SANDIA REPORT
SAND2021-4905
Printed April 2021

Sandia National Laboratories

Using Bayesian Methodology to Estimate Liquefied Natural Gas Leak Frequencies

Garrett W. Mulcahy, Dusty M. Brooks, Brian D. Ehrhart

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

This analysis provides estimates on the leak frequencies of nine components found in liquefied natural gas (LNG) facilities. Data was taken from a variety of sources, with 25 different data sets included in the analysis. A hierarchical Bayesian model was used that assumes that the log leak frequency follows a normal distribution and the logarithm of the mean of this normal distribution is a linear function of the logarithm of the fractional leak area. This type of model uses uninformed prior distributions that are updated with applicable data. Separate models are fit for each component listed. Five order-of-magnitude fractional leak areas are considered, based on the flow area of the component. Three types of supporting analyses were performed: sensitivity of the model to the data set used, sensitivity of the leak frequency estimates to differences in the model structure or prior distributions, and sufficiency of sample sized used for convergence. Recommended leak frequency distributions for all component types and leak sizes are given. These leak frequency predictions can be used for quantitative risk assessments in the future.

ACKNOWLEDGEMENTS

The authors would like to thank Bobby Middleton at Sandia for providing a technical peer review of this report. The authors also gratefully acknowledge financial support from the Department of Transportation (DOT) Pipeline and Hazardous Materials Safety Administration (PHMSA).

CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page left blank

EXECUTIVE SUMMARY

Safety analyses attempt to predict what hazardous scenarios might occur, how likely they are to occur, and what the consequences would be if they did [1]. A quantitative risk assessment is a way to estimate the likelihood and consequences in order to numerically compare different scenarios or system designs. This quantitative risk assessment method was applied previously for a hydrogen refueling station for fuel cell electric vehicles [2]. For each leak size and for each major type of system component, the annual leak frequency was estimated using probability distributions. These leak frequencies were then combined with estimated physical hazard and probability estimations of various scenarios in order to assess the overall system risk. Liquefied natural gas (LNG) facilities could also benefit from more risk-informed justification for safety and siting requirements.

System components, designs, and installations can vary widely, making it difficult to predict an exact and specific leak frequency value for all systems everywhere. Differences can arise from manufacturing defects, installation errors, component damage, operator error, or other reasons. It is important, therefore, for leak frequencies to account for uncertainty in the estimated value used for risk assessment. In this assessment, uncertainty is represented by a probability distribution on the leak frequency, which represents uncertainty due to lack of knowledge about the true leak frequency as well as uncertainty due to physical variability between systems. Thus, the leak frequency is treated as a random variable with a distribution on how likely different values are, rather than an estimate of uncertainty on a single unknown value. The spread of this distribution could theoretically be reduced with improvements in the state of knowledge but could not be reduced to a single leak frequency due to the population variability.

This analysis provides estimates on the leak frequencies of nine components found in LNG facilities. Various types of these components were reported across the different data sets, but for this analysis we considered only these nine broad categories. The nine component types are: flanges and gaskets, heat exchangers, hoses, joints, loading arms, pipes, valves, vaporizers, and vessels. There are varying quantities of leak event data for each component, from as many as 190 observations for vessels to as few as 2 observations for vaporizers. Data were taken from a variety of literature sources, with 25 different data sets included in the analysis. In order to provide transparent estimates, only publicly available data were used. However, the methodology could be applied by system operators to refine leak frequency estimates for their systems using their proprietary data.

The LNG leak frequency model is a hierarchical Bayesian model that assumes that the log leak frequency follows a normal distribution and the logarithm of the mean of this distribution is a linear function of the logarithm of the fractional leak area. Separate models are fit for each component listed. Components are not differentiated by size in this analysis (i.e., the same leak frequency would apply to a large pipe as a small pipe). This type of model uses uninformed prior distributions that are updated with applicable data. Five fractional leak areas are considered, based on the flow area of the component: 0.0001, 0.001, 0.01, 0.1, and 1.0.

Three types of supporting analyses were performed: sensitivity of the model to the data set used, sensitivity of the leak frequency estimates to differences in the model structure or prior distributions, and sufficiency of sample sizes used for convergence. First, the data were characterized as coming from an LNG facility, a facility that should be similar to an LNG facility (LNG-Applicable), or more generic non-LNG data. In general, variability of the predicted leak frequencies more often increases than decreases when the strictly generic data set is added; therefore, only the LNG-Specific and LNG-Applicable data were used to generate the recommended leak frequency estimates. Next, the data were characterized with respect certainty in the assignment of leak size. For most components, including uncertain leak area (LA) data in the leak frequency estimates leads to a higher estimated leak frequency than when these data are excluded. Uncertain LA data describes data for which the assignment to a 0.0001, 0.001, 0.01, 0.1, or 1.0 fractional leak area was not clear from the reference, so analyst judgement was used to make this assignment.

The second type of sensitivity study analyzed the robustness of model with respect to the choice of priors. For most components, the median of the predicted leak frequency distributions does not appear to be sensitive to mild changes in the prior parameter selection; however, the uncertainty around this median (distribution spread) is sensitive to the prior parameter selection. This means that uncertainty around median leak frequency predictions should be used carefully since they are more sensitive to subjective judgements, while the medians of the predicted leak frequency distributions are not as sensitive to such judgements.

The third supporting analysis determined that sufficient sample sizes were used with the statistical software for convergence of the leak frequency distribution estimates.

Recommended leak frequency distributions for all component types and leak sizes are shown in Table ES-1. There were few data points for some component types, especially for joints and vaporizers, and the model inherently builds high uncertainty into the estimates for those components. In the case of joints, this results in high estimated median leak frequencies for smaller leaks with uncertainty spanning multiple orders of magnitude. Such estimates should be used cautiously and, when possible, updated as new data becomes available.

Table ES-1 Recommended point-estimates and uncertainty intervals for predicted LNG leak frequencies by component and fractional leak size

Component	Leak Size	5th	Median	95th
Flange and Gasket	0.0001	1.13E-05	4.18E-05	1.14E-04
	0.001	3.52E-06	2.26E-05	1.82E-04
	0.01	2.84E-07	1.40E-05	7.14E-04
	0.1	8.81E-08	8.68E-06	7.30E-04
	1	4.03E-08	5.24E-06	5.40E-04
Heat Exchanger	0.0001	5.41E-04	2.34E-03	1.22E-02
	0.001	1.03E-04	8.93E-04	7.27E-03
	0.01	3.11E-05	3.24E-04	3.22E-03
	0.1	2.69E-06	1.17E-04	5.14E-03
	1	3.13E-06	4.18E-05	6.27E-04
Hose	0.0001	4.49E-07	1.52E-06	5.13E-06
	0.001	3.16E-06	7.89E-06	1.99E-05
	0.01	4.38E-08	4.13E-05	3.82E-02
	0.1	4.65E-05	2.14E-04	1.00E-03
	1	2.96E-06	1.10E-03	4.34E-01
Joint	0.0001	9.89E+02	3.51E+04	1.25E+06
	0.001	3.20E+01	4.77E+02	7.09E+03
	0.01	9.98E-01	6.46E+00	4.18E+01
	0.1	2.78E-02	8.76E-02	2.76E-01
	1	4.32E-04	1.19E-03	3.26E-03
Loading Arm	0.0001	4.73E-04	1.99E-01	1.04E+01
	0.001	1.67E-04	1.23E-02	1.59E-01
	0.01	1.92E-05	7.45E-04	7.87E-03
	0.1	9.80E-06	3.29E-05	2.58E-04
	1	7.22E-09	3.03E-06	4.44E-04

Component	Leak Size	5th	Median	95th
Pipe	0.0001	3.08E-07	2.67E-06	2.30E-05
	0.001	1.39E-07	1.44E-06	1.52E-05
	0.01	1.17E-07	7.86E-07	5.22E-06
	0.1	4.59E-08	4.25E-07	3.91E-06
	1	1.15E-08	2.30E-07	4.63E-06
Valve	0.0001	2.40E-05	8.43E-05	2.48E-04
	0.001	8.76E-06	4.20E-05	2.18E-04
	0.01	3.54E-06	2.16E-05	1.53E-04
	0.1	4.72E-07	1.18E-05	2.69E-04
	1	2.34E-07	6.42E-06	1.29E-04
Vaporizer	0.0001	1.27E-04	8.19E-03	5.24E-01
	0.001	1.24E-03	2.63E-02	5.57E-01
	0.01	1.15E-02	8.46E-02	6.23E-01
	0.1	8.65E-02	2.72E-01	8.57E-01
	1	2.79E-01	8.75E-01	2.75E+00
Vessel	0.0001	8.18E-05	4.77E-04	3.41E-03
	0.001	3.69E-06	1.39E-04	5.25E-03
	0.01	1.65E-06	3.90E-05	9.14E-04
	0.1	2.03E-07	1.10E-05	5.80E-04
	1	1.67E-08	3.05E-06	5.77E-04

ACRONYMS AND DEFINITIONS

Abbreviation	Definition
CDF	cumulative distribution function
DOT	Department of Transportation
GTI	Gas Technology Institute
HyRAM	Hydrogen Risk Assessment Models
LA	fractional leak area, defined as $\frac{area\ of\ leak}{cross{-}sectional\ area\ of\ component}$
LNG	liquefied natural gas
NFPA	National Fire Protection Association
MAD	median absolute deviation
PHMSA	Pipeline and Hazardous Materials Safety Administration

1. INTRODUCTION

Safety analyses attempt to predict what hazardous scenarios might occur, how likely they are to occur, and what the consequences would be if they did [1]. A quantitative risk assessment is therefore a way to estimate the likelihood and consequences in order to numerically compare different scenarios or system designs. In 2009, this was performed for a hydrogen refueling station for fuel cell electric vehicles [2]. This report detailed a quantitative risk assessment in which different sized leaks of hydrogen gas were hypothesized to occur and the risk was calculated for these different leaks. For each leak size and for each major type of system component, the annual leak frequency was estimated using probability distributions. These leak frequencies were then combined with estimated physical hazard and probability estimations of various scenarios in order to assess the overall risk of the system. This report was used to inform major revisions to the 2011 edition of the National Fire Protection Association (NFPA) 2 Hydrogen Technologies Code. These calculations formed the basis for the Hydrogen Risk Assessment Models (HyRAM) software, a free and open-source toolkit for the simplified estimations of release behavior and risk for gaseous hydrogen systems [3]. In 2020, leak frequencies were estimated for gaseous hydrogen for the analysis of a hydrogen production facility nearby to a nuclear power plant [4]. This analysis followed the same overall structure for estimating leak frequencies and facility risk as in [2].

Liquefied natural gas (LNG) facilities could also benefit from more risk-informed justification for safety and siting requirements. Requirements in fire codes and standards such as NFPA 52, Vehicular Natural Gas Fuel Systems Code, and NFPA 59A, Standard for the Production, Storage, and Handling of Liquefied Natural Gas (LNG) can be updated and improved through a better understanding and quantification of the risk of different system designs and layouts. Quantitative risk assessment calculations would allow for more direct comparisons between different facilities, different facility designs, and potential facility modifications. In order to inform these types of calculations, the frequency of leaks from system components need to be estimated. These leak frequencies should be estimated for different system components and different leak sizes in order to give a more useful estimate of risk that can be used to compare systems with different designs and configurations.

System components and system designs and installations can vary widely, making it difficult to predict an exact and specific leak frequency value for all systems everywhere. A piece of equipment in one system may leak while the exact same component in another system will not. Leak frequency values can estimate how often a piece of equipment might leak in general, but not exactly when any particular piece of equipment may leak. This difference in whether or not a component is expected to leak can be due to manufacturing defects, installation errors, component damage, operator error, or any number of other reasons. It is important, therefore, for leak frequencies to account for uncertainty in the estimate of whatever value might be used for risk assessment. The assessments discussed previously [2] [4], did not estimate single leak frequency values, but rather a probability distribution of different leak frequencies. This probability distribution of the leak frequency value come from uncertainty about what the values may be (which can be reduced with more and better data) but also from uncertainty about the value of the leak frequency itself. As will be discussed later in this report, fairly broad categories of components are used to determine leak frequencies and different sub-types of each component category may have a slightly different leak frequency. Furthermore, the way systems are designed, installed, and maintained may result in different leak frequencies even for similar types of components. Thus, the leak frequency is treated as a random variable with a distribution on how likely different values are, rather than an estimate of uncertainty on a single unknown value.

This report details an effort to estimate leak frequencies for different LNG system components for different release sizes. Section 2 describes components that are considered for leak frequency estimation as well as the sources of the data used to produce estimates. Section 3 describes the statistical methods used, the model structure, and the software used for implementation. Section 4 documents sensitivity analyses in which various assumptions and initial values are modified in order to illustrate the effect analyst decisions have on model results. Finally, Section 5 contains conclusions from this effort, including recommendations for which leak frequency distribution parameters would form useful default values for something like the HyRAM software toolkit.

2. ANALYSIS SCOPE

The goal of this analysis is to provide estimated annual leak frequencies for LNG system components. The components that were included in the study are described in Section 2.1 and the data sources used to obtain the estimates are described in Section 2.2.

2.1. Component List

This analysis provides estimates on the leak frequencies of the following nine components found in LNG facilities. Various types of these components were reported across the different data sets, but for this analysis we considered only these nine broad categories. Additionally, it should be noted that this analysis does not differentiate components by size; that is, a large pipe and a small pipe would both be categorized as "Pipes" and a single leak frequency distribution would apply to both.

- Flanges and gaskets (the two are treated as one component since they are understood by most users of the PHMSA Failure Rate Table to be the same [5])
- Heat exchangers, including fin fans as well as shell, tube, plate, and pipe heat exchangers
- Hoses
- Joints, mainly expansion joints
- Loading Arms, including ship and truck loading arms
- Pipes, including below ground and above ground
- Valves, including manual, attenuated, motor, solenoid, and air-operated valves
- Vaporizers
- Vessels, including process and pressure vessels, as well as atmospheric and pressurized storage tanks

The data sources in Section 2.2 have many different estimates of component leaks for LNG and other systems. This analysis works to estimate the annual leak frequency for a given component type and for different leak sizes. These estimates are informed by data, which in this case is an estimate of the leak frequency itself. Each "observation" or "data point" for leak frequencies is not always a single event or leak; it is rather an observation of how many leaks of a given size occurred for a given component over a given amount of time. The number of data points identified, categorized, and calculated for each component type is given in Table 2-1. As Table 2-1 demonstrates, there are varying amounts of leak event data for each component, from as many as 190 observations for vessels to as few as 2 observations for vaporizers. The different types of subcategories are delineated more expansively in Appendix A.

Table 2-1 Leak frequency observations by component

Component	Number of Observations
Flange & Gasket	32
Heat Exchanger	76
Hose	20
Joint	6

Loading Arm	19
Pipe	125
Valve	100
Vaporizer	2
Vessel	190

2.2. Data Sources

The data sets in [5] are of different ages, with the oldest set from 1975 and the most recent from 2016. Similarly, these sources vary in the components they report, the type of facility (chemical, nuclear, LNG, unspecified, onshore, offshore), and the type of data (predicted, observed, standards). The data sets are listed in Table 2-2, along with the code by which they will be referred. Appendix A lists the components for which each source has data and any assumptions introduced on the data. The majority of the leak frequencies were given in the units of per year (per meter-year for pipes), so when the frequency did not have these units, we performed the appropriate conversions. For the loading arms and hoses, some leak events were reported in units of rupture per transfer, so since there was no way to convert this rate to operational years, we could not use these values.

Table 2-2 Data sources

Data Set	Source
API 581 '16	API, Risked-Based Inspection Methodology, Recommended Practice 581, Third Edition, American Petroleum Institute, 2016.
CCPS '89	CCPS Guidelines for Process Equipment Reliability Data with Data Tables, American Institute of Chemical Engineers, Center for Chemical Process Safety, New York, NY, 1989.
EGIG '18	EGIG Gas Pipeline Incidents, 10th Report of the European Gas Pipeline Incident Data Group (1970 – 2016), Doc. Number EGIG VA 17.R.0395, March 2018.
GRI LNG FRD '81	Johnson, D.W., and Welker, J.R., Applied Technology Corp. Development of an Improved LNG Plant Failure Rate Database, Final Report for Gas Research Institute, GRI-80/0093, 1981.
HSE FRED NOV '17	HSE Failure Rate and Event Data for use within Risk Assessments, UK, November 2017.
INL CHEM '95	Alber, T.G., Hunt, R.C., Fogarty, S.P., Wilson, J.R., Idaho Chemical Processing Plant Failure Rate Database, Idaho National Engineering Laboratory, NEL-95/0422, August 1995.
INL NUC '07	Plants, N. P. Industry-average performance for components and initiating events at US commercial nuclear power plants, Idaho National Engineering Laboratory, NUREG/CR-6928 (2007).
IOGP 434-1	IOGP Risk Assessment Data Directory, Process Release Frequencies, International Association of Oil and Gas Producers Report No. 434 – 1, 2019.
IOGP 434-3	IOGP Risk Assessment Data Directory, Storage Incident Frequencies, International Association of Oil and Gas Producers Report No. 434 – 3, March 2010.
KGSC '06	Lee, S. R. "Safety comparison of LNG tank designs with fault tree analysis",

	International Gas Union 23rd World Gas Conference proceedings, Amsterdam, Holland. 2006.
KJCE '05	Kim, Hyo, Koh, Jae-Sun, Kim, Youngsoo, University of Seoul Department of Chemical Engineering, and Theofanous, Theofanius, University of California at Santa Barbara Center for Risk Studies and Safety, "Risk Assessment of Membrane Type LNG Storage Tanks in Korea – based on Fault Tree Analysis," Korean Journal of Chemical Engineering, Vo. 22., No. 1, 1-8, 2005.
LEES '12	Mannan, Sam, Lees' Loss Prevention in the Process Industries – Hazard Identification, Assessment and Control, Third Edition, Volume 3, Appendix 14, Elsevier, Inc. 2005.
LNE '09	LNE Handbook of Failure Frequencies, and Appendix, Safety Report, Flemish Government LNE Department Environment, Nature and Energy Policy Unit Safety Reporting Division, 2009.
NFPA 59A '19	NFPA Standard for the Production, Storage, and Handling of Liquefied Natural Gas (LNG), NFPA 59A, 2019 Edition.
PHMSA HL GTI '20	PHMSA Hazardous Liquid Pipeline Accident Data
PHMSA NGT GTI '20	PHMSA Natural Gas Transmission Pipeline Incident Data
PNL PSRP '82	Pelto, P.J., Baker, E.G., Holter, G.M, and Powers, T.B., Analysis of LNG Peakshaving Facility Release Prevention Systems, Pacific Northwest Laboratory, PNL-4153, 1982.
RIVM BEVI '09	RIVM Reference Manual Bevi Risk Assessment, Version 3.2, Module C, the Netherlands National Institute of Public Health and Environment (RIVM), July 1, 2009.
SAI '75	SAI, LNG Terminal Risk Assessment Study for Oxnard, California, Science Applications, Inc. SAI-75-615-LJ, 1975.
SERCO AEA '04	O'Donnell, IJ and Phillips, DW, Serco Assurance, and Winter, PW, AEA Technology, "A New Estimate of the Likelihood of Spontaneous Catastrophic Failure of Pressurized LPG Storage Vessels", Hazards XXVIII, Symposium Series No. 150, IChemE, 2004.
SIGTTO IP4 '96	SIGGTO, Accident Prevention – The Use of Hoses & Hard Arm at Marine Terminals Handling Liquefied Gas, Information Paper No. 4, 1996.
TGC '03	Miyazaki, Sinichi, Tokyo Gas Co. Ltd. and Yamada, Yoshihisa, Tokyo Gas Engineering Co., Ltd., "Quantitative Risk Assessment of LNG Above Ground Tanks Based on Past Operating Records of LNG Regasification Terminals and Life Cycle Assessment", International Gas Union 22nd World Gas Conference proceedings, Tokyo, June 1-5, 2003.
TNO PURPLE '05	TNO Guidelines for Quantitative Risk Assessment (TNO Purple Book), Committee for the Prevention of Disasters (CPR), National institute of Public Health and the Environment (RIVM), The Netherlands Organization for Applied Scientific Research (TNO). First edition 1999/2005.
TNO RED '05	TNO Methods for Determining and Processing Probabilities (TNO Red Book),

	Committee for the Prevention of Disasters (CPR), National Institute of Public Health and the Environment (RIVM), The Netherlands Organization for Applied Scientific Research (TNO). Second edition. 1997/2005.
WELKER '76	Welker, J.R., Brown, L.E., Ice, J.N., Martinsen, W.E., and West, H.H., Fire Safety Aboard LNG Vessels, Final Report DOT-CG-42, 355A, Task #1.

Since the publication of [5] in 2017, new editions of EGIG '15, HSE FRED JUN '12, IOGP 4343-1 '10 and NFPA 59A '16 (codes as given in [5]) were released. We replaced those datasets with their updated versions and renamed them EGIG '18, HSE FRED NOV '17, IOGP 434-1, and NFPA 59A '19, respectively.

The data sets PHMSA HL GTI '16 and PHMSA NGT GTI '16 were based on calculations from databases that PHMSA maintains and continually updates. These data sets were last updated in June 2020 at the writing of this report, so we recalculated the leak frequencies. In this recalculation, we conservatively assumed that each reported event corresponded to a fractional leak of 1 (as the previously data set did), "LENGTH_ISOLATED_SEGMENT" was the pipe length, and that the number of operational years for each pipe was the number of years between 2020 and "INSTALLATION_YEAR" (variables name as given in the PHMSA database). We replaced the '16 data sets with sets titled PHMSA HL GTI '20 and PHMSA NGT GTI '20, respectively.

The calculations and assumptions made about these data sets to prepare them for analysis (specifically with the calculation of fractional leak area) are detailed in Appendix A. We had to associate a fractional leak area with each reported leak frequency; oftentimes this fractional leak area (defined as the ratio between leak area and cross-sectional area of the component) was reported by the dataset or easily calculable. Common calculations included squaring the ratio of the reported leak diameter to the reported cross-sectional diameter. Sometimes this was not possible, such as when a "leak" or "rupture" frequency was reported with no indication to size. In these cases, we had to make assumptions on the fractional leak area; these assumptions are discussed in Section 4.1.2. We also had to exclude some data points because of redundancies across the different sources; these are also outlined in Appendix A. For example, some of the data sets cited or combined data for a component from other data sets in the database. When this occurred, we excluded the component data from all but one of the datasets under question in the analysis.

3. METHODS

In this section we introduce the statistical methodology used to analyze the leak frequency database, the model used to predict leak frequencies based on this database, and the software used to perform the analysis. This analysis closely mirrors that of [4].

3.1. Bayesian Statistical Method

The Bayesian statistical paradigm is governed by Bayes Theorem (stated below), where θ is a vector of parameters and x is observed data from the continuous probability distribution governed by θ [6] [7].

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)d\theta} \quad \text{Equation 1}$$

In Equation 1, $p(x|\theta)$ is the likelihood, which is the joint probability distribution of data regarded as a function of θ. The function $p(\theta)$ is the prior, and it is of considerable importance because it allows the analyst to incorporate previous knowledge or assumptions about θ. The denominator is a normalization constant so that the posterior $p(\theta|x)$ integrates to 1 and is thus a probability distribution.

The Bayesian approach is well suited to this analysis. For many components (e.g., joints and vaporizers) there is little observed data regarding leak frequencies, in which case the Bayesian approach provides useful expressions of the uncertainty in parameter estimates with greater flexibility than frequentist approaches, especially with respect to the inclusion of expert knowledge. The Bayesian model used for this analysis predicts leak frequencies for different leak sizes with uncertainty in the parameters defining this relationship and uncertainty about the relationship. This generates different leak frequency estimates as well as different estimates of uncertainty around the leak frequency based on leak size. A linear relationship is assumed within the model, which relates the logarithm of the mean of the underlying normal distribution on log leak frequency to the logarithm of the leak area. The linear nature of the model also allows data from one leak size to inform the frequency estimate for leaks of a different size. It is important to note that when there is an abundance of high-quality data, the Bayesian estimates will tend toward the frequentist estimates.

