The Impact of Catalog Size on Losses in AIR'S Hurricane and Earthquake Models for the United States

AIR WORLDWIDE

Trademarks

AIR Worldwide, CATRADER, and Touchstone are registered trademarks of AIR Worldwide Corporation.

Confidentiality

AIR invests substantial resources in the development of its models, modeling methodologies and databases. This document contains proprietary and confidential information and is intended for the exclusive use of AIR clients who are subject to the restrictions of the confidentiality provisions set forth in license and other nondisclosure agreements.

Contact Information

AIR is headquartered in Boston with additional offices in North America, Europe, and Asia. For information on our office locations, visit **http://www.air-worldwide.com/About-AIR/Offices/**.

AIR welcomes feedback on its documentation. To contact us regarding this document please email us at **ModelDocumentationFeedback@air-worldwide.com**.

Revision History

February 2017	Original document release.
October 2017	In Section 3, updated list of models that have multiple catalogs of different sizes.

Table of Contents

List of Figures

List of Tables

1 Catalog Optimization

Within catastrophe modeling there are many sources of uncertainty impacting loss estimates. These sources of uncertainty are described at length in other documents (e.g., the *Touchstone Financial Module*). The source of uncertainty discussed here is called *sampling variability* or *process risk*. In this context, such uncertainty is the source of differences between the 10,000 year EP curve from a 10,000 year catalog and the 10,000 year EP curve from a larger catalog (e.g., 50,000 years). The uncertainty therefore shows up as the variability in losses at a given exceedance probability brought about by the size of the catalog. In theory this source of uncertainty is reduced if a larger catalog is used. This document serves to discuss the creation of the 10,000 year catalog for AIR's Hurricane and Earthquake Models for the United States, and illustrate the *difference in losses* between the 10,000 year and 50,000 year catalogs as obtained from CATRADER.

In AIR, sampling variability is considered in the construction of the 10,000 year catalog, and significant attempts have been made to reduce the resulting sampling variability through a process known as *catalog optimization*. The 10,000 year AIR U.S. Hurricane and U.S. Earthquake catalogs are each an optimized sample of a much larger catalog.

AIR uses constrained Monte Carlo simulation to generate its stochastic catalogs. For example, the U.S. Hurricane stochastic catalog generation process produces a large number of simulated years of hurricane activity, with each simulated year representing a possible scenario of next year's hurricane activity. A 10,000 year catalog is generated, but larger 50,000 year and 100,000 year catalogs are also available.

1.1 U.S. Hurricane

Acknowledging that 100,000 year or even 50,000 year catalogs may pose unacceptable computational demands when multiple runs are needed, AIR delivers a standard catalog of 10,000 years primarily for performance and storage reasons. In order to expedite convergence and reduce sampling variability, constraints are implemented in the catalog generation process. The goal is to reproduce the theoretical frequency and intensity distributions by 100-mile coastal segment in the different (10,000, 50,000, or 100,000 year) catalogs. The implementation consists of three stages: 1) placing caps on the number of storms that are drawn for each coastal segment, 2) placing caps on the number of storms that are drawn for each Saffir Simpson (SS) category, and 3) convergence tests on losses.

The caps are generated by running the AIR U.S. Hurricane model's landfall frequency and intensity modules for one million years with no caps imposed and scaling the resulting frequencies down to a 10,000 year simulation to produce expected frequencies for each SS category in each coastal segment. To be clear, the 1 million year simulation is done to generate the landfall frequency and intensity statistics, but 1 million years of events are not simulated.

The annual frequency for each year in each 1,000 year segment of the catalog is constrained to be within a small range around the expected value. The annual frequency for each of these 1,000 year periods is drawn from a negative binomial distribution. If the average frequency for that set of years does not fall within the expected range, then the frequency for all 1,000 years is re-sampled from the distribution. In cases where the target frequencies are exceeded, the intensity and/or landfall segment are re-sampled. The caps are implemented in successive 10,000 year blocks in the simulation (i.e., years 1-10,000, 10,001-20,000, 20,001-30,000, etc.) in the generation of the 50,000 and 100,000 year catalogs.

