \text{base}) \text{ deductible} - \text{LC at } ((0.8 * \text{base}) + 80\%) \text{ deductible}) / 0.80.$</p> <p>This calculation was done for each ZIP code, construction, and coverage combination.</p>
LOSS ADJUSTMENT EXPENSE FACTOR	The loss adjustment expense factor of 1.19 was applied to the model output since the model does not account for loss adjustment expense. The data underlying the selection of this factor is shown on Table 5.
AGGREGATION INTO TERRITORIES	The ZIP code loss costs were weighted together to develop territory loss costs. Commercial Property exposure information by ZIP code was used to determine these weights.
CREDIBILITY OF MODEL OUTPUT	In this filing, full credibility has been assigned to the output of the AIR model.
MINIMUM LOSS COST	A minimum loss cost of .001 (per \$100 of value) was used in place of the calculated loss cost in each ZIP code where the indicated loss cost was below the minimum.

TABLE 5

Development of Loss Adjustment Expense Factors

The following table shows a compilation of 20 years of Earthquake Loss adjustment expense experience based on Insurance Expense Exhibit information from A. M. Best's.

Year	Net Incurred Losses	Net Incurred LAE	LAE/ Losses
2001	411,033,000	96,478,000	23.5%
2002	294,643,000	100,127,000	34.0%
2003	254,026,000	69,630,000	27.4%
2004	222,435,000	26,594,000	12.0%
2005	211,685,000	27,269,000	12.9%
2006	14,989,000	11,945,000	79.7%
2007	(38,668,000)	(1,370,000)	3.5%
2008	17,460,000	5,140,000	29.4%
2009	76,266,000	21,488,000	28.2%
2010	198,058,000	23,938,000	12.1%
2011	387,426,000	21,537,000	5.6%
2012	84,378,000	12,216,000	14.5%
2013	(4,680,000)	12,363,000	-264.2%
2014	66,078,000	11,986,000	18.1%
2015	(8,217,000)	7,408,000	-90.2%
2016	40,157,000	7,612,000	19.0%
2017	111,178,000	20,600,000	18.5%
2018	201,166,000	22,637,000	11.3%
2019	(21,141,000)	6,397,000	-30.3%
2020	173,985,000	4,099,000	2.4%
Total	2,692,257,000	508,094,000	18.9%
Selected			19.0%

NEW MEXICO

COMMERCIAL EARTHQUAKE INSURANCE

SECTION C

Commercial Earthquake Loss Costs	C2-3
--	------

73. CAUSES OF LOSS – EARTHQUAKE FORM

E.1.a. Earthquake Loss Costs (Subline Code 930)

Bldg. Class	Base Deduct.	Territory 1 Loss Costs					Territory 2 Loss Costs				
		Bldg.	Personal Property Grade				Bldg.	Personal Property Grade			
			1*	2*	3*	4*		1*	2*	3*	4*
A1	5%	0.002	0.003	0.001	0.001	0.001	0.013	0.011	0.008	0.007	0.007
B1	5	0.002	0.003	0.001	0.001	0.001	0.010	0.010	0.007	0.006	0.006
C1	5	0.002	0.003	0.001	0.001	0.001	0.012	0.010	0.008	0.008	0.007
D1	5	0.002	0.003	0.002	0.002	0.001	0.014	0.011	0.010	0.009	0.008
D2	5	0.002	0.003	0.002	0.002	0.001	0.014	0.012	0.010	0.009	0.008
D3	5	0.003	0.003	0.002	0.002	0.002	0.016	0.013	0.011	0.010	0.009
E1	5	0.002	0.003	0.002	0.001	0.001	0.013	0.011	0.009	0.008	0.008
E2	5	0.004	0.003	0.003	0.002	0.002	0.024	0.018	0.016	0.015	0.013
E3	5	0.004	0.004	0.003	0.003	0.002	0.027	0.020	0.018	0.016	0.015