To better understand the strengths and weakness of the Bayesian model we employ, we perform two data sensitivity analyses, documented in Section 4.1, as well as model sensitivity analysis, documented in Section 4.2.

3.2. LNG Frequency Model

The LNG leak frequency model is a hierarchical Bayesian model that assumes that the logarithm of the mean of the distribution on log leak frequency is a linear function of the logarithm of the fractional leak area. Separate models are fit for each component listed in Section 2.1 and the coefficients of the linear relationship, α_1 and α_2, are assumed to be normally distributed. The model for a fixed component is given below, where the normal distribution is parameterized by mean and precision and the gamma distribution is parameterized by shape and rate:

$$\log(\mu_{LF,j}) = \alpha_1 + \alpha_2 \log(LA_j) \quad \text{Equation 2}$$

$$\alpha_1 \sim N(\alpha_{11}, \alpha_{12}) \quad \text{Equation 3}$$

$$\alpha_2 \sim N(\alpha_{21}, \alpha_{22}) \quad \text{Equation 4}$$

$$\log(LF_j) \sim N(\mu_{LF,j}, \tau_j) \quad \text{Equation 5}$$

$$\tau_j \sim Gamma(r_j, s_j) \quad \text{Equation 6}$$

The variables in the above equations are defined as follows:

- LA is the fractional leak area, defined as $LA = \frac{leak\ area}{cross\text{–}sectional\ area\ of\ component}$
- j is a subscript between 1 and 5 corresponding to the fractional leak area, the correspondence being $LA_1 = 0.0001, LA_2 = 0.0.001, LA_3 = 0.01, LA_4 = 0.1, LA_5 = 1$
- $\mu_{LF,j}$ is the mean of the underlying normal distribution on log leak frequency for leaks with fractional leak area LA_j
- α_1, α_2 are parameters governing the relationship between $\mu_{LF,j}$ and LA_j. Notice that exponentiating Equation 2 gives $\mu_{LF,j} = e^{\alpha_1} \times LA^{\alpha_2}$
- τ_j is the precision of the distribution of the $\log(LF_j)$; the precision of a normal distribution is the multiplicative inverse of the variance
- α_{11}, α_{12}, α_{21}, α_{22}, r_j, and s_j are the hyperparameters of this hierarchical model.

We discretize the fractional leak area variable because most of data provided leak frequencies in ranges of leak areas. Broadly speaking, the fractional leak areas can be interpreted as follows:

- $LA = 0.0001$—Very small leak
- $LA = 0.001$—Small leak
- $LA = 0.01$—Medium leak
- $LA = 0.1$—Large leak
- $LA = 1$—Very large leak (full-bore release)

The priors for α_1, α_2, and τ_j were selected as follows:

$$\alpha_1 \sim N(\alpha_{11} = 0, \alpha_{12} = 10^{-3}) \quad \text{Equation 7}$$

$$\alpha_2 \sim N(\alpha_{21} = 0, \alpha_{22} = 10^{-3}) \quad \text{Equation 8}$$

$$\tau_j \sim Gamma(r_j = 5, s_j = 1) \quad \text{Equation 9}$$

We chose uninformed priors for α_1 and α_2 to avoid making any significant assumptions on sign or magnitude and instead learn what the data suggested. These priors are said to be uninformed because the normal distributions for these parameters are centered around zero with low precision. A gamma distribution prior is assumed for τ_j because τ_j is the precision for the normal distribution on the logarithm of leak frequency and precision is defined to be non-negative. Initially, a value of $r_j = 1$ was used for the gamma distribution, but there was poor convergence in the mean for this specific prior on τ_j; the first parameter was incremented to achieve reasonable convergence in the mean. This is a reasonable tactic because we know that leaks are not common enough to have infinite, or near infinite, frequency. Adjusting the prior on precision to avoid unrealistically large leak frequencies incorporates relevant expert judgements into the model. We chose to increment the first parameter as opposed to the second parameter because this shifts the gamma distribution to the right (allowing for convergence of the mean), while still concentrating most of the distribution at lower precisions than they would be if the second parameter were increased (preserving some lack of knowledge about the precision). The change in the τ_j makes smaller precisions (i.e., larger variances) less likely, so we have a more informed prior for the precision than the initial prior. This prior is a stronger assumption than that of an uninformed prior, so when there is less data available this decision will yield significant influence on the predicted leak frequencies. Sensitivity of the leak frequency predictions to this choice in prior is discussed in Section 4.2.3.

3.3. Software

For each update, the model was fit with JAGS which was called from R using the rjags package [8] [9] [10]. No initial values were provided. The sample size for model generation was 1×10^6 and the sample size used for model updates was 1×10^6. These sample sizes were chosen based on their sufficiency in previous work [4] and may be in excess of what is necessary. However, the cost of such high sample sizes is negligible. The sample size used to define the posterior was 5×10^5; 10^5 samples were taken for each chain, with five total chains. A smaller sample size is used to define the posterior because these samples (as opposed to those used for model generation and updating) must be saved and imported into software used for generating plots. Saving and loading larger samples is cumbersome and leads to no visible differences between plots of the sample distributions. Multiple chains were used, as is common practice [7], because comparison between the empirical distributions of individual chains is a useful diagnostic for assessing sample convergence.

4. SENSITIVITY AND CONVERGENCE ANALYSES

This section details the sensitivity and convergence analyses applied to the model described in Section 3. Three types of analyses were performed. The first, in Section 4.1, addresses sensitivity of the model to the data set used to fit it. Because the data described in Section 2.2 vary with respect to LNG-applicability and level of detail, the analyses in Section 4.1 were performed to better understand how analyst assumptions applied to the data affect analysis results. The analyses in Section 4.2 assess sensitivity of the leak frequency estimates to differences in the model structure or prior distributions. Section 4.3 examines the sufficiency of the sample size used to characterize the posterior distributions from the model.

4.1. Sensitivity Analysis with Respect to Data

This section presents the results of two separate sensitivity studies. In each analysis we perform three sequential Bayesian updates to the model in Section 3.2 to gauge the influence of different data assumptions on the predictive posteriors for the leak frequencies. Sufficiency of the posterior distribution sample size is further analyzed in Section 4.3.

4.1.1. Sensitivity to Generic Data

Table 13 of [5] classifies the various data sets by their source fluid and their applicability to LNG facilities. Only four data sets were reported to have LNG source fluid, meaning that LNG was "the primary fluid service from which specific equipment failure source data/reference was derived." However, many data sets were categorized as being applicable to LNG fluid service, meaning "the failure source data is predicted, specified, or regulated for LNG, in the judgement of the project team." Thus, the data are classified into three groups:

- LNG Source Fluid
- LNG-Applicable
- Strictly Generic

It is important to note that data classified as LNG Source Fluid data are also considered LNG-Applicable. Strictly generic data are defined as data that are not directly LNG-Applicable. The amount of data in each group for each component is shown below.

Table 4-1 Number of data points categorized by LNG applicability

Component	LNG Source Fluid	LNG-Applicable	Strictly Generic
Flange & Gasket	0	32	0
Heat Exchanger	0	58	18
Hose	0	16	4
Joint	0	6	0
Loading Arm	0	18	1
Pipe	1	86	39
Valve	0	61	39
Vaporizer	1	2	0
Vessel	13	168	22

Ideally, an analysis of LNG plants would only rely on data from LNG facilities, but Table 4-1 reveals that there are not sufficient data to do so. Rather, to have meaningful leak frequency prediction we must use some amount of LNG-Applicable or generic data. To investigate the effect of including LNG-Applicable versus strictly generic data, the model was fit in three sequential updates:

1. LNG-Specific: only LNG source fluid data was utilized (for components with no such data, we disregard the uninformative output)
2. LNG-Applicable: we add in data classified as LNG-Applicable and refit the model
3. All data: we add in strictly generic data and refit the model

The frequency estimates are presented in Table C-1 in Appendix C, but for this discussion, the analysis focuses on examination of violin plots of the predictive posteriors created using *Violin Plot* in MATLAB [11]. Violin plots are a way of visualizing multiple distributions on the same axis for easy comparison. A violin plot displays distributions rotated vertically and then reflected across their axes to give the symmetric "violin" shape. The thickness of each "violin" is then normalized so that the center and relative variance of each distribution is easily viewable. Results from two components are provided to illustrate general trends; plots corresponding to rest of the components are in Appendix D.1

The predictive posteriors for valves are given in Figure 4-1. The green distributions are the posteriors from the first update described above, orange the second update, and red the third. Note that the y-axis is in log scale. The disperse leak predictions resulting from using only LNG-Specific data reflects the fact that there was no LNG-Specific data for valves; the domain of the distribution includes much lower and much higher frequencies than are reasonable (effectively never to always). We see that adding LNG-Applicable data immediately narrows this distribution to leak frequency predictions that are consistent with the scale of frequency we would expect at all leak areas. When the strictly generic data are added, there are changes to both the center and spread of the predictive posteriors. It appears that including strictly generic data (red) in addition to LNG data and LNG-Applicable data increases both the median predicted leak frequency and the variance of the distribution for all leak areas.

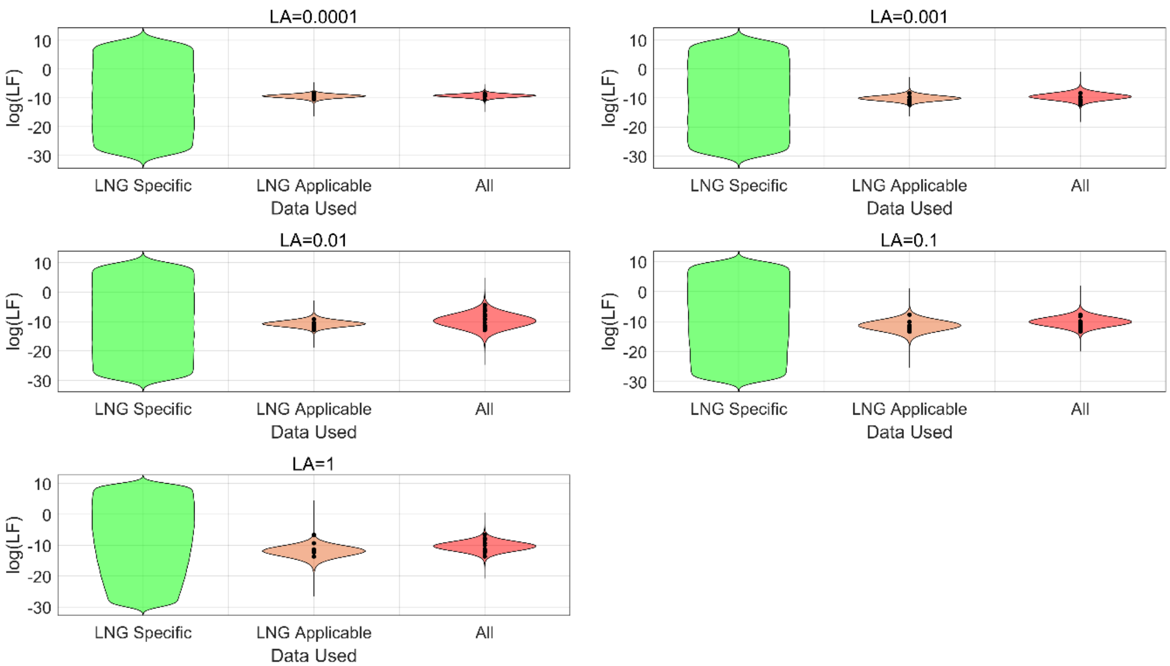

Figure 4-1 Effect of data applicability on valve leak predictions

The predictive posteriors for vessels are given in Figure 4-2. Unlike for valves, there are meaningful leak frequency predictions from LNG-Specific data because there are 13 such points as opposed to 0. As the data set is expanded to include all LNG-Applicable data, the variances of the predicted leak frequencies tend to increase (except when LA = 0.001) and the modes of the distributions tend to increase (except when LA = 0.1 and LA = 1). Lastly, there is little change when the strictly generic data are added, due to the small amount of strictly generic data relative to the amount of LNG-Applicable data.

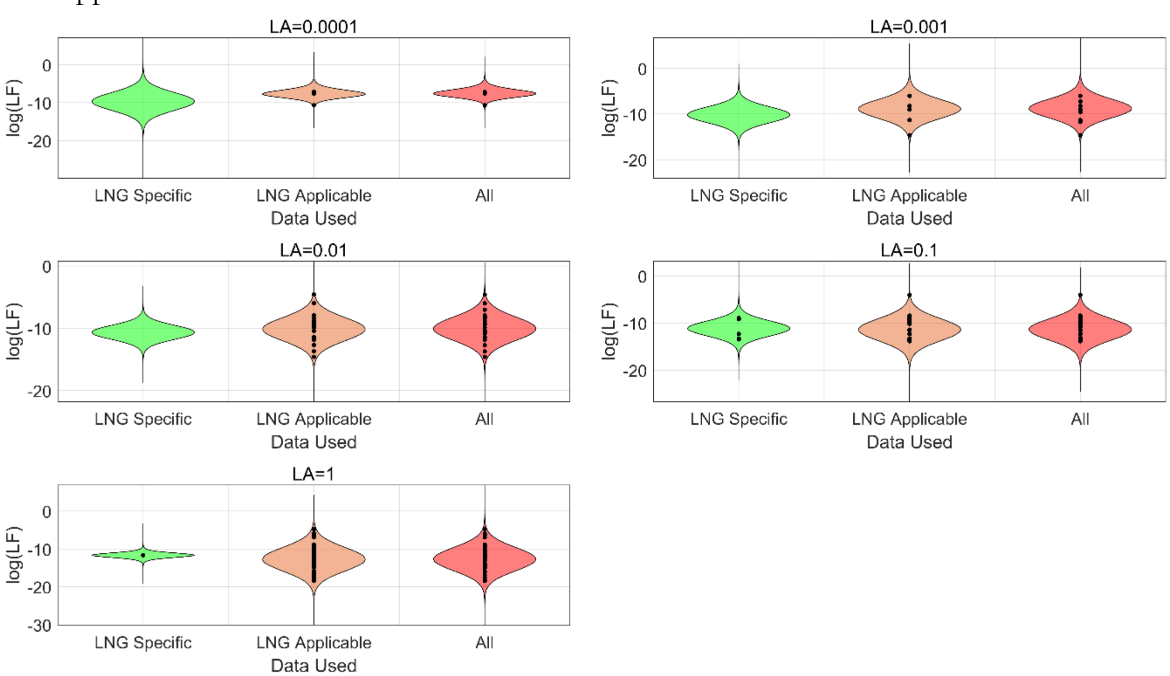

Figure 4-2 Effect of generic data on vessel leak predictions

Although we refer the reader to Appendix D.1 for the rest of the components, there are a few other trends worth noting. For most components, the relationship between fractional leak area and leak frequency was negative, meaning that larger fractional leak areas were predicted to occur less frequently. However, for hoses and vaporizers this trend was reversed, with larger leaks predicted to be more common than smaller leaks. This trend reversal might have occurred because for hoses and vaporizers there was much more data for larger leaks than smaller leaks, and this scarcity of data for smaller leaks makes extrapolations of the trend from larger leaks to smaller leaks unreliable. Alternatively, it may be the case that some components are more prone to larger leaks, for example if smaller leaks sufficiently weaken the component that they quickly widen to large leaks.

There may also be potential for unreliable extrapolation for loading arms and joints; even with all the data included, the model predicts very large leak frequencies for small leaks. Though this is consistent with the intuitive trend that small leaks should be more frequent, the current estimates may nonetheless be artificially large because they are extrapolated from a relative small amount of data exclusively pertaining to larger leaks. The scarcity of data for small leaks for these components could be because repetitive mechanical motion of loading arms and joints leads to larger leaks, in which case the extrapolation over-predicts small leaks. But it could also be the case that small leaks are difficult to detect and may be under-reported, in which case the extrapolation may not be an over-prediction. As such, these leak frequency predications should be interpreted cautiously, as the extrapolation to small leaks may be biased and the direction of this bias is not obvious.

Table 4-2 summarizes the effect of generic data on the center and spread of the predicted leak frequencies at each leak area. The median leak frequency is presented as the measure of central tendency, and the spread is summarized by the median absolute deviation (MAD) [12]. The MAD is the median of the absolute difference between each sample and the sample median. This gives some idea of spread in predictions; however, the median from which the deviance is calculated may also be poorly estimated for small data sets. Values of NA for "not applicable" are reported in the table when there were no data reported for that specific update. The metrics show that, a slight majority of the time, the predictions calculated using the LNG-Applicable data have a lower MAD than the predictions calculated using either the LNG-Specific data or all data. Due to the lack of LNG-Specific data for many components, lower MAD values may support the use of LNG-Applicable data for leak frequency prediction.

The use of generic data is not necessarily precluded by the MAD values comparison. Higher MAD values on average for predictions fit using generic data give us reason to question whether generic data is representative of LNG systems. However, it may also be the case that generic data is representative, and the MAD values reflect variability due to the relatively small number of data points. It could be appropriate to include specific generic data points based on an expert judgement of representativeness.

Table 4-2 Summary of effect of generic data on center and spread of predicted leak frequencies

Component		Median Predicted Leak Frequency (per year)			Median Absolute Deviation		
	Leak Area	LNG-Specific	LNG-Applicable	All	LNG-Specific	LNG-Applicable	All
Flange & Gasket	0.0001	NA	4.18E-05	4.18E-05	NA	1.75E-05	1.75E-05
	0.001	NA	2.26E-05	2.26E-05	NA	1.49E-05	1.49E-05
	0.01	NA	1.40E-05	1.40E-05	NA	1.32E-05	1.32E-05

	0.1	NA	8.68E-06	8.69E-06	NA	8.41E-06	8.42E-06
	1	NA	5.24E-06	5.23E-06	NA	5.12E-06	5.11E-06
Heat Exchanger	0.0001	NA	2.34E-03	1.52E-03	NA	1.29E-03	7.62E-04
	0.001	NA	8.93E-04	6.11E-04	NA	6.33E-04	5.71E-04
	0.01	NA	3.24E-04	2.46E-04	NA	2.44E-04	2.11E-04
	0.1	NA	1.17E-04	9.73E-05	NA	1.09E-04	9.26E-05
	1	NA	4.18E-05	3.92E-05	NA	3.37E-05	3.71E-05
Hose	0.0001	NA	1.52E-06	1.45E-06	NA	6.98E-07	6.40E-07
	0.001	NA	7.89E-06	7.94E-06	NA	2.74E-06	2.76E-06
	0.01	NA	4.13E-05	4.44E-05	NA	4.12E-05	4.43E-05
	0.1	NA	2.13E-04	2.42E-04	NA	1.19E-04	1.24E-04
	1	NA	1.11E-03	1.32E-03	NA	1.10E-03	1.32E-03
Joint	0.0001	NA	3.51E+04	3.51E+04	NA	3.21E+04	3.21E+04
	0.001	NA	4.77E+02	4.76E+02	NA	3.87E+02	3.86E+02
	0.01	NA	6.46E+00	6.45E+00	NA	4.14E+00	4.14E+00
	0.1	NA	8.76E-02	8.76E-02	NA	3.68E-02	3.68E-02
	1	NA	1.19E-03	1.19E-03	NA	4.52E-04	4.53E-04
Loading Arm	0.0001	NA	1.99E-01	1.17E-01	NA	1.96E-01	1.16E-01
	0.001	NA	1.23E-02	8.98E-03	NA	1.12E-02	8.29E-03
	0.01	NA	7.45E-04	6.80E-04	NA	6.35E-04	5.94E-04
	0.1	NA	3.29E-05	3.67E-05	NA	1.72E-05	2.02E-05
	1	NA	3.03E-06	4.19E-06	NA	3.00E-06	4.14E-06
Pipe	0.0001	2.28E-05	2.67E-06	1.11E-06	2.28E-05	1.91E-06	1.05E-06
	0.001	2.30E-05	1.44E-06	6.80E-07	2.30E-05	1.09E-06	5.59E-07
	0.01	2.30E-05	7.86E-07	4.11E-07	2.30E-05	5.16E-07	3.25E-07
	0.1	2.32E-05	4.25E-07	2.50E-07	2.32E-05	3.11E-07	1.86E-07
	1	2.30E-05	2.30E-07	1.50E-07	9.70E-06	1.98E-07	1.40E-07
Valve	0.0001	NA	8.43E-05	9.97E-05	NA	3.65E-05	3.94E-05
	0.001	NA	4.20E-05	7.31E-05	NA	2.41E-05	5.01E-05
	0.01	NA	2.16E-05	5.59E-05	NA	1.39E-05	5.35E-05
	0.1	NA	1.18E-05	4.17E-05	NA	1.04E-05	3.48E-05
	1	NA	6.42E-06	3.22E-05	NA	5.65E-06	2.66E-05
Vaporizer	0.0001	1.22E-03	8.19E-03	8.15E-03	1.22E-03	7.81E-03	7.77E-03
	0.001	7.36E-03	2.63E-02	2.63E-02	7.36E-03	2.27E-02	2.27E-02
	0.01	4.48E-02	8.46E-02	8.46E-02	4.48E-02	5.72E-02	5.72E-02
	0.1	2.72E-01	2.72E-01	2.72E-01	1.14E-01	1.15E-01	1.14E-01
	1	1.66E-00	8.75E-01	8.76E-01	1.66E-00	3.68E-01	3.68E-01
Vessel	0.0001	6.22E-05	4.77E-04	5.30E-04	5.92E-05	3.01E-04	3.38E-04
	0.001	3.81E-05	1.39E-04	1.52E-04	3.28E-05	1.28E-04	1.39E-04
	0.01	2.34E-05	3.90E-05	4.20E-05	1.57E-05	3.44E-05	3.72E-05
	0.1	1.44E-05	1.10E-05	1.16E-05	1.20E-05	1.04E-05	1.09E-05

	1	8.82E-06	3.05E-06	3.17E-06	3.71E-06	3.01E-06	3.12E-06
Colors indicate the ordering of values in each row. There are three shades of blue to distinguish the largest (darker) to smallest (lighter) median leak frequency values in the row. Three shades of green are similarly used for the MAD values in the row. Shades indicate that values within a row are different but do not reflect the magnitude of that difference.							

In summary, the effects of including LNG-Applicable and Strictly Generic data on predicted leak frequencies vary by both component and leak area. Broadly speaking from Table 4-2, though, it appears that variability of the predicted leak frequencies more often increases than decreases when the Strictly Generic data are added. This could be a result of generic data reflecting greater variation in operating conditions as well as a wider variety of age, materials, maintenance conditions, and design type for each component.

The presentation of leak frequency predictions in this manner allows decision-makers to see how the different pieces of available information may affect decisions. This analysis showed that leak frequency predictions are sensitive to the type of data, with respect to LNG applicability, that are included in the model fitting process. Propagation of leak frequency predictions based on different data sets through risk calculations would further demonstrate whether estimated risk is also sensitive to the inclusion of data with varying levels of LNG applicability. This type of analysis could also help focus future data collection by showing where future LNG-Specific data collection would be the most informative with respect to actual risk.

As a result of this sensitivity analysis, the LNG-Applicable data (which includes LNG-Specific data) were used to produce the leak frequency estimates recommended in Section 5.