Finally, in addition to the convergence tests on frequencies just described, convergence tests on losses are also performed as a form of validation to ensure consistency between the 10,000 year and larger catalogs. Further information can be found in Section 1.4 of the *AIR Hurricane Model for the United States* model document.

1.2 U.S. Earthquake

In the AIR U.S. Earthquake model, a 100,000 year catalog is first obtained from a 1 million year catalog. This is done by solving a multi-criteria optimization problem in which three aspects of the catalog are evaluated: 1) the magnitude-rate distributions for each region/zone, 2) ground motion at specific locations and at specific return periods, and 3) loss distributions. The result is a 100,000 year catalog that performs similarly to the 1 million year catalog, not only with respect to the magnitude-rate distributions but also with respect to the ground motion and loss distributions. Acknowledging that even 100,000 year catalogs may pose unacceptable computational demands when multiple runs are needed, AIR delivers a standard catalog of 10,000 years. This smaller 10,000 year catalog is obtained by selecting years from the 100,000 year catalog iteratively while monitoring the difference at state level between the 10,000 year EP curve produced by the 100,000 year catalog and the 10,000 year EP curve produced by the 10,000 year catalog. Losses at state level and for special interest zones such as the New Madrid seismic zone (NMSZ) are also included in the optimization to ensure that the final 10,000 year results are comparable to the 100,000 year results, even in low seismicity areas such as the NMSZ. This is shown graphically in Figure 1 below.

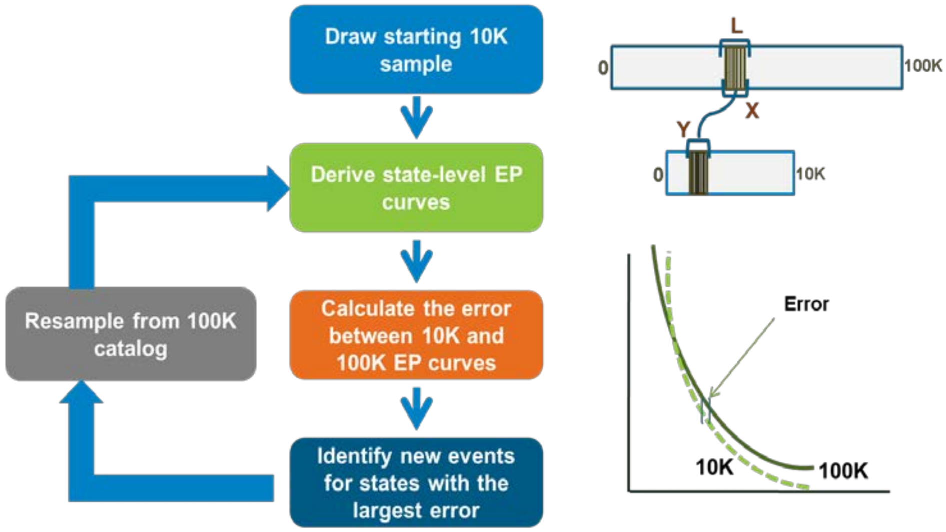

Figure 1. Figure showing schematically how the 10,000-year optimized catalog is obtained through a *constrained sampling* process from the 100,000-year catalog

The resulting optimized catalogs attempt to minimize the differences between the 10,000-, 50,000-, and 100,000-year catalogs. Therefore, the results below show a much smaller difference between the 10,000 and 50,000 year catalogs than if the 10,000 year catalog was selected randomly from the 100,000 year catalog. Also, for ease of use of event output, the 10,000 year catalog is the first 10,000 years of the 100,000 year catalog, and the 50,000 year catalog is the first 50,000 years of the 100,000 year catalog. This allows the three catalogs to be easily compared against each other using losses from the 100,000 year catalog only. Further information can be found in Section 1.6 of the AIR Earthquake Model for the United States model document.