Bldg. Class	Base Deduct.	Territory 3 Loss Costs					Territory 3A Loss Costs				
		Bldg.	Personal Property Grade				Bldg.	Personal Property Grade			
			1*	2*	3*	4*		1*	2*	3*	4*
A1	5%	0.022	0.041	0.012	0.011	0.010	0.012	0.023	0.006	0.006	0.005
B1	5	0.018	0.042	0.012	0.011	0.010	0.010	0.024	0.006	0.006	0.005
C1	5	0.021	0.042	0.014	0.013	0.011	0.012	0.024	0.008	0.007	0.006
D1	5	0.025	0.044	0.016	0.015	0.013	0.014	0.025	0.009	0.008	0.007
D2	5	0.025	0.044	0.016	0.015	0.013	0.014	0.025	0.009	0.008	0.007
D3	5	0.029	0.042	0.019	0.017	0.015	0.016	0.024	0.010	0.009	0.008
E1	5	0.023	0.043	0.015	0.014	0.012	0.013	0.024	0.008	0.008	0.006
E2	5	0.039	0.043	0.025	0.023	0.020	0.022	0.024	0.014	0.013	0.011
E3	5	0.043	0.044	0.027	0.025	0.022	0.024	0.025	0.015	0.014	0.012

E.1.a.(1) Class-Rated Risks – Earthquake Loss Costs

Bldg. Class	Mand. Deduct.	Territory 21 Loss Costs					Territory 22 Loss Costs				
		Bldg.	Contents Grade				Bldg.	Contents Grade			
			1*	2*	3*	4*		1*	2*	3*	4*
1C	5%	0.022	0.166	0.049	0.022	0.016	0.014	0.101	0.033	0.014	0.010
1D	5%	0.022	0.136	0.039	0.018	0.013	0.014	0.095	0.030	0.013	0.010
2A	5%	0.014	0.142	0.042	0.019	0.014	0.009	0.087	0.028	0.012	0.008
2B	5%	0.014	0.117	0.033	0.015	0.011	0.009	0.082	0.026	0.011	0.008
3A	5%	0.022	0.124	0.037	0.017	0.012	0.014	0.081	0.026	0.011	0.008
3B	5%	0.019	0.100	0.029	0.014	0.009	0.012	0.069	0.022	0.011	0.007
3C	10%	0.017	0.057	0.021	0.010	0.005	0.011	0.035	0.015	0.006	0.004
4A	5%	0.018	0.100	0.030	0.014	0.009	0.012	0.066	0.021	0.010	0.006
4B	5%	0.023	0.114	0.034	0.015	0.011	0.014	0.078	0.025	0.011	0.008
4C	10%	0.026	0.082	0.031	0.016	0.009	0.016	0.051	0.021	0.010	0.005
4D	10%	0.028	0.092	0.035	0.017	0.010	0.019	0.058	0.025	0.011	0.006
5A	5%	0.038	0.164	0.048	0.023	0.016	0.023	0.112	0.036	0.016	0.012
5AA	10%	0.027	0.084	0.030	0.017	0.009	0.017	0.051	0.020	0.010	0.005
5B	10%	0.051	0.127	0.061	0.036	0.017	0.032	0.103	0.046	0.027	0.012
5C	10%	0.074	0.188	0.108	0.061	0.026	0.046	0.143	0.076	0.041	0.017

* Contents Grade—See Rule 73.F.

Bldg. Class	Mand. Deduct.	Territory 23 Loss Costs					Territory 24 Loss Costs				
		Bldg.	Contents Grade				Bldg.	Contents Grade			
			1*	2*	3*	4*		1*	2*	3*	4*
1C	5%	0.013	0.095	0.031	0.014	0.009	0.006	0.046	0.015	0.007	0.004
1D	5%	0.013	0.092	0.030	0.012	0.009	0.006	0.045	0.014	0.006	0.004
2A	5%	0.007	0.076	0.025	0.011	0.007	0.003	0.034	0.011	0.005	0.004
2B	5%	0.007	0.073	0.024	0.010	0.007	0.003	0.033	0.011	0.005	0.004
3A	5%	0.011	0.073	0.024	0.010	0.008	0.004	0.032	0.010	0.005	0.004
3B	5%	0.010	0.062	0.020	0.010	0.006	0.003	0.026	0.009	0.004	0.003
3C	10%	0.009	0.032	0.013	0.006	0.004	0.003	0.014	0.006	0.002	0.002
4A	5%	0.010	0.061	0.020	0.009	0.006	0.004	0.028	0.009	0.004	0.003
4B	5%	0.012	0.071	0.023	0.010	0.007	0.004	0.029	0.009	0.004	0.003
4C	10%	0.014	0.048	0.019	0.009	0.004	0.005	0.021	0.008	0.004	0.002
4D	10%	0.015	0.052	0.022	0.009	0.005	0.005	0.021	0.009	0.004	0.002
5A	5%	0.022	0.111	0.036	0.016	0.012	0.010	0.059	0.019	0.008	0.006
5AA	10%	0.016	0.051	0.020	0.009	0.005	0.007	0.026	0.010	0.005	0.002
5B	10%	0.030	0.103	0.045	0.026	0.012	0.012	0.054	0.022	0.012	0.006
5C	10%	0.041	0.142	0.071	0.037	0.016	0.017	0.071	0.032	0.017	0.008
* Contents Grade — See Rule 73.F.											