4.1.2. Sensitivity to Leak Area Assumptions

For most data, associating a reported leak frequency with a leak area was not possible without making assumptions. For example, some references did not report the cross-sectional area of the component, or simply reported a frequency for a "leak" or "rupture" without any more detail about the size. We delineate these assumptions in Appendix A, which were used to categorize the data into three levels of leak area assignment certainty:

- Certain Assignment: In this case, we were able to calculate the fractional leak area and reasonably assign the leak frequency to a leak area bin. For example, pipe leak frequencies were often specified by leak diameter intervals and pipe diameter intervals. The hole and pipe cross-section were assumed to be circular, so the endpoints of the specified intervals could be used to calculate an interval of fractional leak areas. Then a moderately conservative assignment for the fractional leak area was made based off of the leak area that most adequately captured the calculated interval of fractional leak areas.
- Conservative Rupture Assumption: When a rupture frequency was reported with no indication to size, we conservatively assigned the fractional leak area to 1. This was the most common assumption required. For example, if a data set reported a "rupture" frequency and there was no indication of the fractional leak area used to define "rupture" in the body of the report, the leak area was assumed to be $LA = 1$.
- Uncertain Assignment: In this case, the fractional leak area could not be calculated so the leak frequency was assigned in a way that seemed conservative yet reasonable, though conservatism could not be confirmed. For example, if a "leak" frequency was reported and there was no indication in the report as to the size of the leak, we assumed that $LA = 0.01$,

which was consistent with the definition of leak in many other reports which defined "leak". Similarly, some reports gave a frequency of "small leaks", "medium leaks", and "large leaks" without defining the different sizes; in this case, we assigned fractional leak area in a way that was consistent with the leak area interpretation given in Section 3.2 (i.e., a small leak was assumed as $LA = 0.001$, etc.).

Table 4-3 Number of data points categorized by leak area assumption

Component	Certain Assignment	Conservative Rupture Assumption	Uncertain Assignment
Flange & Gasket	24	5	3
Heat Exchanger	35	16	25
Hose	6	7	7
Joint	0	5	1
Loading Arm	6	12	1
Pipe	84	30	11
Valve	48	18	34
Vaporizer	0	1	1
Vessel	8	66	116

The assumptions made to categorize the data include subjectivity, so it is important to understand the effects of those assumptions on the leak frequency prediction model. To do this, we performed three sequential updates, where we included the following data:

1. Only the data classified as Certain Assignment
2. Data classified as Certain Assignment or Conservative Rupture Assumption
3. All data

The results of this analysis are discussed similarly to the results in Section 4.1; the results for select components are discussed to illustrate general trends and results for the remaining components are presented in the appendices. Graphical results are provided in Appendix D and numerical estimates are presented in Table C-2 in Appendix C.

The predictive posteriors for vessel leak frequencies are shown in Figure 4-3. Black dots on the plots indicate the data points used to fit each model. While there are meaningful leak predictions at each update, there are a few interesting changes that occur when the data with uncertain LA assignment is added. First, it appears that with each update the median of the distribution decreases. However, the variance of the distribution also increases. Thus, as data with assumptions are added in the analysis, the predicted leak frequencies appear to become more variable in this case. Reduction in the median may suggest that some of the data was binned non-conservatively. For flanges and gaskets, heat exchangers, joints, and valves, however, the median increases with the inclusion of uncertain LA assignment data (see Table 4-4). This suggests that the assumptions may have led to conservative leak frequency predictions for those components.

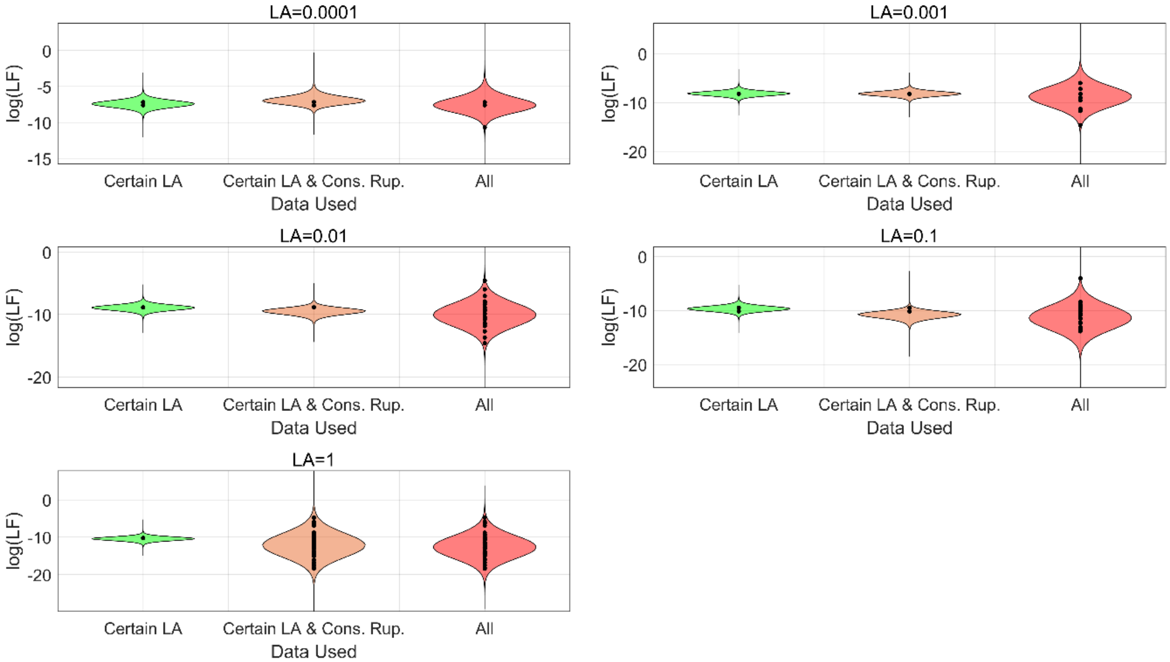

Figure 4-3 Effect of LA assumptions on vessel leak predictions

The predictive posteriors for the leak frequencies of pipes are shown in Figure 4-4. Black dots on the plots indicate the data points used to fit each model. Interestingly, it appears that as data with uncertain LA assignment are added, there is little change in both the center and spread of the distributions of predicted leak frequencies. This is likely attributed to the fact that, compared to the other components, the way the data on pipe leak events were reported was especially conducive to leak area calculations: the data were often reported in ranges of hole size and ranges of pipe diameter, rendering the leak area easily calculable. As a result, roughly two-thirds of the pipe data was classified as “Certain LA”, roughly a quarter as “Conservative Rupture Assumption”, and less than a tenth as “Uncertain LA”.

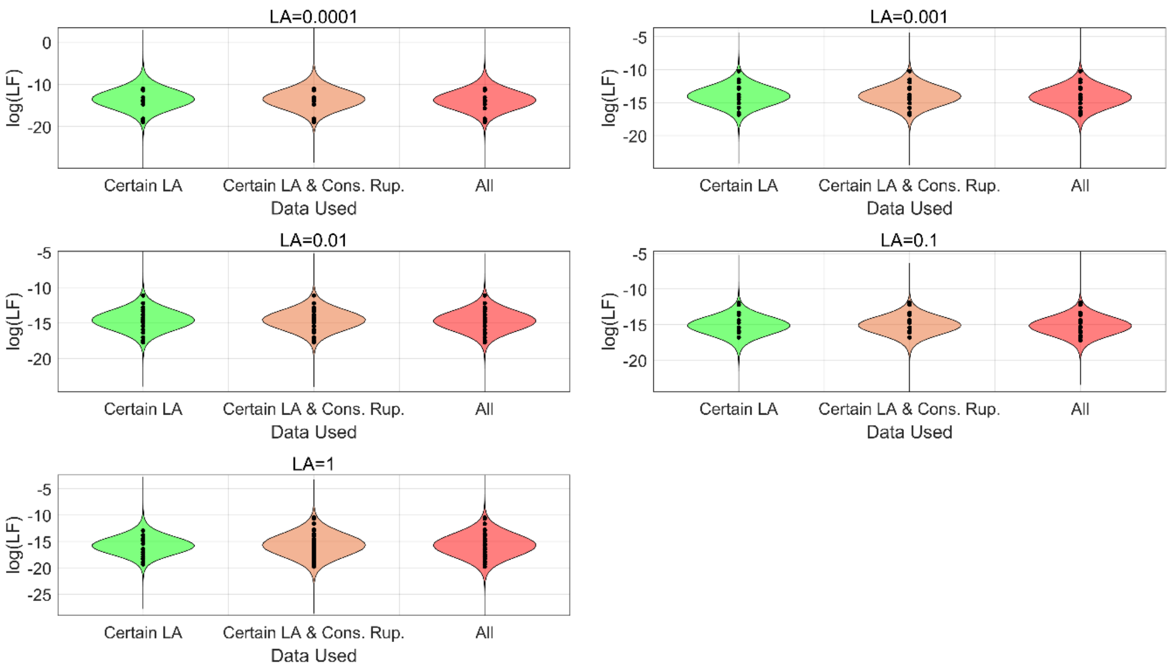

Figure 4-4 Effect of LA assumptions on pipe leak predictions

The predictive posteriors for the leak frequencies of joints are shown in Figure 4-5. This is an interesting case to highlight. There are no data points with certain LA, so the results of the first model update (green) are not useful. However, there are 5 data points for which a conservative assumption about rupture could be made. This is why the second update results in more meaningful predictions for rupture ($LA = 1$) but not for any of the other leak sizes; there is no certainty in the slope and intercept of the linear relationship over leak areas because data only apply to one leak area. A single data point existed in the uncertain LA data set, which was a leak frequency for "leak or rupture". This data point was assumed to correspond to a fractional leak area of 0.1. The addition of this point allows the model to define a more certain linear relationship between the leak areas, which results in less uncertainty around the leak frequency predictions at each leak area, including those for which there is no data. This is an example of a case where the leak frequency prediction is highly dependent on a single analyst assumption due to a lack of data and the assumption is highly subjective.

Though the assumption is subjective, results from the model fit under this assumption are likely conservative. Median joint leak frequencies from the uncertain LA data model may be reasonable to use as conservative estimates in the context of risk analysis. For uncertainty analysis, however, the uncertainty around the median leak frequency estimate may be underestimated. The results are similar for vaporizers, however there is only one data point with a conservative rupture assumption and one data point with an uncertain LA assignment, so those results should be used with more caution.

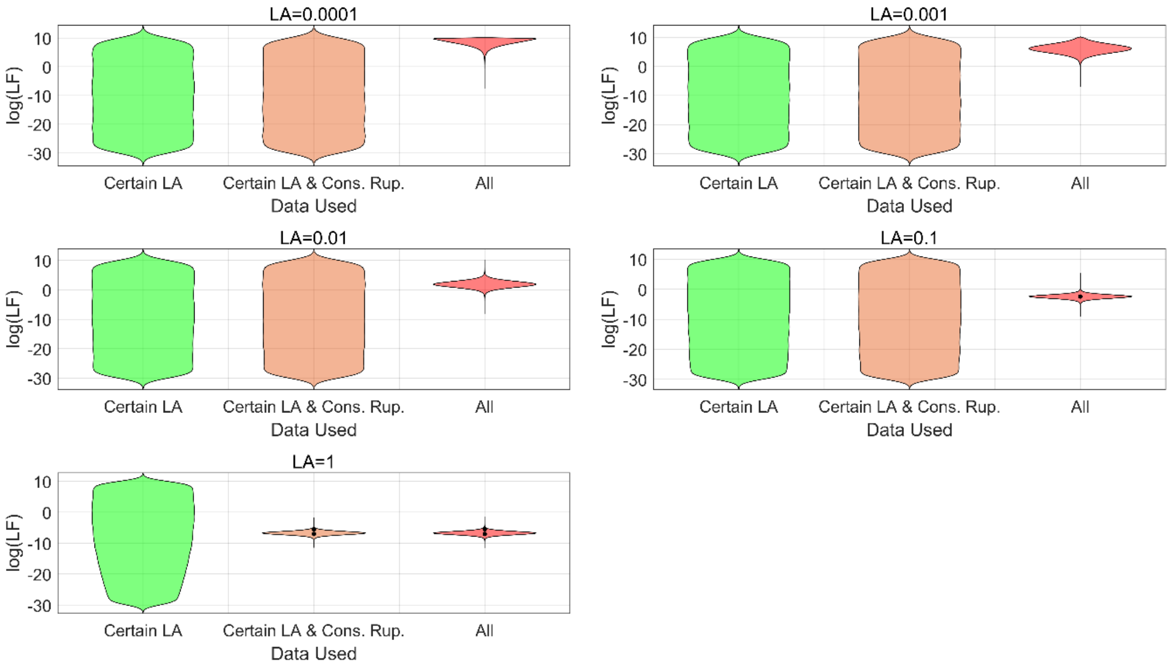

Figure 4-5 Effect of LA assumptions on joint leak predictions

For leak frequency prediction plots for the other components, see Appendix D. Table 4-4 summarizes the effect of fractional leak area assumptions on the center and spread of the predicted leak frequencies. Again, the median is used as the measure of central tendency and the MAD is used as a measure of spread. When a value of NA is reported, it is because there was no data for that component for that update. The MAD values are generally higher when data with all assumptions are included. The MAD values for models that included only certain LA data and the MAD values for models that included certain LA data with conservative assumption/rupture data are comparable.

It makes sense that the spread in predictions would generally be larger from the model fit that included uncertain LA assignment data. The assumptions for this data set are more subjective and increase the variability in the data. Though none of the updates result in a model that is strictly conservative for all components, the median leak frequency is higher for most estimates when the uncertain LA assignment data is included.

In summary, the effect of leak area assumptions on predicted leak frequencies varies by both component and leak area, but it may be conservative for the majority of components to use leak frequency estimates that include uncertain LA assignment data. This analysis also identified components for which the lack of data combined with analyst assumptions generates a solution that may be overly conservative with a poor estimate of uncertainty. This conclusion can inform future data collection if those components (joints and vaporizers) are found to drive risk.

Table 4-4 Effect of LA assumptions on center and spread of predicted leak frequencies

Component	Leak Area	Median Predicted Leak Frequency (per year)			Median Absolute Deviation		
		Certain LA	Certain LA & Cons. Rup.	All	Certain LA	Certain LA & Cons. Rup.	All
Flange & Gasket	0.0001	2.86E-05	2.87E-05	4.18E-05	1.48E-05	1.52E-05	1.75E-05
	0.001	1.16E-05	1.13E-05	2.26E-05	6.09E-06	5.91E-06	1.49E-05
	0.01	5.36E-06	5.17E-06	1.40E-05	2.82E-06	2.70E-06	1.32E-05
	0.1	2.52E-06	2.45E-06	8.65E-06	1.01E-06	1.06E-06	8.38E-06
	1	1.15E-06	1.11E-06	5.25E-06	3.94E-07	1.10E-06	5.13E-06
Heat Exchanger	0.0001	1.45E-03	1.50E-03	1.66E-03	7.16E-04	7.43E-04	8.39E-04
	0.001	5.55E-04	5.55E-04	6.71E-04	2.25E-04	2.25E-04	5.94E-04
	0.01	2.15E-04	2.07E-04	2.69E-04	1.41E-04	1.37E-04	2.22E-04
	0.1	8.11E-05	7.57E-05	1.06E-04	3.35E-05	3.11E-05	1.00E-04
	1	3.14E-05	2.82E-05	4.20E-05	1.27E-05	2.55E-05	3.84E-05
Hose	0.0001	5.19E+03	1.25E+02	1.44E-06	4.05E+03	9.07E+01	6.39E-07
	0.001	4.41E+01	6.85E+00	7.96E-06	2.48E+01	3.63E+00	2.75E-06
	0.01	3.74E-01	3.75E-01	4.43E-05	1.32E-01	1.32E-01	4.42E-05
	0.1	3.18E-03	2.05E-02	2.43E-04	1.57E-03	9.35E-03	1.24E-04
	1	2.70E-05	1.12E-03	1.33E-03	2.60E-05	1.11E-03	1.32E-03
Joint*	0.0001	NA	1.56E-03	3.52E+04	NA	1.56E-03	3.21E+04
	0.001	NA	1.44E-03	4.77E+02	NA	1.44E-03	3.87E+02
	0.01	NA	1.35E-03	6.47E+00	NA	1.35E-03	4.15E+00
	0.1	NA	1.27E-03	8.77E-02	NA	1.27E-03	3.68E-02
	1	NA	1.19E-03	1.19E-03	NA	4.52E-04	4.53E-04
Loading Arm	0.0001	7.36E+00	7.35E-02	1.17E-01	5.90E+00	6.70E-02	1.16E-01
	0.001	9.24E-02	6.23E-03	8.96E-03	5.82E-02	4.74E-03	8.25E-03
	0.01	1.16E-03	5.24E-04	6.79E-04	6.31E-04	3.51E-04	5.92E-04
	0.1	1.46E-05	3.50E-05	3.67E-05	5.27E-06	1.80E-05	2.03E-05
	1	1.83E-07	3.86E-06	4.19E-06	9.93E-08	3.80E-06	4.13E-06
Pipe	0.0001	1.41E-06	1.36E-06	1.11E-06	1.34E-06	1.29E-06	1.05E-06
	0.001	8.10E-07	8.01E-07	6.79E-07	6.51E-07	6.44E-07	5.58E-07
	0.01	4.58E-07	4.66E-07	4.12E-07	3.59E-07	3.65E-07	3.26E-07
	0.1	2.63E-07	2.73E-07	2.50E-07	1.95E-07	2.01E-07	1.86E-07
	1	1.43E-07	1.56E-07	1.49E-07	1.22E-07	1.46E-07	1.40E-07
Valve	0.0001	8.63E-05	8.02E-05	9.98E-05	3.54E-05	3.46E-05	3.93E-05
	0.001	3.40E-05	3.40E-05	7.31E-05	1.81E-05	1.81E-05	5.01E-05
	0.01	1.37E-05	1.49E-05	5.59E-05	7.46E-06	8.33E-06	5.35E-05
	0.1	5.57E-06	6.54E-06	4.17E-05	2.64E-06	3.19E-06	3.48E-05
	1	2.30E-06	2.99E-06	3.22E-05	1.10E-06	2.94E-06	2.66E-05
Vaporizer*	0.0001	NA	1.12E-00	8.08E-03	NA	1.12E-00	7.70E-03
	0.001	NA	1.06E-00	2.61E-02	NA	1.06E-00	2.25E-02
	0.01	NA	1.01E-00	8.42E-02	NA	1.01E-00	5.70E-02
	0.1	NA	9.45E-01	2.71E-01	NA	9.45E-01	1.14E-01
	1	NA	8.76E-01	8.76E-01	NA	3.68E-01	3.68E-01
Vessel	0.0001	5.94E-04	9.68E-04	5.31E-04	2.01E-04	3.65E-04	3.39E-04
	0.001	2.80E-04	2.70E-04	1.51E-04	8.91E-05	9.04E-05	1.39E-04
	0.01	1.32E-04	7.44E-05	4.20E-05	3.98E-05	2.70E-05	3.72E-05
	0.1	6.23E-05	2.03E-05	1.16E-05	2.05E-05	8.87E-06	1.08E-05
	1	2.97E-05	5.25E-06	3.16E-06	1.06E-05	5.20E-06	3.11E-06

* Results for this component should be used with extra caution due to small quantities of data and high sensitivity to analyst assumptions.

Colors indicate the ordering of values in each row. There are three shades of gray to distinguish the largest (darker) to smallest (lighter) median leak frequency values in the row. Three shades of orange are similarly used for the MAD values in the row. Shades indicate that values within a row are different but do not reflect the magnitude of that difference.

4.2. Sensitivity Analyses with Respect to the Model

In this section, we analyze how robust the model is with respect to the choice of priors. The sensitivity is assessed prior-by-prior, perturbing the parameters and gauging the effect on the resultant leak frequency predictions. All of the available data were used in this section, regardless of LNG applicability and LA assumptions. To gauge if the posterior predicted leak frequencies are different under another choice of prior, we employ the Kolmogorov-Smirnov Test [13] on samples taken from the calculated posteriors. Although the Kolmogorov-Smirnov test is a frequentist method, it is appropriate given the large number of samples we obtain from the posterior distributions. For two distributions (or rather, samples from the distributions), the Kolmogorov-Smirnov test establishes whether there is sufficient evidence, at a pre-specified significance level, to conclude that two distributions are different. Difference is detected using the maximum absolute difference between the empirical cumulative distribution functions. In other words, the Kolmogorov-Smirnov test will conclude that the distributions are significantly different if the vertical distance between the cumulative distribution curves is large enough at any point.

For each of the different priors considered, we perform the Kolmogorov-Smirnov test for each pair of predictive leak distributions for each leak area. These results are placed in a matrix, where the entry is 1 if the predicted leak frequency distributions are different and is 0 if the predicted leak frequency distributions are the same. Then we sum the five matrices corresponding to each leak area for a given component. This summation is done component-wise for each of the five leak sizes, so each entry in the resultant matrix is an integer value between 0 (color coded as green, indicating that none of the leak size distributions are significantly different) and 5 (dark red, indicating that all of the leak sizes are significantly different), inclusive. Lower values are preferable because it means that the predictive posteriors for leak frequencies are less sensitive to changes in the prior under investigation.

The diagonal entries in such a matrix are always zero, since it is a comparison of the distribution with itself, and the matrix will always be symmetric. A significance level of 0.05 was used for each test of distribution difference. To speed up computational time, only 5×10^4 samples from the predictive posteriors were used. The sufficiency of this sample size is addressed in Section 4.3.

Tables with the percent change in the median leak frequency between cases with the Original prior and cases with modified priors are also provided for comparison. Because the Kolmogorov-Smirnov test is based on the maximum vertical distance between two cumulative distributions, it can detect that two distributions are different from each other even if they are centered around the same median. The percent change in the median only highlights differences between the centers of two distributions. Both pieces of information are important for assessing the significance of sensitivity relative to the practical application. For example, if two distributions are centered around the same median but are substantially different in the tails, this will only affect a risk analysis if leak frequency uncertainty is included in risk calculations. In this case, the Kolmogorov-Smirnov test would imply sensitivity but the percent change in the median would not.

4.2.1. Prior sensitivity of α_1

For α_1, consider the modifications outlined in Table 4-5. Recall that α_1 is the intercept in the linear model relating log leak area to the logarithm of the mean of the normal distribution on log leak frequency (see Equation 3). The Original prior was used in Section 4.1, Mod1 is a less disperse uninformative prior, and Mod2 is a more disperse uninformative prior. Each prior represents a different extent of lack of knowledge regarding α_1, so we are investigating the sensitivity of predicted leak frequencies to mild changes in the α_1 prior.

Table 4-5 Priors on α_1

Name	Prior
Original	$\alpha_1 \sim N(0, 10^{-3})$
Mod1	$\alpha_1 \sim N(0, 10^{-2})$
Mod2	$\alpha_1 \sim N(0, 10^{-4})$

The results of the sensitivity analysis are displayed in Figure 4-6. We see that for most components, the predictive posteriors for leak frequency are not sensitive to mild changes in the α_1 prior. The notable exception is the loading arm: the predicted leak frequency distributions are statistically significantly different at every leak area when the priors Original and Mod1 are compared as well as Mod1 and Mod2. This may be because there is no very small leak or small leak data for loading arms and there are only 19 data points total (see Figure 4-7). This means the data have less of an influence over the intercept (α_1). The higher precision of the Mod1 assumption may simply be high enough to overcome the influence of limited data.