2 Assumptions

We will now illustrate the benefits of catalog optimization through comparing the difference in losses between the 10,000- and 50,000-year catalogs at several return periods. These runs were performed for the hurricane peril; using the North Atlantic Standard Tropical Cyclone (Standard) and North Atlantic Warm Sea Surface Temperature Tropical Cyclone (WSST) catalogs, the U.S. Industry Exposure Database, and 2016 Indexed Industry Take-Up Rates in CATRADER 18.0, with demand surge and storm surge damage included. For the earthquake peril, the U.S. Time Dependent and Time Independent earthquake models were used in CATRADER 18.0 with 2016 Indexed Industry Take-Up Rates and demand surge turned on.

The percentage differences were computed on an *aggregate* loss basis after assumed industry deductibles and limits were applied. Please note that the differences on an occurrence basis are similar to the results shown here, being in some cases a few percentage points different from the aggregate loss differences.

3 Results

The 10,000- and 50,000-year catalog results, presented in Table 1 are within 8% when run on an industry basis for both the hurricane and earthquake perils at the 100 to 200 year return periods. This is shown for the U.S. as a whole as well as for the key earthquake state of California and the key hurricane state of Florida. This good agreement is expected due to the constrained sampling process highlighted above. Nevertheless, as expected, larger differences are seen at the low frequency, high return periods (e.g., 10,000 years).

Table 1. Comparison by peril between the 10,000- and 50,000-year catalogs U.S.-wide and for a single key loss causing state for the earthquake and hurricane perils

Aggregate	(50k-10k)/10k							
	Earthquake (Time Dep)		Earthquake (Time Indep)		Hurricane (Standard)		Hurricane (WSST)	
Return Period	**US-Wide**	**California**	**US-Wide**	**California**	**US-Wide**	**Florida**	**US-Wide**	**Florida**
AAL	4%	1%	3%	1%	1%	1%	1%	1%
20	2%	1%	2%	-2%	-1%	-3%	1%	-1%
30	4%	2%	0%	-4%	0%	1%	2%	0%
50	2%	3%	-1%	1%	-2%	2%	0%	3%
100	0%	2%	6%	-3%	3%	2%	2%	-3%
200	8%	0%	7%	3%	1%	0%	-1%	0%
250	14%	4%	12%	5%	1%	2%	-2%	-3%
500	4%	3%	3%	4%	2%	0%	1%	0%
1000	-8%	13%	-6%	-2%	21%	5%	7%	11%
10000	2%	0%	4%	5%	45%	38%	52%	56%

Intuitively, one would expect that the differences between the 10,000- and 50,000-year catalogs would be greater for regions which have lower activity compared with those which have higher activity. We could imagine that the number of large loss causing years is low in regions of low activity and hence the high return periods of the resulting EP curve have a larger amount of scatter compared to regions of high activity. This leads to greater uncertainty about the resulting EP curve for less active regions. Taking a less active region for the earthquake peril such as New York, as shown in Table 2 below, we can see at the 100 to 250 year return periods the difference in losses between the 10,000 and 50,000 year catalogs is up to 45%, larger than what is seen for California in Table 1 above. (Note that there are no losses up to the 50 year return period for New York.)

Table 2. Loss difference between the 10,000- and 50,000-year earthquake catalogs for New York

Aggregate	(50K-10K)/10K
	Earthquake
Return Period	New York
100	44%
200	15%
250	16%
500	-19%
1,000	12%
10,000	37%

Looking now at hurricanes in Table 3, we see good agreement (as in Table 1 above) with both U.S.-wide and other loss causing regions for the 10,000 and 50,000 year catalogs. The differences do not exceed 5% at the 100 to 250 year return periods. As expected there is more volatility when comparing the higher return periods (e.g., 10,000 year) between the two catalogs; this is due to the small number of large loss causing events and the lower likelihood of convergence as a result.