These territories are assigned to deductible tier as follows:

Territories: 21 — 24	Tier: 3
Territory: 1	Tier: 1
Territories: 2, 3, 3A	Tier: 2

These territories are assigned to height territory group as follows:

Territories: 1, 2, 3, 3A	Group: 1
--------------------------	----------

* Personal Property Rate Grade — See Rule 73.D.5.

Appendix A

K-means Clustering

1. DESCRIPTION

K-means clustering aims to group similar data points into clusters and simultaneously maximize the difference between clusters. This method begins with the initial selection of a cluster centroid for each cluster. The distance of each data point from the centroid is calculated and data points are assigned to the closest cluster based on the distance assigned. After the first iteration, a new centroid is calculated for each cluster by taking the straight/weighted average of all data points within the cluster. This process is repeated until all data points converge and the centroids stop moving.

2. DATA USED IN CLUSTERING

The k-means clustering algorithm is both flexible and robust regarding the input data used. The data used can range from single-valued data points to multi-valued data points (e.g. a curve). In the case where multi-valued data points/vector of values/a curve is used, a multi-dimensional k-means approach is utilized, where each value of the vector/curve can be considered as co-ordinates of a data point on a multi-dimensional plane.

3. SELECTING THE NUMBER OF CLUSTERS - HOMOGENEITY INDEX

The clustering process does not determine the optimal or appropriate number of clusters. In order to select the number of clusters, the homogeneity index is considered (along with various other factors which vary based on the data that is being clustered).

The homogeneity index is a measure of how much of the total variance of the data is explained by the clustering system. In order to calculate the homogeneity index, it is necessary to calculate the following:

- Overall within cluster variation - A measure of the average, over all clusters, of each cluster's within variation, where each clusters' within variation is the average, over all data points in the cluster, of the squared deviation of each data point from the average of the cluster.
- Between cluster variation - A measure of the average, over all clusters, of the squared deviation of each cluster's average from the average of all data points i.e. the total average.
- Total variation - Sum of the overall within cluster variation and the between cluster variation.

For each clustering scheme, the homogeneity index is calculated as the ratio of the between cluster variation of the clustering scheme to the total variation.

Regardless of the number of clusters, the total variation remains unchanged. As the number of clusters decrease, the between cluster variation decreases, and, therefore, the homogeneity index decreases. Since it is desired to produce a relatively high homogeneity index (for example, 90%) in an efficient manner (meaning using a small number of clusters), selecting the optimal number of clusters involves examining the homogeneity index at each iteration of the clustering process. After each iteration of the clustering process, the number of clusters decrease by one. Candidates for the desired number of clusters (n) were selected such that the clustering iteration that produced n-1 clusters from n clusters resulted in a large decrease in the homogeneity index as compared to the decrease in the homogeneity index from the previous clustering iteration that produced n clusters from n+1 clusters.

For example, suppose that several iterations of the clustering procedure resulted in the following homogeneity indices:

Number of Clusters	Within Cluster Variation	Between Cluster Variation	Total Variation	Homogeneity Index
7	7	93	100	0.93
6	8	92	100	0.92
5	9	91	100	0.91
4	20	80	100	0.80

In this example, as the clustering process moved from five clusters to four clusters, there was a large decrease in the homogeneity index from 0.91 to 0.80. However, the previous few iterations of the clustering process produced decreases of only 0.01 in the homogeneity index. Therefore, in this example, "five" would be selected as a candidate for the desired number of clusters.