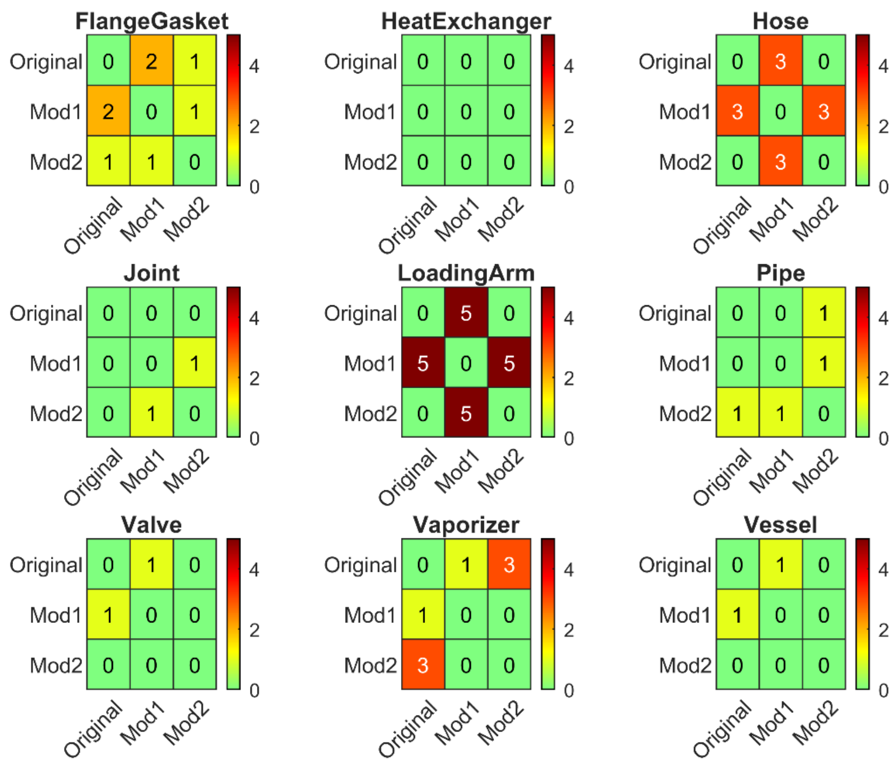

Figure 4-6 Sensitivity to α_1 prior

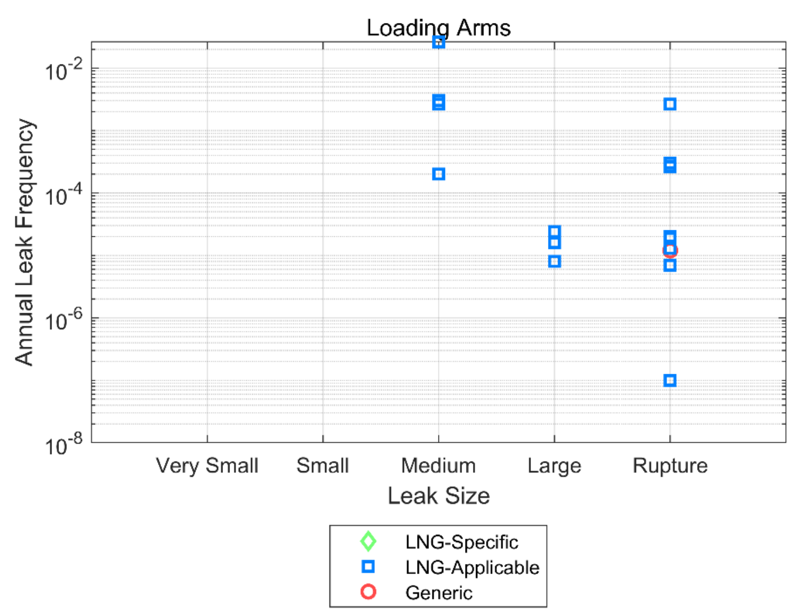

Figure 4-7 Scatter plot of the loading arm data; note the lack of data to influence the model for very small and small leaks as well as the lack of LNG-Specific data.

For hoses, the data for small leaks strongly suggest a positive trend (larger leaks are more frequent, see Figure 4-8). For the Original or Mod2 priors, the precision is low enough that the data can pull the model into this positive-slope regime. However, Mod1 assumes higher confidence that the intercept of the mean should be near 0, which would force the model to have a negative slope. This may be why the model for hoses is sensitive to Mod1.

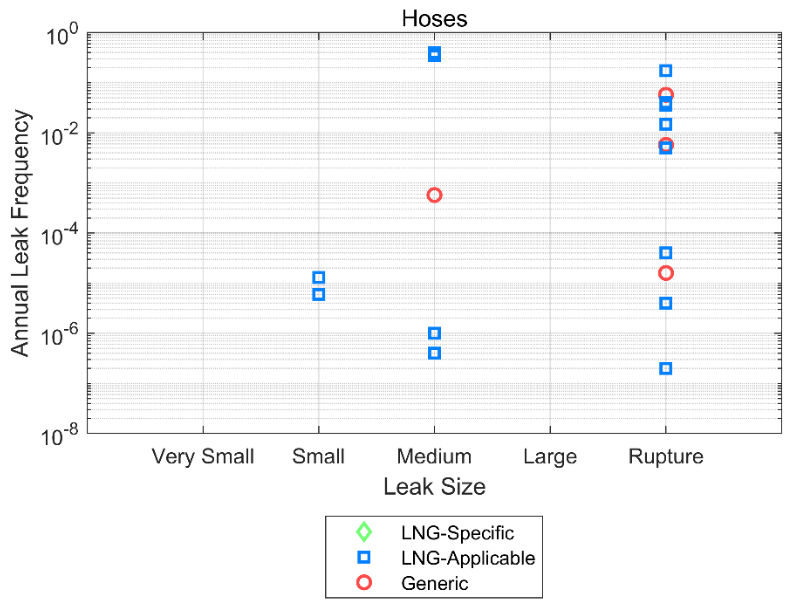

Figure 4-8 Scatter plot of the hose data; note that the data may have a positive trend due to the low frequencies for small leaks.

Table 4-6 is included to assess whether the different predictive posteriors given by the different priors are practically significant (as opposed to statistically significant, which the Kolmogorov-Smirnov test tells us). Table 4-6 presents the percent change between the median leak frequency predicted by the original model and the median leak frequencies predicted by the modified models. This is calculated as described in Equation 10, where m_0 is the median from the original model and m_{mod} is the median from the modified model.

$$100 \times (m_0 - m_{mod})/m_0 \qquad \text{Equation 10}$$

The original median is also presented in the table for context. With the exception of loading arms and ruptures for flanges, the median changes by less than 10%. This suggests that even though changing the prior may result in statistically different predictive posteriors, the centers of the distributions are typically not practically sensitive to the prior assumptions.

Table 4-6 Sensitivity of median predicted leak frequency to α_1 prior

Component		Median Predicted Leak Frequency	Percent Change in Median Predicted Leak Frequency	
	Leak Area	Original Median	Mod1	Mod2
	0.0001	4.18E-05	-0.4%	0.6%

Flange & Gasket	0.001	2.26E-05	2.4%	-0.4%
	0.01	1.40E-05	8.3%	4.7%
	0.1	8.69E-06	9.7%	5.0%
	1	5.23E-06	15.9%	5.3%
Heat Exchanger	0.0001	1.52E-03	-0.3%	-0.3%
	0.001	6.11E-04	0.2%	0.3%
	0.01	2.46E-04	4.6%	1.1%
	0.1	9.73E-05	0.6%	-1.3%
	1	3.92E-05	2.3%	0.9%
Hose	0.0001	1.45E-06	-0.1%	-1.7%
	0.001	7.96E-06	-0.1%	-0.9%
	0.01	4.44E-05	1.3%	-1.7%
	0.1	2.42E-04	-0.2%	1.8%
	1	1.32E-03	2.0%	4.7%
Joint	0.0001	3.51E+04	1.0%	2.2%
	0.001	4.76E+02	1.2%	1.8%
	0.01	6.45E+00	-0.4%	0.5%
	0.1	8.76E-02	0.5%	0.6%
	1	1.19E-03	0.7%	-0.8%
Loading Arm	0.0001	1.17E-01	-19.4%	5.7%
	0.001	8.98E-03	-10.2%	3.9%
	0.01	6.80E-04	-2.7%	2.1%
	0.1	3.67E-05	2.6%	-1.8%
	1	4.19E-06	15.3%	-2.7%
Pipe	0.0001	1.11E-06	-0.7%	-2.8%
	0.001	6.80E-07	-0.8%	-1.6%
	0.01	4.11E-07	-0.5%	-1.0%
	0.1	2.50E-07	-0.6%	-0.9%
	1	1.50E-07	0.1%	-1.0%
Valve	0.0001	9.97E-05	0.8%	0.3%
	0.001	7.31E-05	1.7%	1.5%
	0.01	5.59E-05	-0.4%	-0.6%
	0.1	4.17E-05	-0.6%	0.6%
	1	3.22E-05	1.3%	-0.5%
Vaporizer	0.0001	8.15E-03	1.5%	2.4%
	0.001	2.63E-02	2.5%	2.0%
	0.01	8.46E-02	1.3%	0.9%
	0.1	2.72E-01	-0.3%	-0.1%
	1	8.76E-01	-0.6%	-0.8%
Vessel	0.0001	5.30E-04	0.1%	0.2%
	0.001	1.52E-04	1.5%	2.9%
	0.01	4.20E-05	0.6%	2.4%
	0.1	1.16E-05	-0.3%	1.6%
	1	3.17E-06	-0.6%	-0.9%

4.2.2. Prior sensitivity of α_2

For α_2, consider the modifications outlined in Table 4-7. As previously, Original is the prior used in Section 4.1, while Mod1 and Mod2 represent mild changes in the assumed state of knowledge compared to the Original. We also examine an additional prior in this analysis, Mod3. The prior given by Mod3 insists that α_2 is negative and explicitly builds into the model the intuition that smaller leaks are more frequent than larger leaks. It should be stressed that Mod3 is a very different assumption than the other three priors.

Table 4-7 Priors on α_2

Name	Prior
Original	$\alpha_2 \sim N(0, 10^{-3})$
Mod1	$\alpha_2 \sim N(0, 10^{-2})$
Mod2	$\alpha_2 \sim N(0, 10^{-4})$
Mod3	$-\alpha_2 \sim Gamma(1,1)$

The results from this sensitivity analysis are shown in Figure 4-9. For most components, it appears that the predicted leak frequency distributions are not sensitive to mild (Mod1, Mod2) or substantial (Mod3) changes to the α_2 prior. For the joint and hose, the predicted distributions do not appear to be sensitive to mild changes in the α_2 prior, whereas the vastly different set of assumptions represented by Mod3 appear to give different distributions. Lastly, the predicted leak frequency distributions for the loading arm and vaporizer are sensitive to both mild and substantial changes on the α_2 prior. Note that these are the components for which the least amount of data are available. It makes sense that the model is more sensitive to changes in the priors when data are sparse.

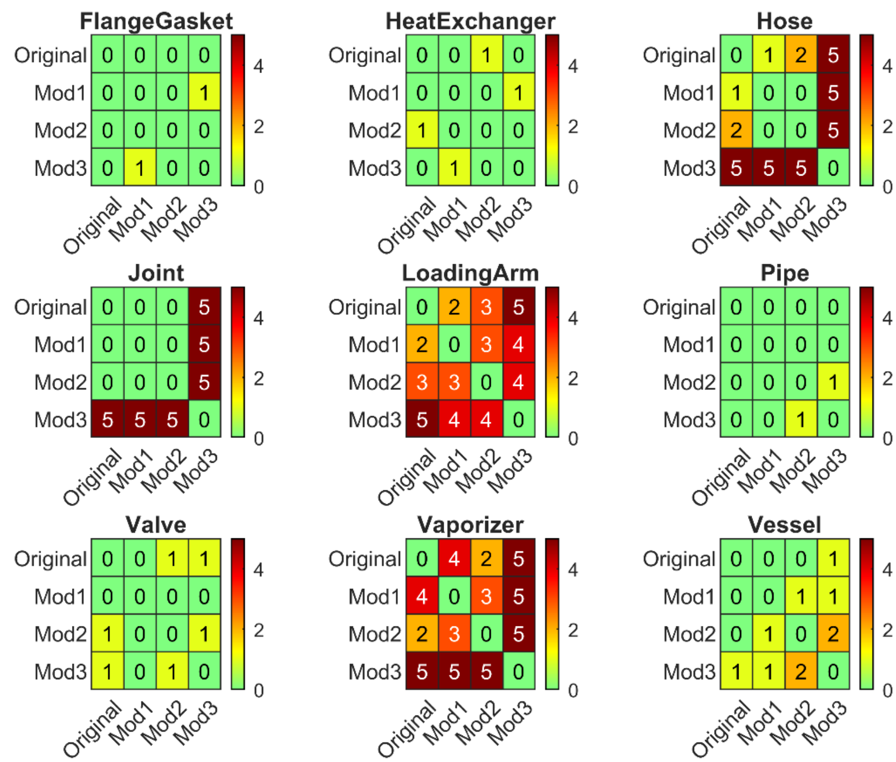

Figure 4-9 Sensitivity to α_2 prior

Table 4-8 presents the percent change (see Equation 10) of the predicted leak frequency at each leak area for each component. For all components, the median changed by less than 5% for the Mod1 and Mod2 cases. This suggests that even though there may be some statistically significantly different predictive posteriors for these changes to prior parameters, the medians are not practically sensitive to the prior distributions.

The Mod3 case also had small practical changes in the median for all components except for hoses, joints, loading arms, and vaporizers, which saw more drastic changes in the median predicted leak frequencies. For hoses and vaporizers, there are medians that changed by more than 100%, and in some cases by much more than that. These substantially different predictions occur because the posteriors of α_2 when no sign assumptions are made (i.e. Original, Mod1, and Mod2) are predominantly positive, so the assumption that α_2 is negative completely changes the direction of the relationship between log leak area and log leak frequency. For joints and loading arms, the negative slope is consistent with the data and the other models. The significantly different behavior may be caused by the gamma distribution on α_2 being concentrated around zero. The negative trends for joints and loading arms in the data are steeply negative, for example compared to the negative trend for pipes (see Figure 4-10). A more disperse prior may have resulted in less significant change.

Table 4-8 Sensitivity of median predicted leak frequency to α_2 prior

Component	Leak Area	Median Predicted Leak Frequency: Original Median	Percent Change in Median Predicted Leak Frequency: Mod1	Mod2	Mod3
Flange & Gasket	0.0001	4.18E-05	0.5%	0.3%	-0.3%
	0.001	2.26E-05	1.5%	0.5%	2.5%
	0.01	1.40E-05	2.3%	4.2%	4.1%
	0.1	8.69E-06	2.8%	3.4%	7.0%
	1	5.23E-06	0.7%	-1.3%	8.1%
Heat Exchanger	0.0001	1.52E-03	0.2%	-0.3%	-0.8%
	0.001	6.11E-04	-0.2%	-0.5%	0.4%
	0.01	2.46E-04	1.8%	2.7%	3.3%
	0.1	9.73E-05	-1.1%	-1.8%	0.8%
	1	3.92E-05	-2.9%	-2.2%	1.8%
Hose	0.0001	1.45E-06	-0.4%	0.6%	821.9%
	0.001	7.96E-06	-0.5%	-0.5%	26.6%
	0.01	4.44E-05	2.6%	-4.4%	-82.1%
	0.1	2.42E-04	0.4%	-1.1%	-97.5%
	1	1.32E-03	-1.0%	-0.3%	-99.6%
Joint	0.0001	3.51E+04	0.4%	1.8%	-35.9%
	0.001	4.76E+02	-0.3%	0.8%	-28.2%
	0.01	6.45E+00	-0.8%	0.4%	-19.4%
	0.1	8.76E-02	0.2%	-0.1%	-8.6%
	1	1.19E-03	0.2%	0.7%	2.3%
Loading Arm	0.0001	1.17E-01	-1.4%	1.0%	-66.8%
	0.001	8.98E-03	0.8%	1.9%	-52.3%
	0.01	6.80E-04	-0.6%	1.4%	-28.3%
	0.1	3.67E-05	0.7%	-0.5%	-0.7%
	1	4.19E-06	1.5%	0.0%	41.1%
Pipe	0.0001	1.11E-06	-0.4%	1.2%	-1.4%
	0.001	6.80E-07	-1.4%	-2.5%	-0.8%
	0.01	4.11E-07	-0.9%	-0.3%	-0.3%
	0.1	2.50E-07	-0.4%	0.6%	0.7%
	1	1.50E-07	-2.7%	-0.8%	-1.2%
Valve	0.0001	9.97E-05	0.2%	-0.4%	-0.1%
	0.001	7.31E-05	2.2%	0.1%	0.9%
	0.01	5.59E-05	-1.0%	0.0%	-0.7%
	0.1	4.17E-05	0.9%	0.1%	-0.4%
	1	3.22E-05	0.2%	0.3%	1.2%
Vaporizer	0.0001	8.15E-03	3.4%	1.2%	14386.9%
	0.001	2.63E-02	3.7%	1.4%	3518.6%
	0.01	8.46E-02	1.8%	1.2%	780.9%
	0.1	2.72E-01	0.6%	0.5%	99.6%
	1	8.76E-01	-0.4%	-0.6%	-50.7%
Vessel	0.0001	5.30E-04	0.6%	-0.7%	-3.7%
	0.001	1.52E-04	3.0%	3.2%	-0.5%
	0.01	4.20E-05	-1.0%	-1.4%	-0.1%
	0.1	1.16E-05	0.3%	0.1%	0.6%

	1	3.16E-06	-1.3%	0.1%	-0.5%

These sensitivity analyses show that the model is not sensitive to small changes in the precision of the prior on the slope (a_2) of the model mean. However, the model is sensitive to changes that enforce a sign condition. This is because the original model allows the data to determine whether larger leaks are more frequent or less frequent and this trend differs by component. When the prior distribution only allows for a negative trend, the model becomes more inaccurate for data with a positive trend or for components with small quantities of data. As Figure 4-10 shows, the joint and loading arm data demonstrate pronounced negative sloping trends dominated by large leak data and a lack of small leak data, compared to the pipe data which demonstrates a less pronounced trend due to high variability at each leak size.

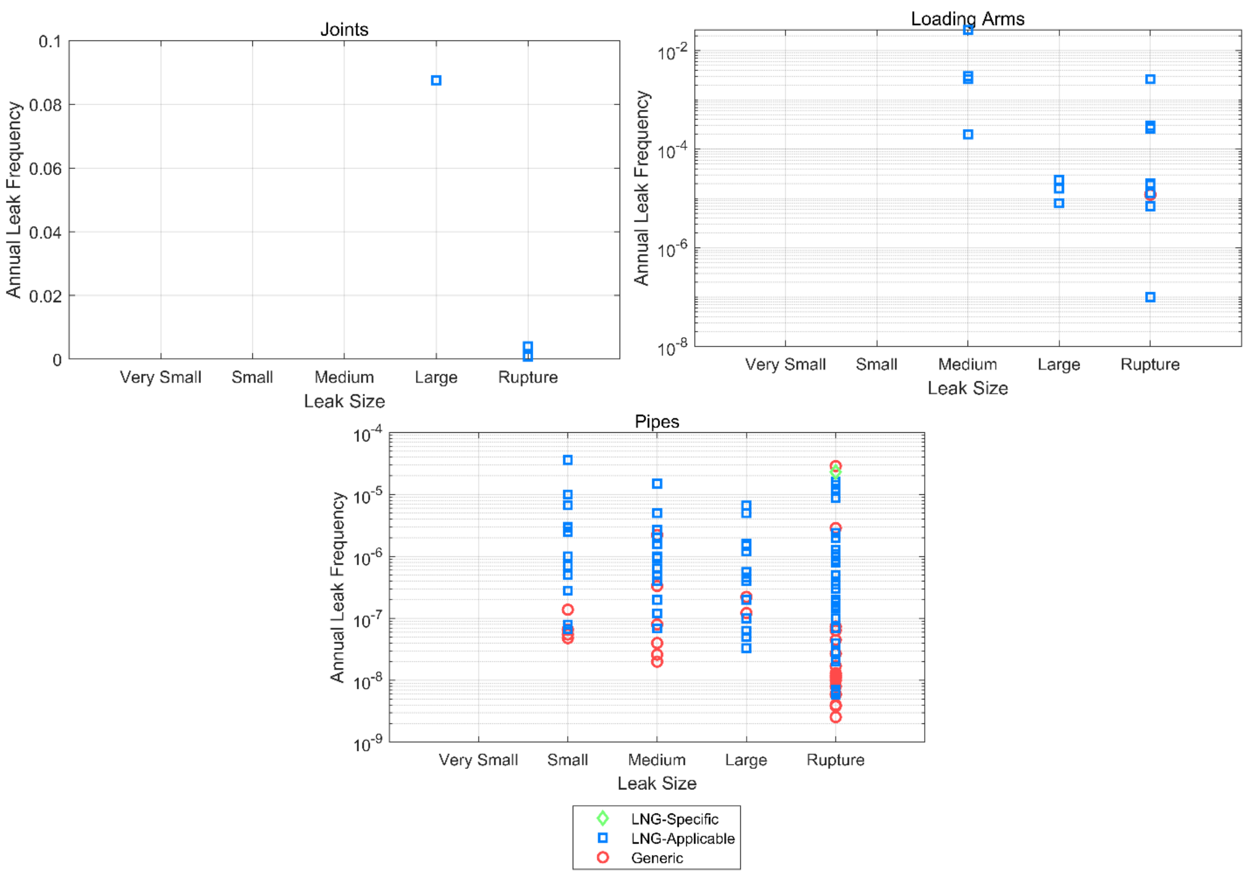

Figure 4-10 Leak frequency data for joints, loading arms, and pipes

4.2.3. *Prior Sensitivity of τ_j*

For τ_j we consider the modifications outlined in Table 4-9. Again, Original denotes the weakly informed prior used in Section 4.1, while Mod1, Mod2, Mod3, and Mod4 are all slight perturbations to the Original prior. Increases in the first parameter of the gamma distribution shift the distribution to the right (leading to higher values of τ_j) with slight widening of the distribution spread, with the opposite effect for decreases in the parameter. Hence, changing the first parameter represents a moderate change in the assumed precision of the model. Increases in the second parameter of the gamma distribution shift the distribution to the right with more pronounced stretching of the distribution spread; decreases have the opposite effect. Thus, changing the second parameter represents a large change in the assumed precision of the model.

Table 4-9 Priors on τ_j

Name	Prior	Interpretation
Original	$\tau_j \sim Gamma(5,1)$	
Mod1	$\tau_j \sim Gamma(5,1.5)$	Large increase in precision
Mod2	$\tau_j \sim Gamma(5,0.5)$	Large decrease in precision
Mod3	$\tau_j \sim Gamma(5.5,1)$	Small increase in precision
Mod4	$\tau_j \sim Gamma(6,1)$	Moderate increase in precision

The results from this sensitivity analysis are presented in Figure 4-11. For all components, the predicted leak frequency distribution is sensitive to perturbations to the τ_j priors, with this sensitivity being especially pronounced for components with little data (vaporizers, loading arms, and joints). It is worth noting that the Original prior was selected for numerical reasons: it allowed for reasonable convergence of the mean. Thus, the Original prior represents a boundary on which the numerical behavior changes, which may be an explanation for the observed sensitivity.

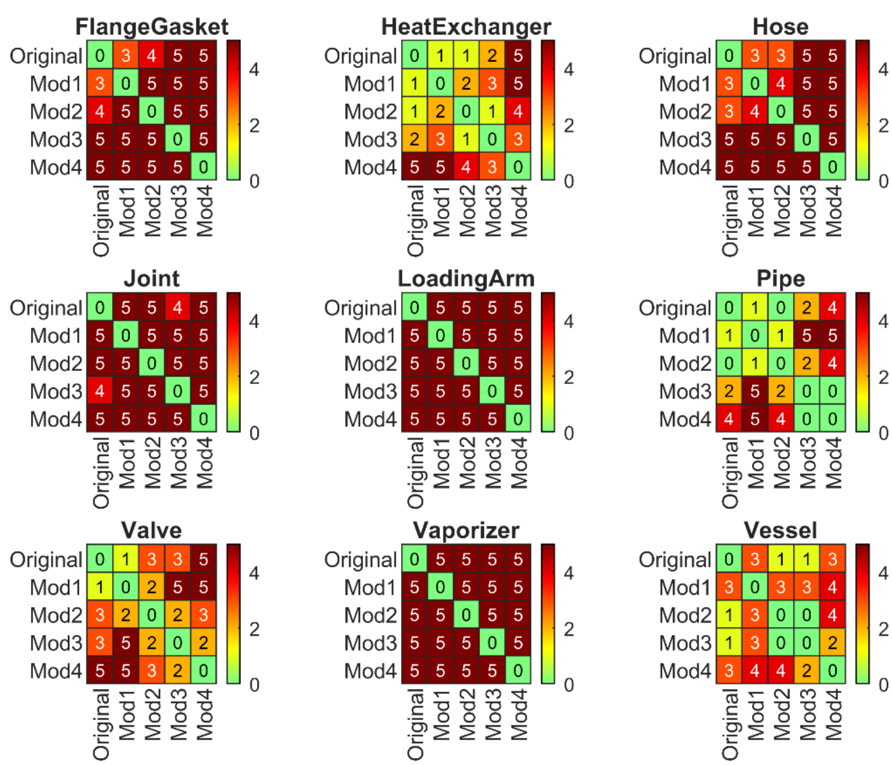

Figure 4-11 Sensitivity to τ_j prior

Table 4-10 presents the percent change in (see Equation 10) median leak frequency predictions at each leak area for each component. The baseline for these comparisons is the median of the predicted leak frequencies given by the Original prior, also provided in the table for context. As in the other sensitivity analyses, most of the medians are not sensitive to changes in this prior, even though the leak frequency distributions are sensitive. The largest difference is again seen for the loading arm, with percent changes ranging from 3.3% to 58.5%.