Table 3. Comparison for hurricane between the 10,000- and 50,000-year Standard and Warm Sea Surface Temperature catalogs U.S.-wide and for other key loss causing regions

Aggregate	(50K-10K)/10K							
	Hurricane (Standard)				Hurricane (WSST)			
Return Period	U.S.	Gulf States	Northeast	Florida	U.S.	Gulf States	Northeast	Florida
AAL	1%	1%	2%	1%	1%	1%	1%	1%
20	-1%	2%	-1%	-3%	1%	2%	-3%	-1%
30	0%	1%	-5%	1%	2%	1%	-4%	0%
50	-2%	1%	0%	2%	0%	1%	-2%	3%
100	3%	2%	2%	2%	2%	3%	2%	-3%
200	1%	-5%	0%	0%	-1%	-1%	-1%	0%
250	1%	-1%	2%	2%	-2%	1%	1%	-3%
500	2%	-2%	3%	0%	1%	1%	3%	0%
1,000	21%	2%	0%	5%	7%	5%	0%	11%
10,000	45%	31%	27%	38%	52%	31%	28%	56%

As with the earthquake peril, for states with low hurricane frequency such as Massachusetts and New York, there are larger differences in loss results between the 10,000 and 50,000 year catalogs, compared with Florida. This is seen below where the 10,000 and 50,000 year catalogs produce very similar losses for Florida but less similar results for Massachusetts and New York.

Table 4. Comparison for hurricane between the 10,000- and 50,000-year catalogs for Florida, Massachusetts, and New York

Aggregate	(50,000-10K)/10K					
	Hurricane (Standard)			Hurricane (WSST)		
Return Period	Florida	Massachusetts	New York	Florida	Massachusetts	New York
AAL	1%	-4%	4%	1%	-4%	3%
SD	7%	-2%	17%	6%	-2%	17%
20	-3%	6%	-12%	-1%	9%	-13%
30	1%	2%	-3%	0%	2%	-6%
50	2%	-4%	-2%	3%	-3%	-1%
100	2%	-4%	-2%	-3%	-4%	-1%
200	0%	-12%	4%	0%	-12%	3%
250	2%	-8%	4%	-3%	-8%	-3%
500	0%	2%	-1%	0%	2%	-1%
1,000	5%	-8%	15%	11%	-8%	16%
10,000	38%	0%	22%	56%	0%	22%

Finally, it should be noted that the loss results given here are based on model runs with the U.S. Industry Exposure Database, and the differences between the 10,000 and 50,000 year catalogs can be much larger for a given portfolio. Portfolios with few locations and in regions with low activity (hurricane or earthquake) will be expected to produce larger deviations than industry-wide portfolios. The 50,000 year catalogs for U.S. Earthquake, North Atlantic Standard Tropical Cyclone, and North Atlantic Warm Sea Surface Temperature (WSST) Tropical Cyclone catalogs are separately licensable in CATRADER; there are also 50,000- and 100,000-year catalogs for these perils in Touchstone. Furthermore, there are 100,000-year catalogs for AIR's Typhoon and Earthquake Models for Japan and AIR's Bushfire, Earthquake and Severe Thunderstorm models for Australia in Touchstone. Please contact your AIR representative if you would like to investigate the impact of catalog size on your own portfolio.

4 About AIR Worldwide

AIR Worldwide (AIR) provides catastrophe risk modeling solutions that make individuals, businesses, and society more resilient. AIR founded the catastrophe modeling industry in 1987, and today models the risk from natural catastrophes, terrorism, cyber attacks, and pandemics globally. Insurance, reinsurance, financial, corporate, and government clients rely on AIR's advanced science, software, and consulting services for catastrophe risk management, insurance-linked securities, site-specific engineering analyses, and agricultural risk management. AIR Worldwide, a Verisk Analytics (Nasdaq:VRSK) business, is headquartered in Boston with additional offices in North America, Europe, and Asia. For more information, please visit **www.air-worldwide.com.**