Table 4-10 Sensitivity of median predicted leak frequency to τ_j prior

Component		Median Predicted Leak Frequency	Percent Change in Median Predicted Leak Frequency			
	Leak Area	Original Median	Mod1	Mod2	Mod3	Mod4
Flange & Gasket	0.0001	4.18E-05	-9.6%	10.6%	1.0%	2.2%
	0.001	2.26E-05	-4.0%	7.8%	2.1%	2.7%
	0.01	1.40E-05	3.4%	6.6%	3.1%	6.7%
	0.1	8.69E-06	6.9%	0.6%	3.4%	4.9%
	1	5.23E-06	14.4%	-1.8%	2.3%	3.6%
Heat Exchanger	0.0001	1.52E-03	1.5%	-1.8%	-0.5%	-1.1%
	0.001	6.11E-04	0.7%	-1.7%	-2.1%	-2.5%
	0.01	2.46E-04	0.8%	2.0%	3.1%	1.0%
	0.1	9.73E-05	-2.0%	-1.8%	-0.7%	1.5%
	1	3.92E-05	-1.0%	-0.2%	-0.8%	-2.2%
Hose	0.0001	1.45E-06	0.2%	-2.0%	-0.6%	-1.5%
	0.001	7.96E-06	-0.1%	-1.3%	-0.2%	-0.3%
	0.01	4.44E-05	2.8%	0.5%	-1.8%	-1.2%
	0.1	2.42E-04	3.0%	0.4%	0.0%	2.0%
	1	1.32E-03	4.9%	-0.1%	1.2%	4.3%
Joint	0.0001	3.51E+04	-0.3%	3.8%	2.5%	0.4%
	0.001	4.76E+02	-0.8%	1.5%	2.4%	-0.7%
	0.01	6.45E+00	-0.1%	1.1%	0.4%	-0.7%
	0.1	8.76E-02	-0.1%	0.6%	0.6%	0.4%
	1	1.19E-03	0.0%	0.0%	0.1%	0.4%
Loading Arm	0.0001	1.17E-01	-18.4%	58.5%	20.3%	73.7%
	0.001	8.98E-03	-4.1%	15.0%	11.5%	37.3%
	0.01	6.80E-04	12.9%	-14.0%	3.3%	8.6%
	0.1	3.67E-05	37.1%	-33.5%	-4.8%	-8.4%
	1	4.19E-06	48.7%	-47.0%	-15.9%	-28.5%
Pipe	0.0001	1.11E-06	-3.0%	-1.0%	-0.8%	-2.7%
	0.001	6.80E-07	0.1%	-1.6%	-2.3%	-1.4%
	0.01	4.11E-07	0.1%	0.6%	2.1%	-0.4%
	0.1	2.50E-07	0.1%	-0.2%	-1.2%	0.2%
	1	1.50E-07	0.2%	-0.7%	-2.5%	1.3%
Valve	0.0001	9.97E-05	-2.9%	4.0%	0.2%	0.6%
	0.001	7.31E-05	-0.1%	4.7%	2.5%	2.9%
	0.01	5.59E-05	-0.7%	2.4%	-1.3%	0.2%
	0.1	4.17E-05	0.2%	-2.1%	-1.3%	-0.9%
	1	3.22E-05	3.6%	-1.0%	0.5%	1.9%
Vaporizer	0.0001	8.15E-03	0.7%	1.5%	0.4%	1.4%
	0.001	2.63E-02	1.8%	1.8%	0.9%	1.3%

	0.01	8.46E-02	0.8%	1.4%	1.3%	1.1%
	0.1	2.72E-01	-0.6%	-0.1%	0.3%	0.0%
	1	8.76E-01	-0.8%	-1.3%	-0.6%	-1.0%
Vessel	0.0001	5.30E-04	4.1%	-3.6%	-4.3%	-6.9%
	0.001	1.52E-04	4.1%	-0.4%	-1.8%	-2.3%
	0.01	4.20E-05	0.0%	-1.1%	-2.6%	1.9%
	0.1	1.16E-05	-3.2%	1.6%	0.4%	-0.7%
	1	3.16E-06	-5.0%	-2.2%	-2.4%	0.1%

Figure 4-11 suggests sensitivity to the prior on τ_j for most components and most leak sizes, while Table 4-10 does not suggest much sensitivity in the median. This is expected since τ_j affects the spread in the final leak frequency distribution, not the center of the distribution. The Kolmogorov-Smirnov test uses the maximum vertical distance between two distributions is the test statistic. This maximum is taken over the domains of the distributions. If a sensitivity case narrows the leak frequency distribution, the maximum distance between the tails of the two distribution will increase. This means that the Kolmogorov-Smirnov test is particularly sensitive to detecting this type of difference between distributions, even if the center of the distribution does not change much.

4.2.4. Model Form Sensitivity

The appropriateness of the mean function (Equation 2) can be assessed by testing for curvature. This can be done by introducing a quadratic term in the model and testing if the corresponding coefficient is nonzero [14]—if it is, there is curvature in the data that is not accurately characterized by the linear function. Thus, we consider the following hypotheses:

$$H_0:\ log(\mu_{LF}) = \alpha_1 + \alpha_2 log\ (LA) \qquad \text{Equation 11}$$

$$H_1: log(\mu_{LF}) = \alpha_1 + \alpha_2\ log(LA) + \alpha_3[log\ (LA)]^2 \qquad \text{Equation 12}$$

which are equivalent to

$$H_0: \alpha_3 = 0 \qquad \text{Equation 13}$$

$$H_1: \alpha_3 \neq 0 \qquad \text{Equation 14}$$

The model is fit with an uninformative prior $\alpha_3 \sim N(0, 10^{-3})$ for the quadratic term. Figure 4-12 displays the interval from the 2.5[th] percentile to the 97.5[th] percentile of the posterior distribution on the quadratic coefficient (α_3) for each component; the numerical values are given in Appendix E. If zero is not included in the interval, H_0 can be rejected. If zero is included in the interval, H_0 cannot be rejected and the conclusion is that there may be significant curvature in the data.

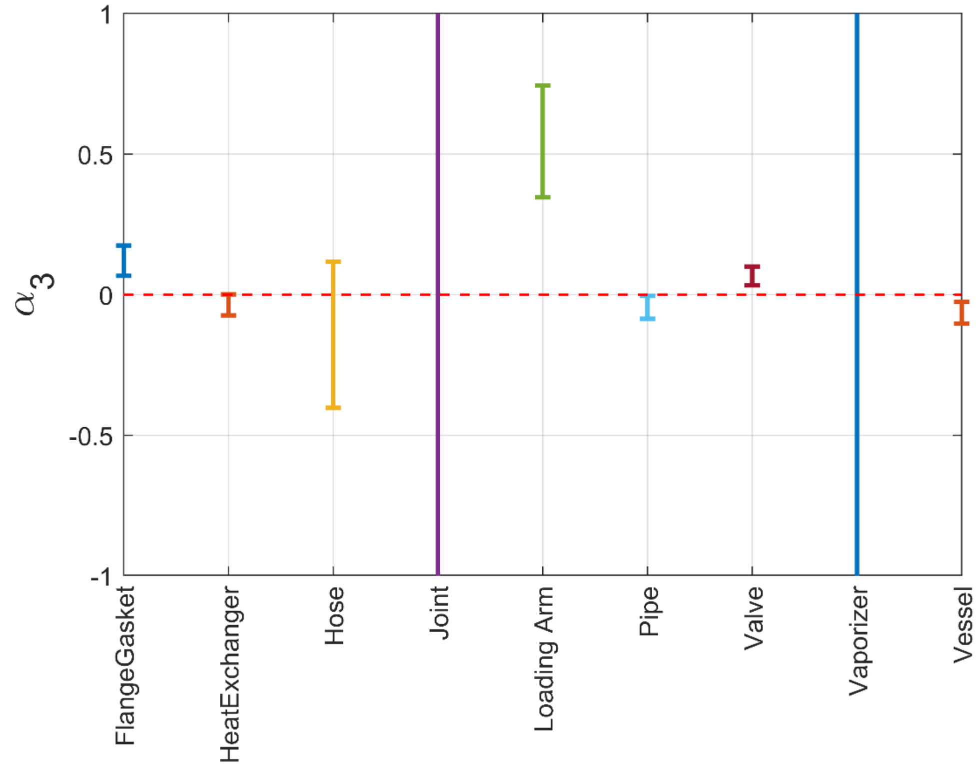

Figure 4-12 95% Intervals (from the 2.5th percentile to the 97.5th percentile) of the posterior distribution on α_3

According to Figure 4-12, there is evidence for curvature in the mean function of leak frequency for flanges and gaskets, heat exchangers, loading arms, pipes, valves, and vessels. These results suggest that future studies may want to consider different mean functions to produce models that more accurately fit the data. However, the curvature may also be a statistical phenomenon due to reliance on data that required several assumptions to be used. These assumptions may have introduced the curvature, so curvature should be reassessed as additional data are accumulated.

4.2.5. Sensitivity Analysis Summary

For most components, the predicted leak frequency distributions do not appear to be sensitive to mild changes in the prior on α_1 and α_2. However, for most components, the distributions appear to be sensitive to slight changes in the prior on τ_j. This means that uncertainty around median leak frequency predictions should be used carefully. The medians of the predicted leak frequency distributions, which are the recommended point estimates for these frequencies, are not sensitive to mild changes in any of the priors.

In future analyses, choosing the prior on τ_j based on physical reasons as opposed to numerical considerations could lead to improvements within the model. Additionally, changing the form of the model may enable convergence without the use of informed priors. Leak frequency predictions are sensitive to the model form for some components, though care should also be taken to avoid over-specifying the model; small data sizes, variability in frequencies at leak sizes, and uneven data coverage over different leak sizes may result in trends (such as curvature) in the data that may not reflect reality. This analysis investigated sensitivity at the level of leak frequency predictions; sensitivity may be different if investigated at the level of scenario risk estimates depending on how sensitive risk measures are to the predicted leak frequencies.

4.3. Convergence Analysis

Leak frequency distributions predicted using the model described in Section 3.2 are characterized by samples. To use these distributions, it is important to check that the sampling uncertainty is low. In other words, it must be established that a sufficient number of samples have been taken to characterize the distribution. This is convergence with respect to sampling uncertainty.

Recall that the Bayesian model outlined in Section 3.2 is fit using the JAGS software, which applies Gibbs sampling to obtain a large number of samples from the posterior distributions in such a way that the samples provide a very good approximation to these posterior distributions (whose analytic forms are often not tractable). In this convergence analysis, we demonstrate that the samples generated in Sections 4.1 and 4.2, and used to generate the final results, were sufficient to approximate the posteriors for the various model parameters. This is done by demonstrating that fewer samples would have sufficed.

For each of the nine components, we examined the samples obtained from the posterior distributions of the model parameters obtained when all the data were included, regardless of LNG applicability and LA certainty (so the last update from the analyses in Sections 4.1 or 4.2). Samples of sizes 10, 100, 1000, 10000, and 100000 were drawn with replacement from each posterior distribution to compare to the full distribution, which was estimated using 2.5×10^6 samples. An empirical cumulative distribution function (CDF) can be calculated from each of these samples. An empirical CDF is a CDF calculated from data; the word "empirical" emphasizes that the data form an approximation to the CDF of the random variable governing the data. Our leak frequency model results in a distribution for each model parameter, but we cannot characterize that distribution with an analytical function, so we instead estimate it empirically. The empirical CDFs calculated using different sample sizes can be plotted on the same axes for comparison to each other; visible differences between the CDFs can be attributed to sampling uncertainty in the estimate. When increasing the sample size no longer results in a perceivable difference in CDFs, we conclude that our sample size is large enough to accurately empirically estimate the distribution returned by our model for that parameter (i.e. sampling uncertainty has been sufficiently reduced).

Plots were generated of the empirical cumulative distribution function (CDF) of each sample against the full posterior distribution (2.5×10^6 samples). If the CDFs generated with fewer samples are indistinguishable from posterior distributions generated with the full 2.5×10^6 samples, then convergence with respect to sampling uncertainty has likely been reached.

Figure 4-13 shows plots of empirical CDFs estimated using samples of size 100 against the posterior distributions for each parameter in the model for heat exchangers. Ten such samples of size 100 were drawn because repeated sampling can also help assess convergence; if the empirical CDF changes with each unique sample of size 100, then 100 is not a sufficient sample size. Additionally, if only one sample were drawn, it could match the posterior due to random chance. If repeated samples result in empirical CDFs that match the posterior well, it is likely not random chance.

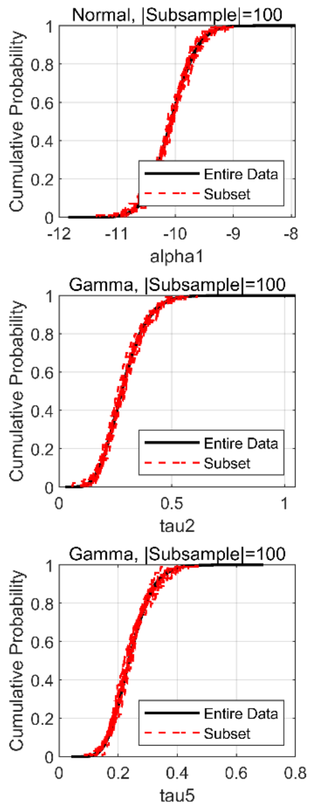

Figure 4-13 Heat exchanger convergence analysis, sample size = 100

Since the empirical CDFs in Figure 4-13 do not match up closely with the posterior distributions (evidenced by the distance between the dashed red curves and the black curve), 100 samples is not adequate to characterize the posteriors. However, this process was repeated using sequentially larger sample sizes. Figure 4-14 shows 10 empirical CDFs from samples of size 10000 with the posterior distribution. These 10 empirical CDFs tightly match up with the posterior distribution. This suggests that the posterior distributions are well characterized by 10000 samples, so the 2.5×10^6 samples used to generate the full empirical CDFs was more than adequate. While we only present the results for the heat exchanger, this trend holds across all nine components. Plots demonstrating the sufficiency of 10000 samples for all components are included in Appendix F.

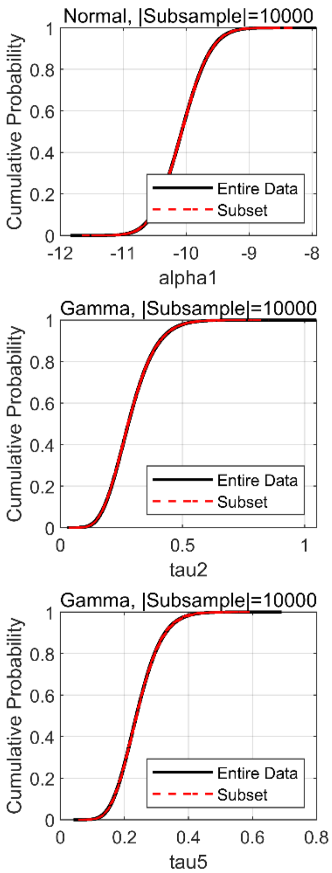

Figure 4-14 Heat exchanger convergence analysis, sample Size = 10000

5. RESULTS AND CONCLUSIONS

The leak frequency prediction model described in Section 3.2 was studied extensively for sensitivity to different types of data (Section 4.1.1), analyst assumptions (Section 4.1.2), prior distribution definitions (Section 4.2.1 through 4.2.3), and model form (Section 4.2.4). A convergence study was presented (Section 4.3) to demonstrate that model results are not significantly affected by sampling uncertainty.

The data sensitivity analyses show that the model can be fit using LNG-Specific and LNG-Applicable data. It can also be fit using Strictly-Generic data, but this was shown to increase uncertainty in leak frequency estimates, potentially because such data may not be relevant for LNG systems. These analyses also show that the effect of data with unknown leak areas (that also cannot be assumed to represent ruptures) is different based on component; inclusion of these data sometimes increases leak frequency estimates and sometimes decreases them. Based on these studies, it is recommended to use all LNG-Applicable data, including the data for which uncertain leak areas were assumed. However, results for joints and vaporizers should be used with caution due to lack of data; estimates for joints are very conservative and estimates for vaporizers are very non-conservative.

The sensitivity analyses related to the prior distributions show that, generally, the median estimated leak frequency is not sensitive to mild changes in prior assumptions. This supports use of the median as a point-estimate for leak frequency. Uncertainty around the median, however, is more sensitive to prior assumptions. Though this does not preclude use of these uncertainties, it should be noted that, for some components, this means that the uncertainty estimates should be expected to change if more data becomes available; the estimates are sensitive to prior assumptions because there are insufficient data.

Recommended leak frequency distributions for all component types and leak sizes are shown in Table 5-1. The full models are plotted in Figure 5-1 through Figure 5-9 with the data set used to fit the models. Note that both hoses (Figure 5-3) and vaporizers (Figure 5-8) demonstrate a positive slope relating leak size to leak frequency. In the cases of hoses, the data for individual leak sizes covers a wide range of leak frequencies, so there may be more inherent variation components in this category. The quantity of data for hoses is low, however, so we cannot know whether additional data would reverse this trend or reinforce it. For vaporizers, there are only two data points and no leak size with more than one data point. This means that the positive trend for vaporizers could easily change with the addition of even one more point. This is why it is recommended that vaporizer results be used cautiously.

Table 5-1 Recommended point-estimates and uncertainty intervals for predicted LNG leak frequencies by component and fractional leak size

Component	Leak Size	5th	Median	95th	Component	Leak Size	5th	Median	95th
Flange and Gasket	0.0001	1.13E-05	4.18E-05	1.14E-04	Pipe	0.0001	3.08E-07	2.67E-06	2.30E-05
	0.001	3.52E-06	2.26E-05	1.82E-04		0.001	1.39E-07	1.44E-06	1.52E-05
	0.01	2.84E-07	1.40E-05	7.14E-04		0.01	1.17E-07	7.86E-07	5.22E-06
	0.1	8.81E-08	8.68E-06	7.30E-04		0.1	4.59E-08	4.25E-07	3.91E-06
	1	4.03E-08	5.24E-06	5.40E-04		1	1.15E-08	2.30E-07	4.63E-06
Heat Exchanger	0.0001	5.41E-04	2.34E-03	1.22E-02	Valve	0.0001	2.40E-05	8.43E-05	2.48E-04
	0.001	1.03E-04	8.93E-04	7.27E-03		0.001	8.76E-06	4.20E-05	2.18E-04
	0.01	3.11E-05	3.24E-04	3.22E-03		0.01	3.54E-06	2.16E-05	1.53E-04
	0.1	2.69E-06	1.17E-04	5.14E-03		0.1	4.72E-07	1.18E-05	2.69E-04
	1	3.13E-06	4.18E-05	6.27E-04		1	2.34E-07	6.42E-06	1.29E-04
Hose	0.0001	4.49E-07	1.52E-06	5.13E-06	Vaporizer	0.0001	1.27E-04	8.19E-03	5.24E-01
	0.001	3.16E-06	7.89E-06	1.99E-05		0.001	1.24E-03	2.63E-02	5.57E-01
	0.01	4.38E-08	4.13E-05	3.82E-02		0.01	1.15E-02	8.46E-02	6.23E-01
	0.1	4.65E-05	2.14E-04	1.00E-03		0.1	8.65E-02	2.72E-01	8.57E-01
	1	2.96E-06	1.10E-03	4.34E-01		1	2.79E-01	8.75E-01	2.75E+00
Joint	0.0001	9.89E+02	3.51E+04	1.25E+06	Vessel	0.0001	8.18E-05	4.77E-04	3.41E-03
	0.001	3.20E+01	4.77E+02	7.09E+03		0.001	3.69E-06	1.39E-04	5.25E-03
	0.01	9.98E-01	6.46E+00	4.18E+01		0.01	1.65E-06	3.90E-05	9.14E-04
	0.1	2.78E-02	8.76E-02	2.76E-01		0.1	2.03E-07	1.10E-05	5.80E-04
	1	4.32E-04	1.19E-03	3.26E-03		1	1.67E-08	3.05E-06	5.77E-04
Loading Arm	0.0001	4.73E-04	1.99E-01	1.04E+01					
	0.001	1.67E-04	1.23E-02	1.59E-01					
	0.01	1.92E-05	7.45E-04	7.87E-03					
	0.1	9.80E-06	3.29E-05	2.58E-04					
	1	7.22E-09	3.03E-06	4.44E-04					

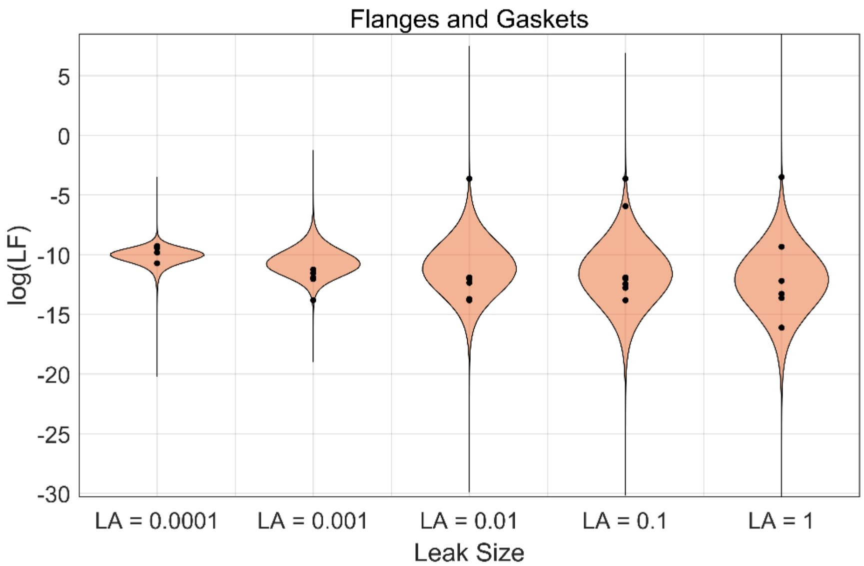

Figure 5-1 Final model for flanges and gaskets

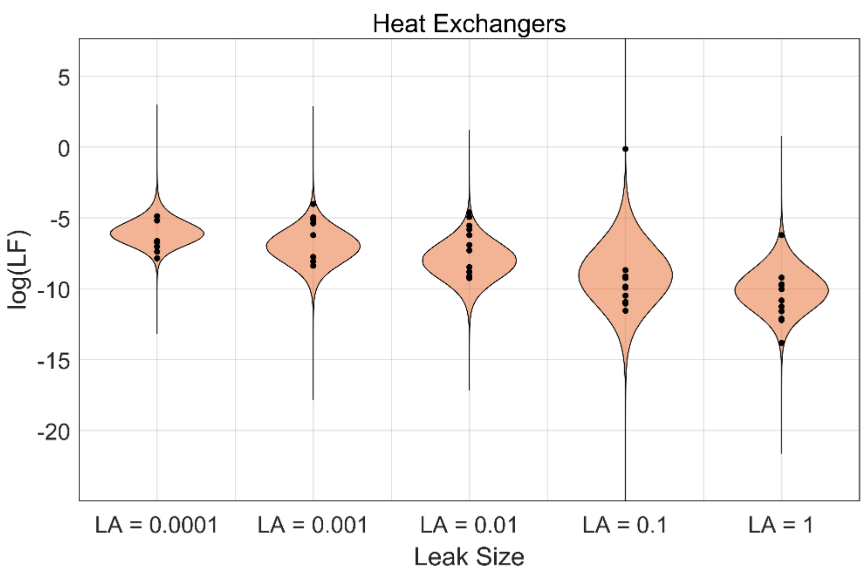

Figure 5-2 Final model for heat exchangers

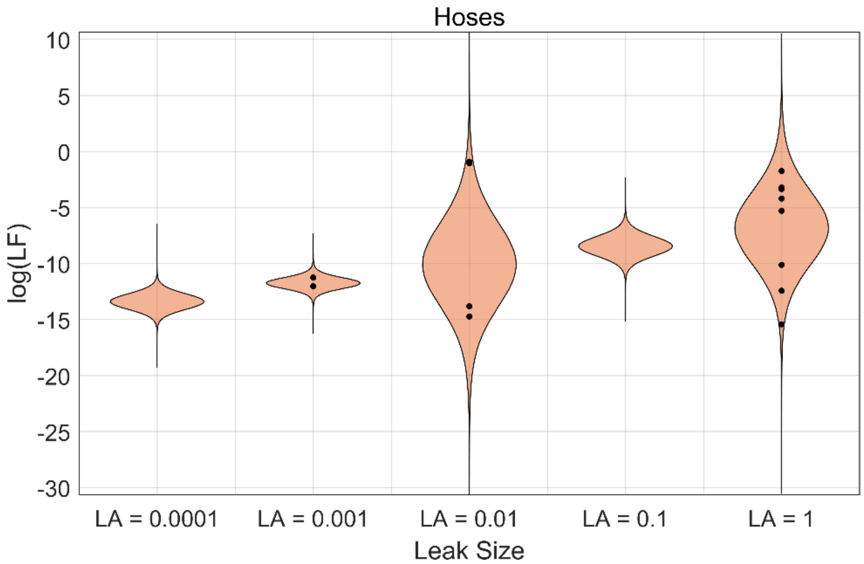

Figure 5-3 Final model for hoses

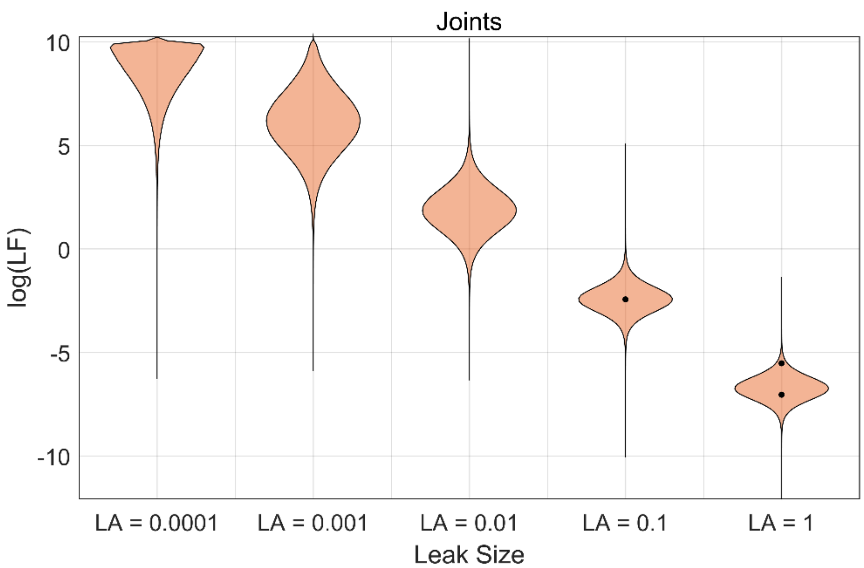

Figure 5-4 Final model for joints

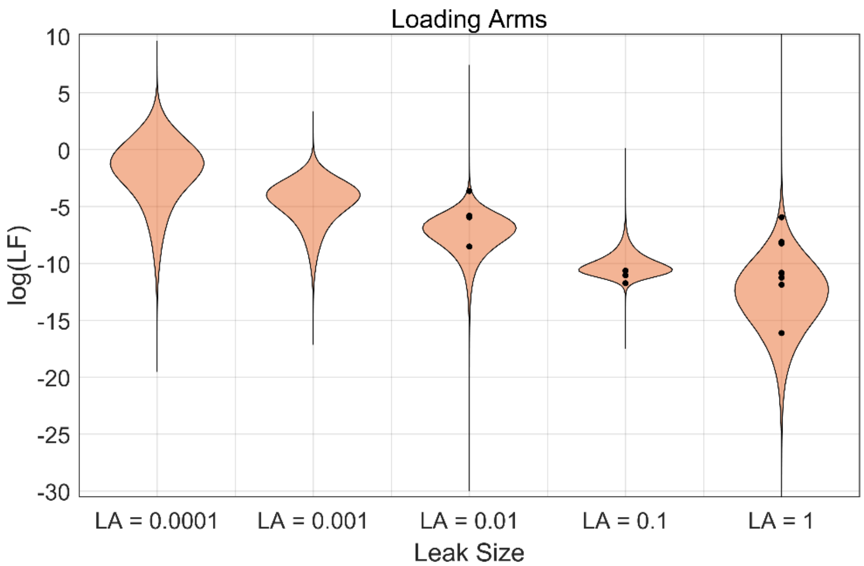

Figure 5-5 Final model for loading arms

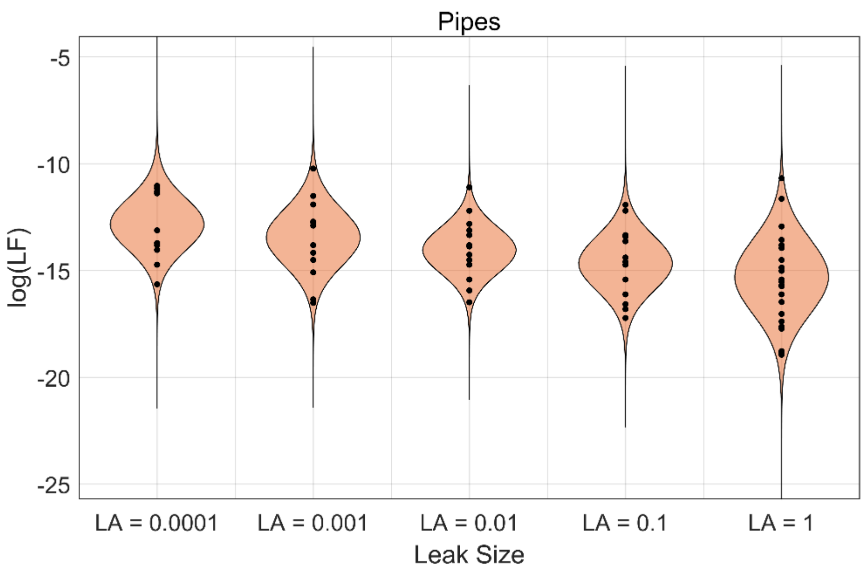

Figure 5-6 Final model for pipes

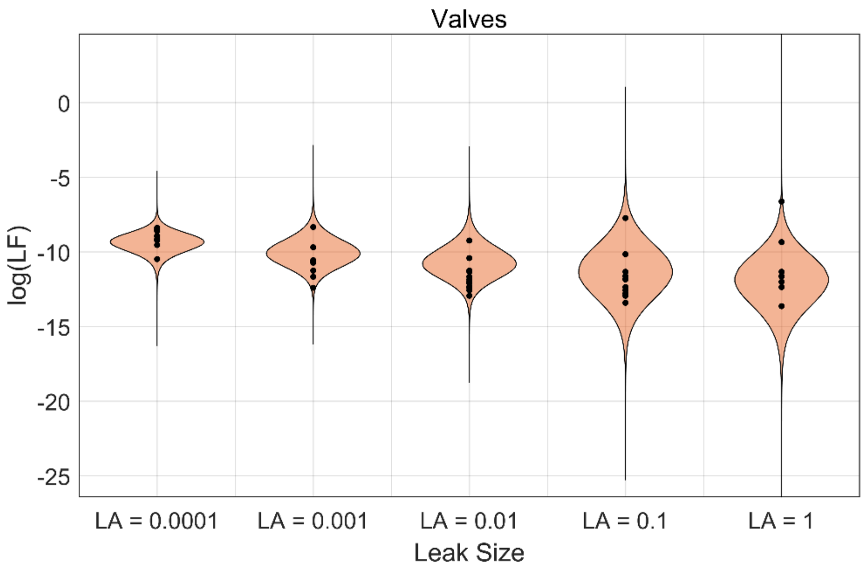

Figure 5-7 Final model for valves

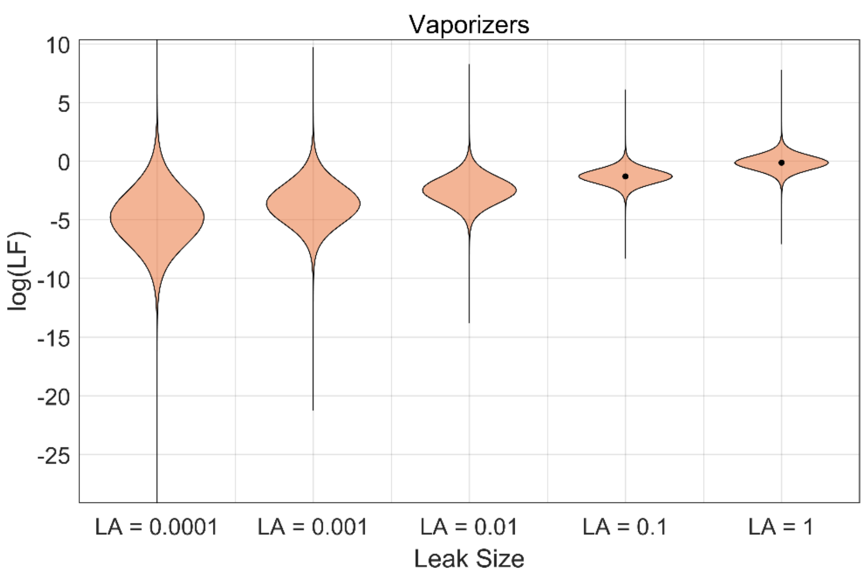

Figure 5-8 Final model for vaporizers

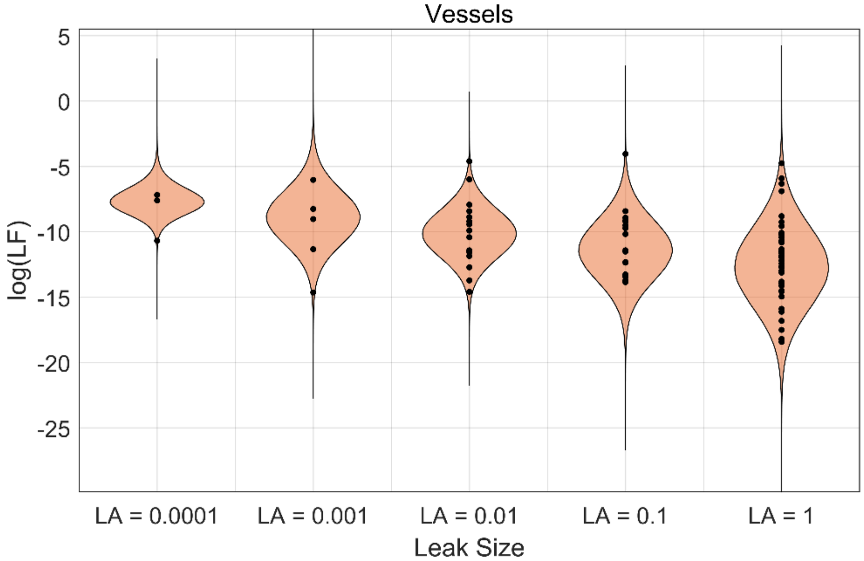

Figure 5-9 Final model for vessels

6. REFERENCES

[1] S. Kaplan and J. B. Garrick, "On the Quantitative Definition of Risk," *Risk Analysis,* vol. 1, no. 1, pp. 11-27, 1981.

[2] J. LaChance, W. Houf, B. Middleton and L. Fleur, *Analysies to Support Development of Rik-Informed Separation Distances for Hydrogen Codes and Standards,* Albuquerque, New Mexico: Sandia National Laboratories, 2009, SAND2009-0874.

[3] B. D. Ehrhart and E. S. Hecht, "Hydrogen Risk Assessment Models (HyRAM) Version 3.0 Technical Reference Manual," Sandia National Laboratories, Albuquerque, New Mexico, 2020, SAND2020-10600.

[4] A. Glover, A. Baird and D. Brooks, "Final Report on Hydrogen Plant Hazards and Risk Analysis Supporting Hydrogen Plant Siting Near Nuclear Power Plants," Sandia National Laboratories, Albuquerque, New Mexico, 2020, SAND2020-10828.

[5] Gas Technology Institute, "Statistical Review and Gap Analysis of LNG Failure Rate Table," Des Plaines, IL, 2017, Final Report DTPH5615T00008..

[6] A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari and D. Rubin, Bayesian Data Analysis, Boca Raton: Chapman and Hall, 2013.

[7] J. Kruschke, Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan, Academic Press, 2015.

[8] M. Plummer, *JAGS: Just Another Gibbs Sampler,* mcmc-jags.sourceforge.net, 2012.

[9] R Core Team, *R: A Language and Environment for Statistical Computing,* Vienna, Austria: R Foundation for Statistical Computing, 2013.

[10] M. Plummer, A. Stukalov and M. Denwood, *rjags: Bayesian Graphical Models Using MCMC,* 2019.

[11] H. Hoffman, *Violin Plot,* MATLAB Central File Exchange, 2020.

[12] F. R. Hampel, "The Influence Curve and Its Role In Robust Estimation," *Journal of the American Statistical Association,* vol. 69, no. 346, pp. 383-393, 1974.

[13] F. Massey, "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American Statistician Association,* vol. 46, no. 253, pp. 68-78, 1951.

[14] S. Weisberg, "Testing for Curvature," in *Applied Linear Regression*, Hoboken, John Wiley & Sons, Inc., 2014, pp. 212-213.

APPENDIX A. DATA ASSUMPTIONS

The appendix documents the data sources and how the data were processed for the analyses in this report. The authors of [5] outline several redundancies in the database they collected. As a result, we made the following modifications to the provided data (see Table A-1 for names used to refer to each data set).

- General Notes
  - Since RIVM BEVI '09 is supposed to replace TNO PURPLE '05, when RIVM BEVI and TNO PURPLE supply data about the same component, we only include values from RIVM BEVI '09 since it is more recent
- For Vessels, Flanges and Gaskets, Heat Exchangers, and Vaporizers
  - No redundancies aside from ones mentioned in general notes
- For Expansion Joints
  - PNL PRSP '82 cites WELKER '76 and SAI '75, which is possibly based on the 1975 WASH-1400 Rasmussen Report (Section 4.7 of [5]). LEES '12 is also based on the WASH-1400 Rasmussen Report. Thus, we remove the WELKER '76 datum and the LEES '12 datum and keep the PNL PRSP '82 data because there are more data in that report
- For Valves
  - SAI '75, WELKER '76, and LEES '12 valve data all rely on the 1975 WASH-1400 Rasmussen Report (Section 4.6 of [5]), so we use just the LEES '12 valve data and remove the SAI '75 and WELKER '76 data since there are more data in LEES '12
- For Hoses & Loading Arms
  - LNE '09 cites RIVM BEVI '09 (Section 4.5 of [5]), so we remove the LNE '09 data
- For Pipes
  - We must remove any data point that reports leaks in "per section" since rates are not in the correct unit (no length of section is given)
    - This includes the pipe data from PNL PSRP '82, WELKER '76, LEES '12, and TNO RED '05

The data sets in this analysis are listed along with the components they contain and any assumptions made for this analysis. Unless an assumption is stated, the fractional leak area (LA) was either calculated from the data (using $LA = \frac{area\ of\ leak}{cross\text{-}sectional\ area\ of\ component}$) or provided from the data (i.e. some reports defined a leak as a hole in a component with effective diameter 10% of the component diameter, so $LA = 0.1^2 = 0.01$).

Table A-1 Data Sources and Analysis Assumptions

Data Set	Components (and component types)	LA Assumptions	Source
API 581 '16	Heat exchanger (HEXSS, HEXTS, Fin fans) Vessel (Distillation column, Tank, Tank bottom)	For all components, we assumed the following assignments: • Small Leak -> LA = 0.001 • Medium Leak -> LA = 0.01 • Large Leak -> LA = 0.1 • Rupture -> LA = 1	API, Risked-Based Inspection Methodology, Recommended Practice 581, Third Edition, American Petroleum Institute, 2016.
CCPS '89	Hose Pipe (metal straight sections) Valve (manual) Vessel (metallic, atmospheric/pressurized)	Pipe: catastrophic rupture -> LA = 1 All other components: Rupture -> LA =1	CCPS Guidelines for Process Equipment Reliability Data with Data Tables, American Institute of Chemical Engineers, Center for Chemical Process Safety, New York, NY, 1989.
EGIG '18	Pipe (separated by diameter ranges)	The report provided the following leak size definitions: • Pinhole/crack: the effective diameter of the hole is smaller than or equal to 2 cm • Hole: the effective diameter of the hole is larger than 2 cm and smaller than or equal to the diameter of the pipe • Rupture: the effective diameter of the hole is larger than the pipeline diameter. Based off of these definitions and the pipe diameter, we calculated a range of possible	EGIG Gas Pipeline Incidents, 10th Report of the European Gas Pipeline Incident Data Group (1970 – 2016), Doc. Number EGIG VA 17.R.0395, March 2018.

		LA values and assigned the LA bin that best captured the calculated range	
GRI LNG FRD '81	Pipe Vaporizer	For pipes we conservatively assigned "failure" -> LA = 1 For vaporizers, the rate was given for "leak or rupture" which we moderately assigned as LA = 0.1	Johnson, D.W., and Welker, J.R., Applied Technology Corp. Development of an Improved LNG Plant Failure Rate Database, Final Report for Gas Research Institute, GRI-80/0093, 1981.
HSE FRED NOV '17	Flange & Gasket Hose (reported as hoses and coupling, separated by facility) Loading Arm (ship hardarms for liquefied gas, separated by how many arms were in use) Pipe (aboveground and non-aboveground, separated in pipe diameter ranges) Valve Vessel (Large atmospheric tank, LNG and generic refrigerated ambient pressure single-walled vessel, generic refrigerated ambient pressure double-walled vessel, LNG refrigerated double-walled vessel, LNG refrigerated full containment vessel, liquid oxygen refrigerated single-walled vessel, generic pressure vessel, chlorine pressure vessel, LPG pressure vessel, tank container)	For hoses, we assumed • Guillotine -> LA = 1 • 15 mm diameter hole -> LA = 0.01 (calculated from typical diameter of a chemical hose) • 5 mm diameter hole -> LA = 0.001 For valves • Valve spray -> LA = 0.001, justified by report definition For loading arm • Guillotine -> LA = 1 For flange and gasket • Failure -> LA = 1 For aboveground piping • Rupture -> LA = 1 • Large hole -> LA = 0.1	HSE Failure Rate and Event Data for use within Risk Assessments, UK, November 2017.

		• Small hole -> LA = 0.001 • Pin -> LA=0.0001 • For other piping data, we were able to calculate LA For atmospheric tank, LNG, generic refrigerated vessels, and tank containers • Catastrophic failure -> LA = 1 • Effective hole diam 1m -> LA = 0.1 (assumed conservatively) • Effective hole diam 0.3m -> LA = 0.01 (assumed conservatively) • Release of vapor only-> Not useful for analyzing leak events, so excluded this data from analysis For liquid oxygen vessels • Catastrophic failure-> LA=1 • Effective hole diam 0.4 m -> LA = 0.1 (assumed conservatively) • Effective hole diam 0.12m -> LA = 0.01 (assumed conservatively) For generic, chlorine, and LPG vessels • Catastrophic release & BLEVE -> LA = 1 (chose the median value) • 50 mm and 25 mm hole -> LA = 0.1 (assumed conservatively)	

		• 13 mm and 6 mm hole -> LA = 0.01 (assumed conservatively)	
INL CHEM '95	Heat Exchanger (Tube, shell) Hose Pipe Valve (Manual, check, motor, control, solenoid) Vessel (Vessel Tank) (For each component, separate frequencies were reported for chemical processes and compressed gas, both of which were included in the analysis)	For all components, we assumed • Rupture -> LA = 1 • Leakage -> LA = 0.01	Alber, T.G., Hunt, R.C., Fogarty, S.P., Wilson, J.R., Idaho Chemical Processing Plant Failure Rate Database, Idaho National Engineering Laboratory, NEL-95/0422, August 1995.
INL NUC '07	Heat Exchanger (Shell, tube) Pipe (Non-service water) Valve (Manual, check, air, hydraulic, motor, solenoid) Vessel (Pressurized tank, unpressurized tank)	In this report "External leakage is subdivided into two modes: small (ELS), covering 1 to 50 gallons per minute (gpm) and large (ELL), covering > 50 gpm (for water systems)," so LA could not be inferred. We made the following assumptions for each component: • Small Leak -> LA = 0.01 • Large leak -> LA = 0.1	Plants, N. P. Industry-average performance for components and initiating events at US commercial nuclear power plants, Idaho National Engineering Laboratory, NUREG/CR- 6928 (2007).
IOGP 434-1	Flange & Gasket Heat Exchanger (Shell and tube, plate, air cooled) Pipe (steel process) Valve (manual, actuated) Vessel (Process vessel)	For heat exchangers and vessels, LA was calculated using the inlet cross-sectional area provided	IOGP Risk Assessment Data Directory, Process Release Frequencies, International Association of Oil and Gas Producers Report No. 434 – 1, 2019.
IOGP 434-3	Vessel (Pressure storage vessel, pressure small container vessel, single/double containment new/existing refrigerated storage tank)	Cross-sectional area not reported, so we assumed the following LA assignments: • For all components ○ Catastrophic rupture -> LA = 1 • For pressure storage vessel	IOGP Risk Assessment Data Directory, Storage Incident Frequencies, International Association of Oil and Gas Producers Report No. 434 – 3, March 2010.

		○ 1-3 mm -> LA = 0.0001 ○ 3-10 mm -> LA = 0.001 ○ 10-50 mm -> LA = 0.01 ○ 50-150 mm -> LA = 0.1 ○ >150 mm -> LA =1 • For small container ○ 1-3 mm -> LA = 0.001 ○ 3-10 mm -> LA = 0.01 ○ >150/Catastrophic -> LA = 1	
KGSC '06	Vessel (Full containment LNG tanks, four different modifications to membrane LNG tanks)	No cross-sectional area provided, so we conservatively assigned the leak frequencies to LA = 0.1	Lee, S. R. "Safety comparison of LNG tank designs with fault tree analysis", International Gas Union 23rd World Gas Conference proceedings, Amsterdam, Holland. 2006.
KJCE '05	Vessel (PC membrane)	No cross-sectional area provided, so we conservatively assumed the failure rate as LA = 1	Kim, Hyo, Koh, Jae-Sun, Kim, Youngsoo, University of Seoul Department of Chemical Engineering, and Theofanous, Theofanius, University of California at Santa Barbara Center for Risk Studies and Safety, "Risk Assessment of Membrane Type LNG Storage Tanks in Korea – based on Fault Tree Analysis," Korean Journal of Chemical Engineering, Vo. 22., No. 1, 1-8, 2005.
LEES '12	Flange & Gasket Joint (expansion joints)	For valves: • Rupture -> LA =1	Mannan, Sam, Lees' Loss Prevention in the Process

	Loading Arm Pipe Valve (manual, air, motor, solenoid) Vessel (pressure vessel, process vessel, separated by fluid)	For (Expansion) joints • Serious leak -> LA=0.1 For loading arm: • Catastrophic Failure -> LA=1 For flanges and gaskets: • Serious leak -> LA = 0.1 For pipes: • Rupture -> LA = 1 • Guillotine -> LA=1 For vessels: • Lees classifies catastrophic and disruptive failure as the worst events • Catastrophic/disruptive failure -> LA = 1 • To be conservative, when an interval was given we picked the upper bound	Industries – Hazard Identification, Assessment and Control, Third Edition, Volume 3, Appendix 14, Elsevier, Inc. 2005.
LNE '09	Heat Exchanger (Pipe, plate by pressure) Hose (LPG) Loading Arm Pipe (aboveground and underground) Vessel (Pressure tank, process tank, atmospheric storage tank of four types, storage tank)	For heat exchangers and vessels we assigned: • Small leak -> LA = 0.001 • Medium leak -> LA = 0.01 • Large leak -> LA = 0.1 • Rupture -> LA = 1 For loading arms we assigned • Leakage-> LA = 0.01 • Rupture-> LA=1 For aboveground pipes: • Small -> LA = 0.001 • Medium -> LA = 0.01 • Large -> LA = 0.1 • Rupture -> LA=1	LNE Handbook of Failure Frequencies, and Appendix, Safety Report, Flemish Government LNE Department Environment, Nature and Energy Policy Unit Safety Reporting Division, 2009.

		For underground pipes: • Crack -> LA = 0.001 • Hole -> LA = 0.01 • Rupture -> LA = 1	
NFPA 59A '19	Flange & Gasket Heat Exchanger Joint (Expansion joints) Loading arm (truck transfer, ship transfer) Pipe (by diameter ranges) Vessel (single containment atmospheric storage tank, double containment atmospheric storage tank, full containment and membrane storage tank, other atmospheric storage tanks, pressurized storage vessel, process vessel, distillation column, condenser)	For heat exchangers and pipes: • Catastrophic failure -> LA = 1 • 10 mm hole -> LA = 0.01 For valves, hoses, loading arms and joints: • Rupture -> LA=1 For flanges and gaskets: • Failure-> LA=1 For pipes: • Catastrophic rupture -> LA =1 For vessels: • Catastrophic rupture -> LA = 1 • When a specific hole diameter was given, there was no way to determine LA, so we conservatively assigned LA = 0.01	NFPA Standard for the Production, Storage, and Handling of Liquefied Natural Gas (LNG), NFPA 59A, 2019 Edition.
PHMSA HL GTI '20	Pipe (by diameter ranges)	Conservatively assumed rupture -> LA =1	PHMSA Hazardous Liquid Pipeline Accident Data
PHMSA NGT GTI '20	Pipe (by diameter ranges)	Conservatively assumed rupture -> LA =1	PHMSA Natural Gas Transmission Pipeline Incident Data
PNL PSRP '82	Flange & Gasket Hose (flexible metal) Joints (Expansion joints) Loading Arm Pipe Valve (control, check, manual, air-operated) Vaporizer	When given an interval of values, we chose the most conservative end For all components we made the following assignments: • Rupture-> LA=1 • When given a rate of "leak or rupture", we set LA = 0.1 as a	Pelto, P.J., Baker, E.G., Holter, G.M, and Powers, T.B., Analysis of LNG Peakshaving Facility Release Prevention Systems, Pacific Northwest Laboratory, PNL-4153, 1982.

	Vessel (storage tank)	moderately conservative assignment	
RIVM BEVI '09	Heat Exchanger (pipe, plate) Hose Loading Arm Pipe (aboveground, underground, with some diameter ranges) Vessel (tank wagon, distillation column, gas container, process vessel, pressure vessel, storage tank above ground, underground pressurized storage tank, single containment atmospheric storage tank, full containment atmospheric storage tank, membrane tank, mounded atmospheric storage tank, storage tank with protective outer shell)	For loading arms and hoses we assumed: • Rupture -> LA = 1 For pipes we assumed: • Rupture -> LA=1 • 20 mm -> LA = 0.1 (conservative assumption since cross-sectional area not provided) For vessels and heat exchanger we assumed the following (the same assumptions as TNO PURPLE since RIVM BEVI is an update to that document) • Instantaneous release->LA=1 • Continuous 10 min release->LA=1 • Continuous 10 mm release->LA=0.01 • Leak -> LA = 0.01 • To be conservative, added together the 10 pipe ruptures and 1 pipe rupture for heat exchanger pipes • For Distillation Unit Column, added together 10mm rectifying section and 10mm stripping section since they are in the same part of vessel • To be conservative, combined the frequencies from two-walled vessels when each wall was given separately • For tanks, to be conservative we added together instantaneous	RIVM Reference Manual Bevi Risk Assessment, Version 3.2, Module C, the Netherlands National Institute of Public Health and Environment (RIVM), July 1, 2009.

		release frequency with release from largest connection and assigned LA = 1	
SAI '75	Flange & Gasket Joint (Expansion joint) Pipe Valve Vessel (tank)	Conservatively assumed rupture -> LA =1	SAI, LNG Terminal Risk Assessment Study for Oxnard, California, Science Applications, Inc. SAI-75-615-LJ, 1975.
SERCO AEA '04	Vessel (large vessel)	Chose the mean rupture frequency, assumed rupture -> LA = 1	O'Donnell, IJ and Phillips, DW, Serco Assurance, and Winter, PW, AEA Technology, "A New Estimate of the Likelihood of Spontaneous Catastrophic Failure of Pressurized LPG Storage Vessels", Hazards XXVIII, Symposium Series No. 150, IChemE, 2004.
SIGTTO IP4 '96	Hose Loading Arm (Hard arm)	Conservatively assumed rupture -> LA =1	SIGGTO, Accident Prevention – The Use of Hoses & Hard Arm at Marine Terminals Handling Liquefied Gas, Information Paper No. 4, 1996.
TGC '03	Vessel (PC Membrane, steel/pre-stressed concrete double shell tank, single containment tank)	Conservatively assumed failure -> LA =1	Miyazaki, Sinichi, Tokyo Gas Co. Ltd. and Yamada, Yoshihisa, Tokyo Gas Engineering Co., Ltd., "Quantitative Risk Assessment of LNG Above Ground Tanks Based on Past

			Operating Records of LNG Regasification Terminals and Life Cycle Assessment", International Gas Union 22nd World Gas Conference proceedings, Tokyo, June 1-5, 2003.
TNO PURPLE '05	Heat Exchanger Hose Loading Arm Pipe (by diameter) Vessel (Process vessel, pressure vessel, reactor vessel, pressurized tank, atmospheric tank, single containment tank, double containment tank, full containment tank, tank with protective outer shell, in-ground tank, mounded tank)	For heat exchangers, there was little basis for LA assignment. Since "catastrophic rupture of a heat exchanger with the dangerous substance outside the pipes is modelled partly as an instantaneous release and partly as a continuous release of the complete inventory within 10 minutes" (Section 3.A.2.6), for we made the following assignments • Instantaneous release->LA=1 • Continuous 10 min release->LA=1 • Continuous 10 mm release->LA=0.01 • Leak -> LA = 0.01 • To be conservative, added together the 10 pipe ruptures and 1 pipe rupture for heat exchanger pipes For vessels there was also little basis for LA assignment. Since "catastrophic rupture is modelled partly as an instantaneous release and partly as a continuous release within ten min" (Section 3.A.2.3), we made the following assignments • Release to atmosphere -> LA =1 • Instantaneous release->LA=1 • Continuous 10 min release->LA=1	TNO Guidelines for Quantitative Risk Assessment (TNO Purple Book), Committee for the Prevention of Disasters (CPR), National institute of Public Health and the Environment (RIVM), The Netherlands Organization for Applied Scientific Research (TNO). First edition 1999/2005.

		• Continuous 10 mm release->LA=0.01 • To be conservative, combined the frequencies from two-walled vessels when each wall was given separately For piping we assumed • Rupture/Full Bore rupture -> LA = 1	
TNO RED ‘05	Heat Exchanger Pipe Valve (not included in analysis because data were not for leak events) Vessel (pressure vessel, single wall vessel)	There was a little basis for LA calculation, so we made the following assignments. For heat exchangers and vessels: • 10 mm leak -> LA = 0.01 • Basic Failure -> LA = 0.1 • Catastrophic Failure -> LA = 1 For pipes: • Leakage -> LA = 0.01 • Breakage -> LA = 1	TNO Methods for Determining and Processing Probabilities (TNO Red Book), Committee for the Prevention of Disasters (CPR), National Institute of Public Health and the Environment (RIVM), The Netherlands Organization for Applied Scientific Research (TNO). Second edition. 1997/2005.
WELKER ‘76	Flange & Gasket Hose Joint (Expansion joint) Loading Arm (Ship transfer) Valve (motor, manual) Pipe (by cross-section size)	If the rate was for “leak or rupture”, we gave moderately conservative assignment of LA = 0.1 For all components, we assumed o Rupture -> LA =1 o Leakage -> LA = 0.01	Welker, J.R., Brown, L.E., Ice, J.N., Martinsen, W.E., and West, H.H., Fire Safety Aboard LNG Vessels, Final Report DOT-CG-42, 355A, Task #1.

APPENDIX B. DATA SCATTER PLOTS

This appendix contains plots of the data by component and leak size. The model studied in this report is characterized as linear in this space (in the log scale of the plots) with uncertainty. The first section shows the data categorized by applicability to LNG systems and the second section shows the same data categorized by certainty in the leak area bin assignment.

B.1. Data by LNG Applicability

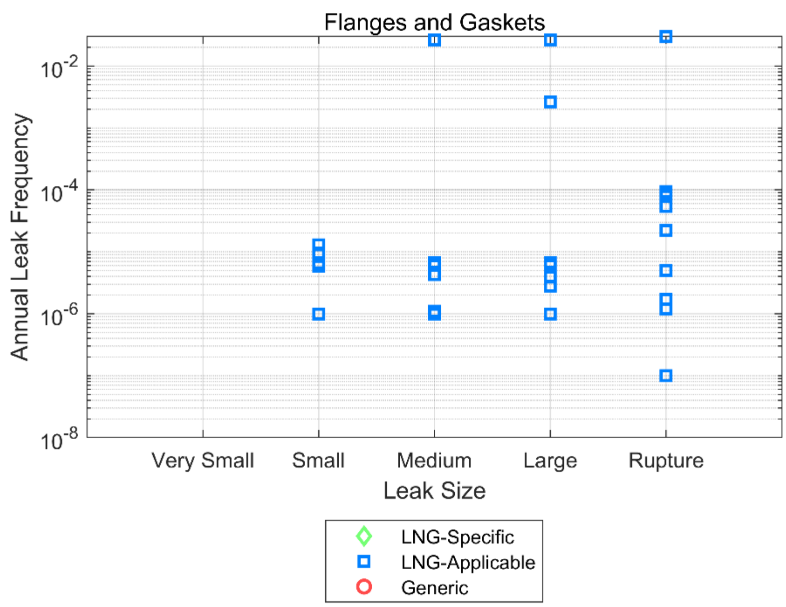

Figure B-1 Scatter plot of flange and gasket data by LNG applicability

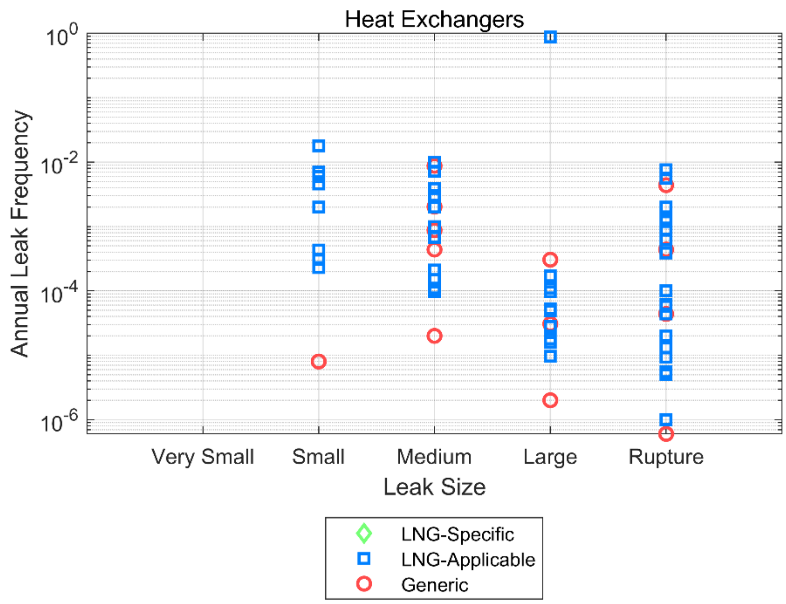

Figure B-2 Scatter plot of heat exchanger data by LNG applicability

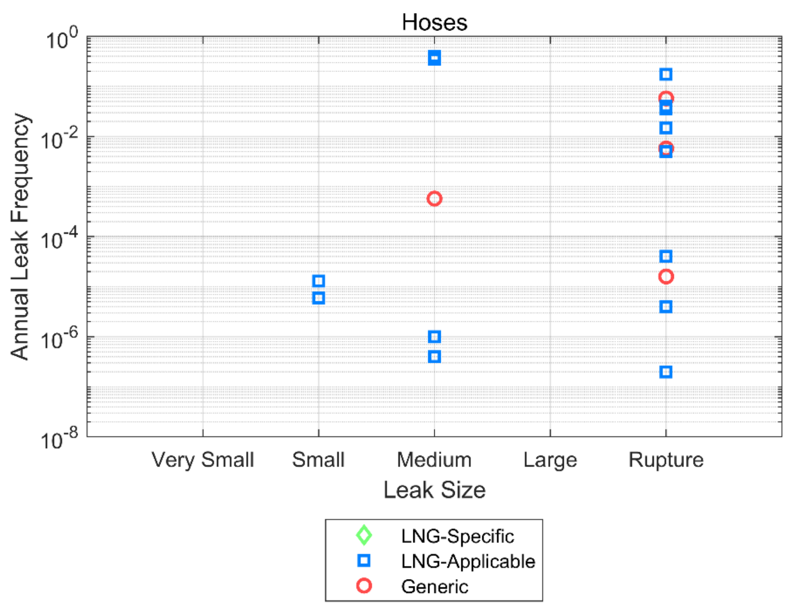

Figure B-3 Scatter plot of hose data by LNG applicability

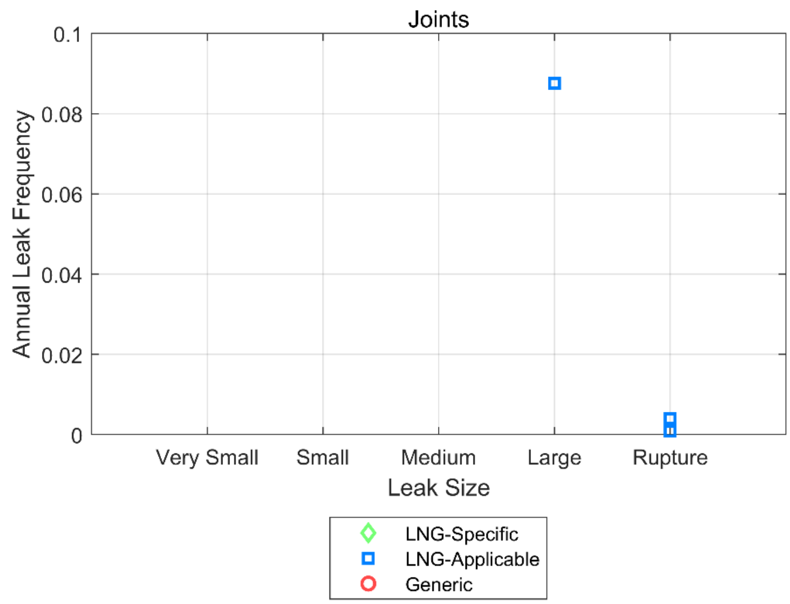

Figure B-4 scatter plot of joint data by LNG applicability

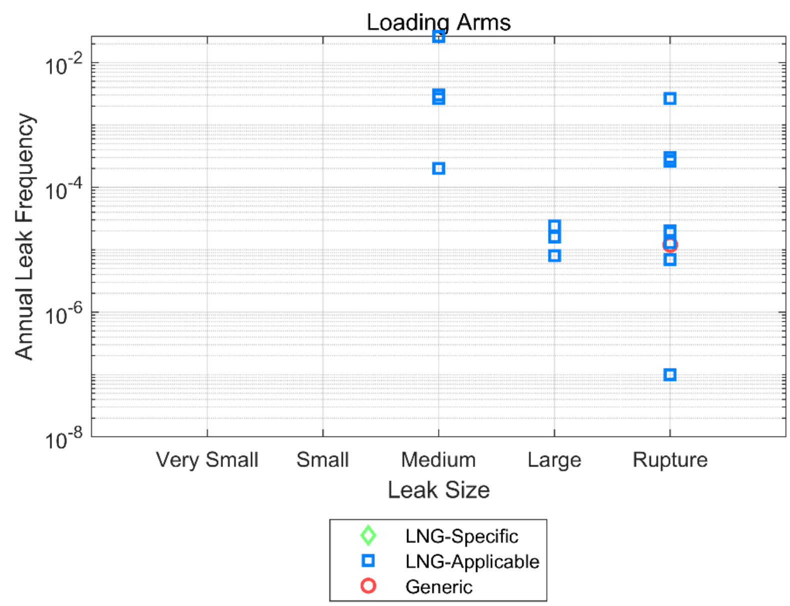

Figure B-5 Scatter plot of loading arm data by LNG applicability

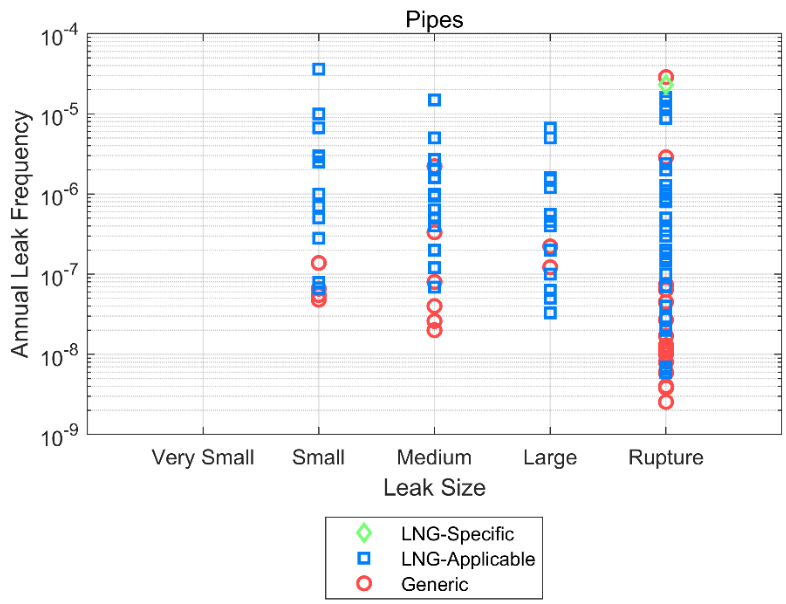

Figure B-6 Scatter plot of pipe data by LNG applicability

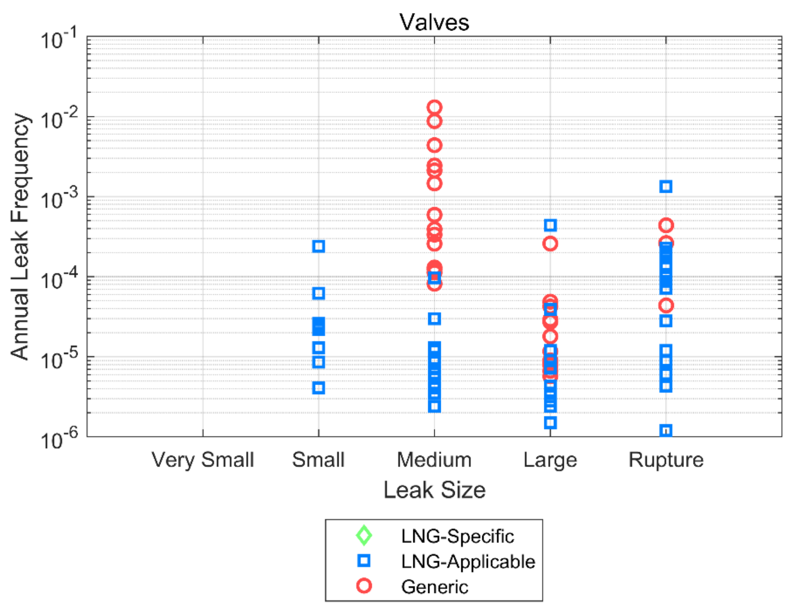

Figure B-7 Scatter plot of valve data by LNG applicability

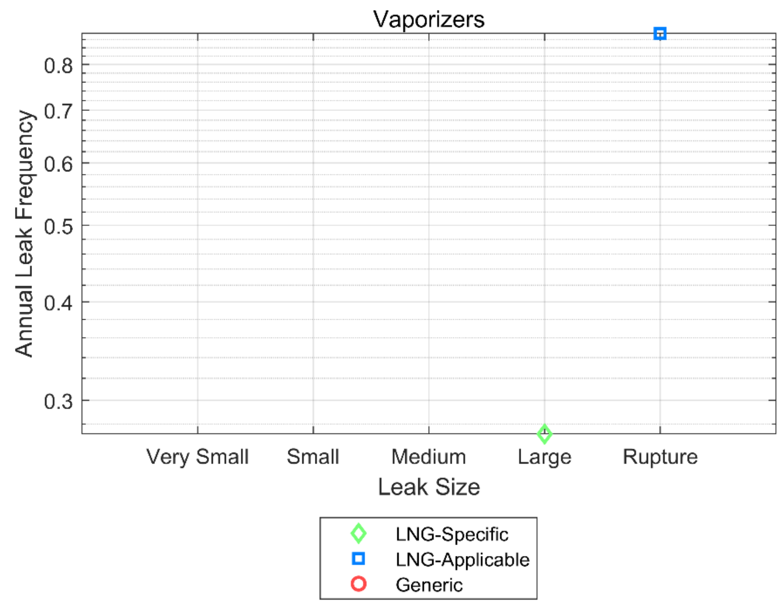

Figure B-8 Scatter plot of vaporizer data by LNG applicability

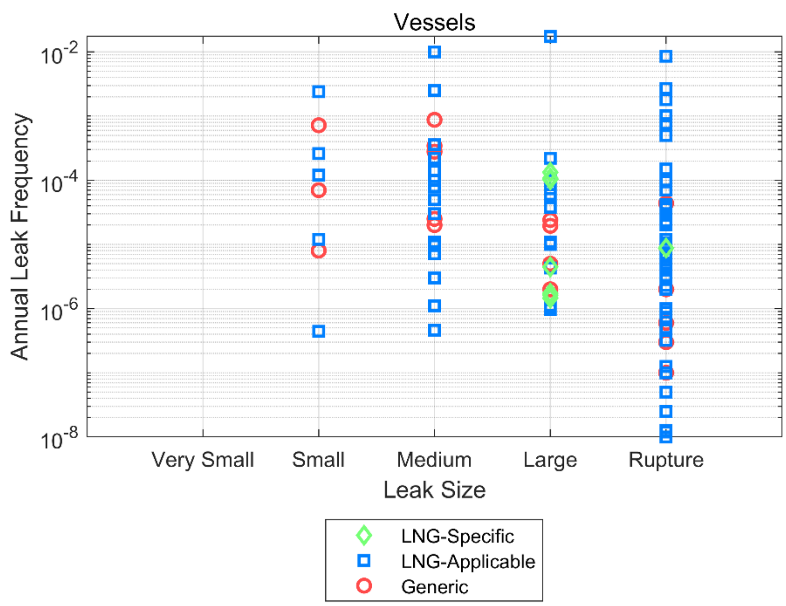

Figure B-9 Scatter plot of vessel data by LNG applicability

B.2. Data by Leak Area Bin Certainty

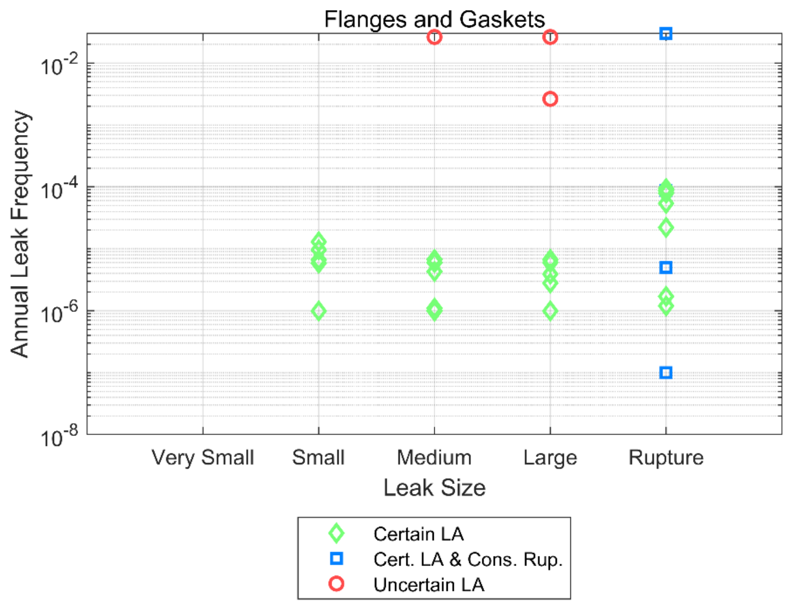

Figure B-10 Scatter plot of flange and gasket data by leak area bin certainty

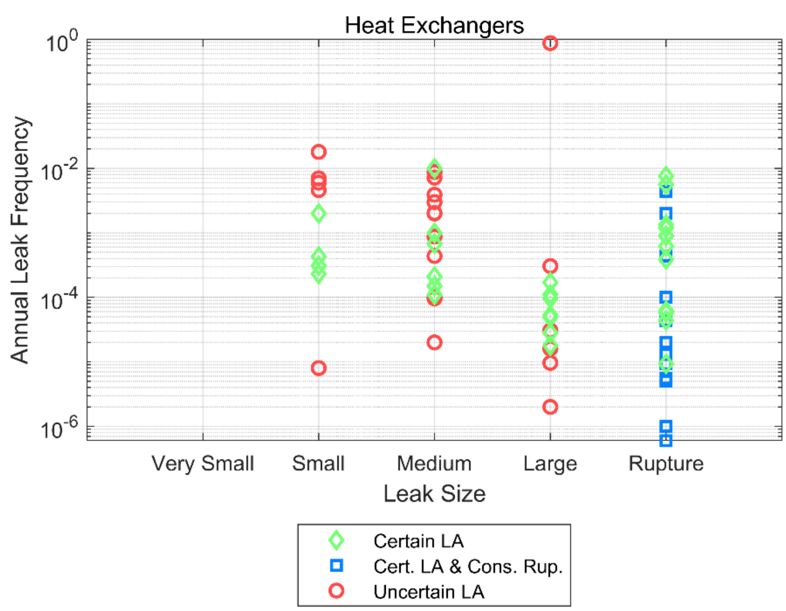

Figure B-11 Scatter plot of heat exchanger data by leak area bin certainty

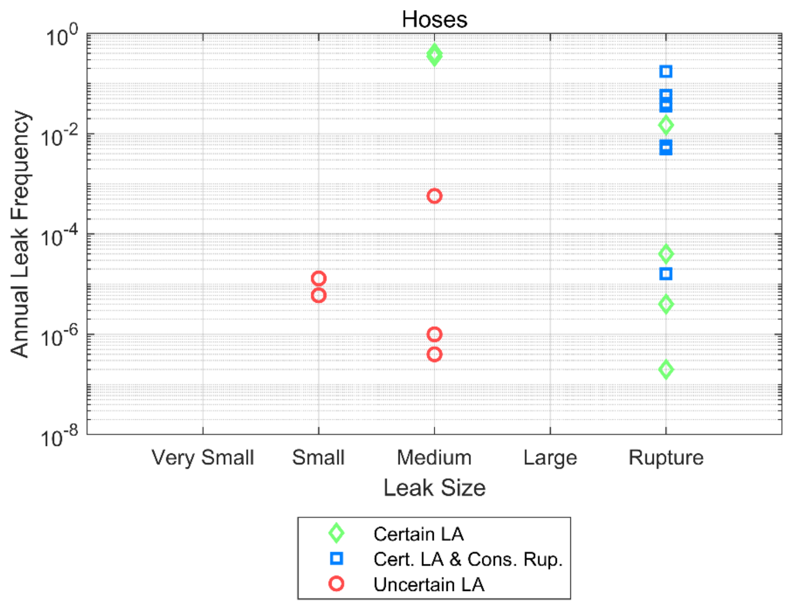

Figure B-12 Scatter plot of hose data by leak area bin certainty

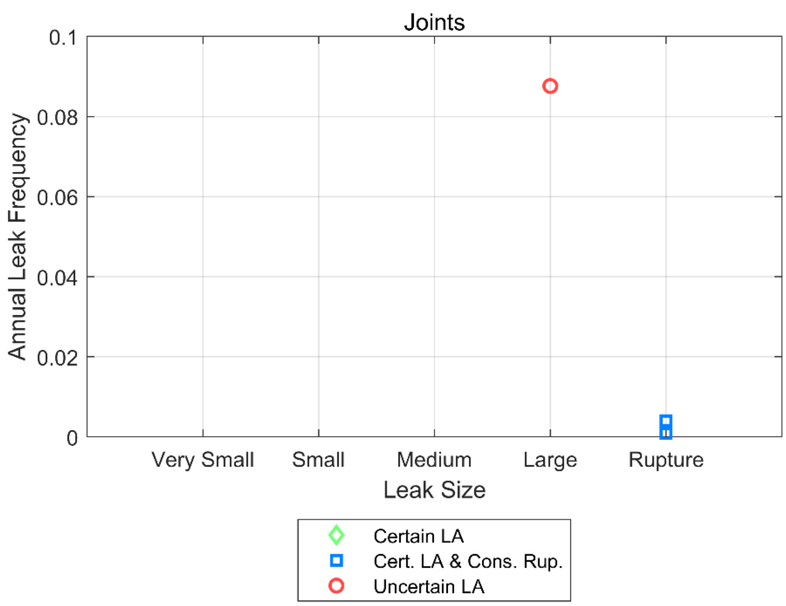

Figure B-13 Scatter plot of joint data by leak area bin certainty

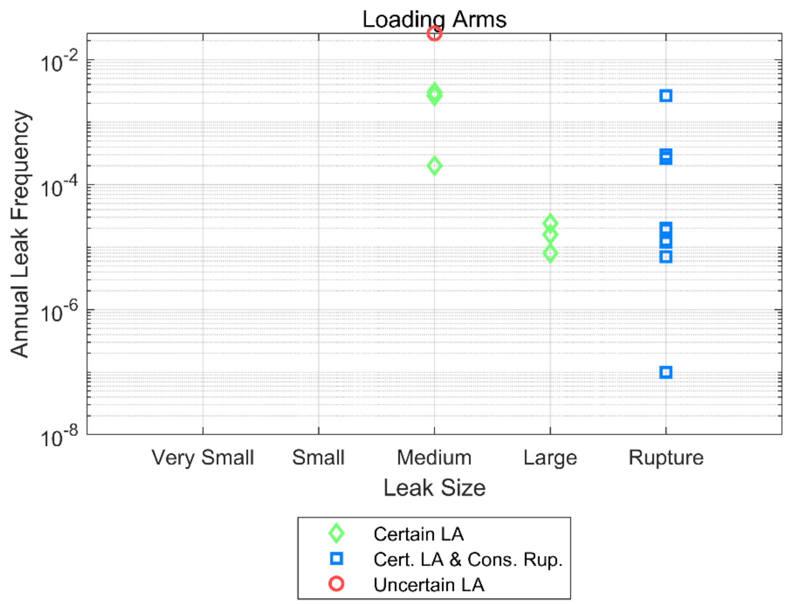

Figure B-14 Scatter plot of loading arm data by leak area bin certainty

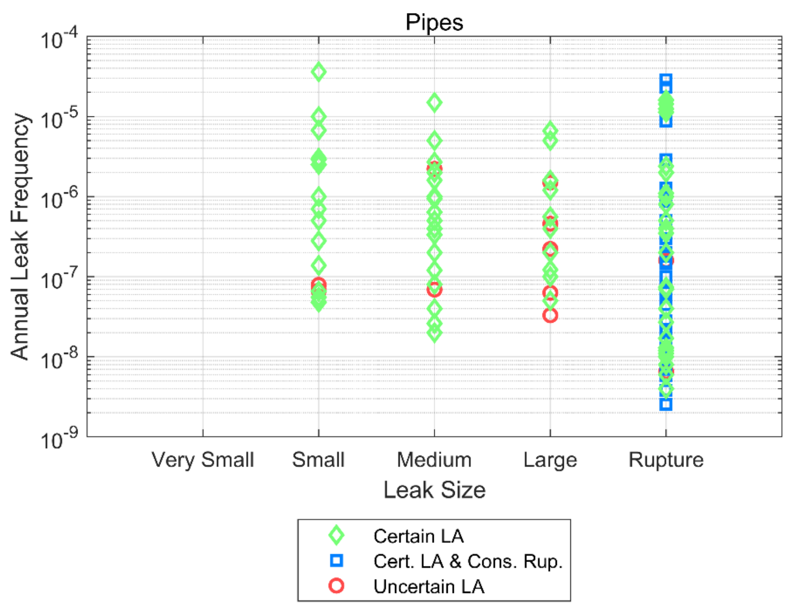

Figure B-15 Scatter plot of pipe data by leak area bin certainty

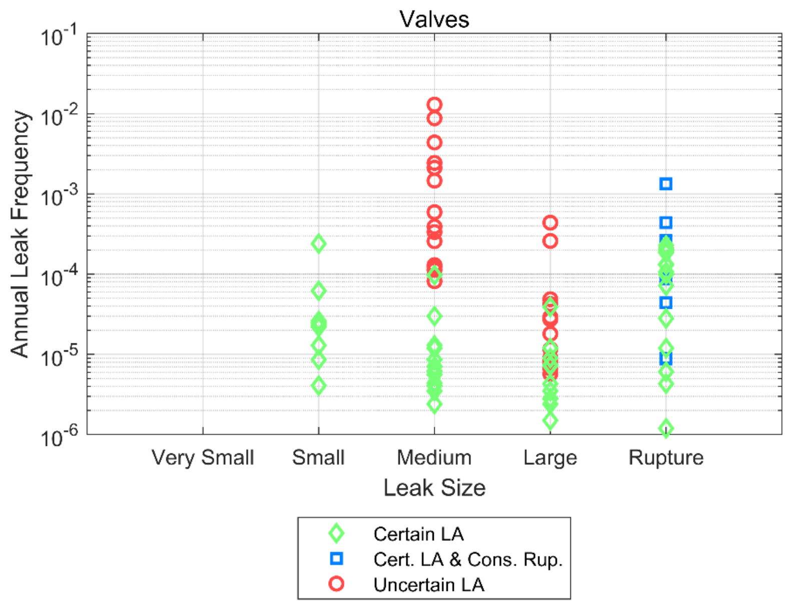

Figure B-16 Scatter plot of valve data by leak area bin certainty

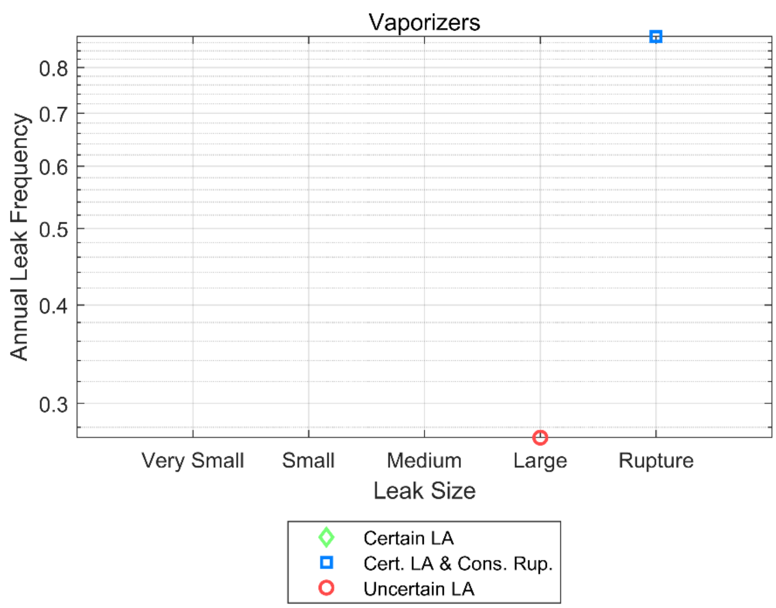

Figure B-17 Scatter plot of vaporizer data by leak area bin certainty

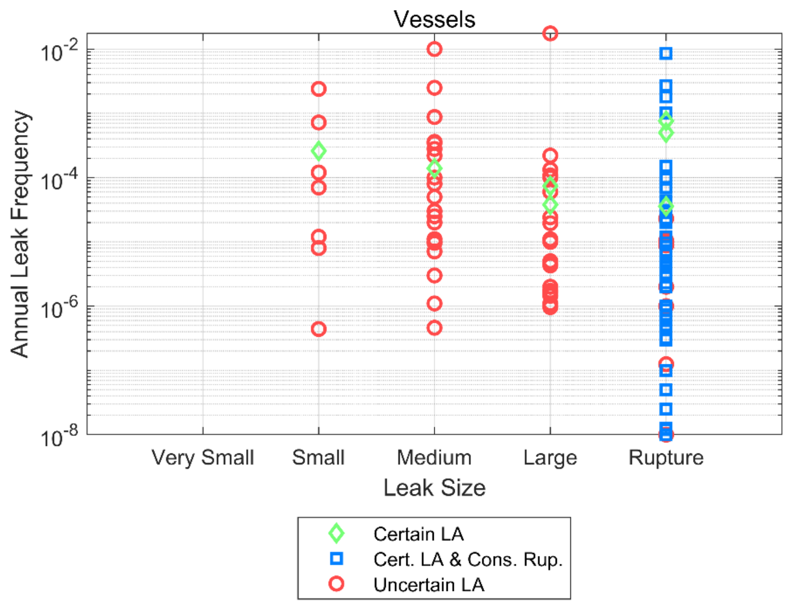

Figure B-18 Scatter plot of vessel data by leak area bin certainty

APPENDIX C. LEAK FREQUENCY PREDICTION TABLES

This appendix contains leak frequency predictions from the different analysis in Section 4. Recall that LNG Source Fluid refers to data that are known to come from LNG systems. LNG-Applicable refers to data that apply to LNG systems but may not all come from LNG systems; the LNG Source Fluid data set is a subset of the LNG-Applicable data. Finally, LNG-Applicable & Generic refers to all data; it includes the LNG-Applicable data that are either from LNG systems or determined to be applicable to LNG systems, as well as data that are generic with respect to source fluid. Table C-1 contains the mean, 5th percentile, median, and 50th percentile of leak frequency prediction models fit to each of these data sets. Leak sizes are relative to flow area.

Table C-1 Predicted leak frequencies with generic data

Component	Leak Size	LNG Source Fluid				LNG-Applicable				LNG-Applicable & Generic			
		Mean	5th	Median	95th	Mean	5th	Median	95th	Mean	5th	Median	95th
Flange and Gasket	0.0001	NA	NA	NA	NA	5.00E-05	1.13E-05	4.18E-05	1.14E-04	5.00E-05	1.13E-05	4.18E-05	1.14E-04
	0.001	NA	NA	NA	NA	5.68E-05	3.52E-06	2.26E-05	1.82E-04	5.74E-05	3.53E-06	2.26E-05	1.83E-04
	0.01	NA	NA	NA	NA	1.17E-03	2.84E-07	1.40E-05	7.14E-04	2.03E-03	2.84E-07	1.40E-05	7.12E-04
	0.1	NA	NA	NA	NA	1.80E-03	8.81E-08	8.68E-06	7.30E-04	2.36E-03	8.78E-08	8.69E-06	7.29E-04
	1	NA	NA	NA	NA	6.20E-02	4.30E-08	5.24E-06	5.40E-04	1.58E-03	4.04E-08	5.23E-06	5.45E-04
Heat Exchanger	0.0001	NA	NA	NA	NA	4.12E-03	5.41E-04	2.34E-03	1.22E-02	2.20E-03	3.93E-04	1.52E-03	6.04E-03
	0.001	NA	NA	NA	NA	2.16E-03	1.03E-04	8.93E-04	7.27E-03	2.63E-02	1.36E-05	6.11E-04	2.76E-02
	0.01	NA	NA	NA	NA	9.04E-04	3.11E-05	3.24E-04	3.23E-03	1.36E-03	1.25E-05	2.46E-04	4.74E-03
	0.1	NA	NA	NA	NA	3.48E-03	2.69E-06	1.17E-04	5.14E-03	4.33E-03	1.65E-06	9.73E-04	5.94E-03
	1	NA	NA	NA	NA	1.96E-04	3.13E-06	4.18E-05	6.27E-04	1.03E-03	7.62E-07	3.92E-05	2.04E-03
Hose	0.0001	NA	NA	NA	NA	2.01E-06	4.49E-07	1.52E-06	5.13E-06	1.88E-06	4.49E-07	1.45E-06	4.67E-06
	0.001	NA	NA	NA	NA	9.31E-06	3.16E-06	7.89E-06	1.99E-05	9.40E-06	3.20E-06	7.96E-06	2.01E-05
	0.01	NA	NA	NA	NA	1.99E+03	4.38E-08	4.13E-05	3.82E-02	9.96E+02	5.46E-08	4.44E-05	3.46E-02
	0.1	NA	NA	NA	NA	3.38E-04	4.65E-05	2.14E-04	1.00E-03	3.51E-04	6.10E-05	2.42E-04	9.80E-04
	1	NA	NA	NA	NA	1.28E+02	2.96E-06	1.10E-03	4.34E-01	4.35E+03	3.75E-06	1.32E-03	4.94E-01
Joint	0.0001	NA	NA	NA	NA	2.09E+08	9.89E+02	3.51E+04	1.25E+06	4.70E+06	9.82E+02	3.51E+04	1.25E+06
	0.001	NA	NA	NA	NA	1.37E+04	3.20E+01	4.77E+02	7.09E+03	3.50E+03	3.19E+01	4.76E+02	7.10E+03
	0.01	NA	NA	NA	NA	1.35E+01	9.98E-01	6.46E+00	4.18E+01	1.32E+01	9.97E-01	6.45E+00	4.18E+01
	0.1	NA	NA	NA	NA	1.14E-01	2.78E-02	8.76E-02	2.76E-01	1.14E-01	2.79E-02	8.76E-02	2.76E-01
	1	NA	NA	NA	NA	1.44E-03	4.32E-04	1.19E-03	3.26E-03	1.45E-03	4.32E-04	1.19E-03	3.27E-03
Loading Arm	0.0001	NA	NA	NA	NA	2.99E+00	4.73E-04	1.99E-01	1.04E+01	1.64E+00	2.41E-04	1.17E-01	5.31E+00
	0.001	NA	NA	NA	NA	3.98E-02	1.67E-04	1.23E-02	1.59E-01	2.84E-02	1.06E-04	8.98E-03	1.11E-01
	0.01	NA	NA	NA	NA	3.58E-03	1.92E-05	7.45E-04	7.87E-03	2.77E-03	1.39E-05	6.80E-04	7.66E-03
	0.1	NA	NA	NA	NA	9.16E-05	9.80E-06	3.29E-05	2.58E-04	1.12E-04	1.02E-05	3.67E-05	3.19E-04
	1	NA	NA	NA	NA	2.20E-02	7.22E-09	3.03E-06	4.44E-04	1.18E-02	1.62E-08	4.19E-06	4.51E-04
Pipe	0.0001	6.55E+04	0.00E+00	2.28E-05	4.6E+203	6.66E-06	3.08E-07	2.67E-06	2.30E-05	4.39E-05	2.06E-08	1.11E-06	6.32E-05
	0.001	6.55E+04	0.00E+00	2.30E-05	3.7E+151	4.32E-06	1.39E-07	1.44E-06	1.52E-05	2.98E-06	4.50E-08	6.80E-07	1.03E-05
	0.01	5.7E+282	0.00E+00	2.30E-05	3.25E+99	1.57E-06	1.17E-07	7.86E-07	5.22E-06	1.41E-06	3.30E-08	4.11E-07	5.11E-06

Component	Leak Size	LNG Source Fluid				LNG-Applicable				LNG-Applicable & Generic			
		Mean	5th	Median	95th	Mean	5th	Median	95th	Mean	5th	Median	95th
	0.1	1.5E+136	0.00E+00	2.32E-05	2.70E+47	1.11E-06	4.59E-08	4.25E-07	3.92E-06	6.80E-07	2.48E-08	2.50E-07	2.42E-06
	1	3.00E-05	7.33E-06	2.30E-05	7.26E-05	1.35E-06	1.15E-08	2.30E-07	4.63E-06	2.66E-06	3.52E-09	1.50E-07	6.59E-06
Valve	0.0001	NA	NA	NA	NA	1.04E-04	2.40E-05	8.43E-05	2.48E-04	1.20E-04	3.37E-05	9.97E-05	2.71E-04
	0.001	NA	NA	NA	NA	7.13E-05	8.76E-06	4.20E-05	2.18E-04	1.72E-04	9.82E-06	7.31E-05	5.80E-04
	0.01	NA	NA	NA	NA	4.68E-05	3.54E-06	2.16E-05	1.53E-04	2.00E-03	8.65E-07	5.59E-05	3.56E-03
	0.1	NA	NA	NA	NA	8.80E-05	4.72E-07	1.18E-05	2.69E-04	2.00E-04	2.60E-06	4.17E-05	6.97E-04
	1	NA	NA	NA	NA	7.72E-05	2.34E-07	6.42E-06	1.29E-04	1.33E-04	2.03E-06	3.22E-05	4.72E-04
Vaporizer	0.0001	1.2E+164	0.00E+00	1.22E-03	6.10E+56	2.51E+12	1.27E-04	8.19E-03	5.24E-01	1.94E+02	1.28E-04	8.15E-03	5.20E-01
	0.001	2.7E+107	0.00E+00	7.36E-03	4.68E+37	8.62E+07	1.24E-03	2.63E-02	5.57E-01	1.55E-00	1.24E-03	2.63E-02	5.55E-01
	0.01	1.01E+51	0.00E+00	4.48E-02	3.63E+18	8.51E+02	1.15E-02	8.46E-02	6.23E-01	2.02E-01	1.14E-02	8.46E-02	6.21E-01
	0.1	3.53E-01	8.64E-02	2.72E-01	8.56E-01	4.09E-01	8.65E-02	2.72E-01	8.57E-01	3.54E-01	8.65E-02	2.72E-01	8.56E-01
	1	1.94E+47	0.00E+00	1.66E-00	1.97E+22	1.14E+00	2.79E-01	8.75E-01	2.75E+00	1.14E+00	2.79E-01	8.76E-01	2.75E+00
Vessel	0.0001	3.21E-03	1.06E-06	6.22E-05	3.60E-03	1.11E-03	8.18E-05	4.77E-04	3.41E-03	1.25E-03	8.87E-05	5.30E-04	3.92E-03
	0.001	2.32E-04	1.93E-06	3.81E-05	7.50E-04	2.68E-03	3.69E-06	1.39E-04	5.25E-03	3.19E-03	4.21E-06	1.52E-04	5.55E-03
	0.01	4.84E-05	3.33E-06	2.34E-05	1.64E-04	2.67E-04	1.65E-06	3.90E-05	9.14E-04	2.91E-04	1.72E-06	4.20E-05	1.01E-03
	0.1	8.18E-05	8.52E-07	1.44E-05	2.41E-04	2.49E-04	2.03E-07	1.10E-05	5.80E-04	2.06E-04	2.53E-07	1.16E-05	5.21E-04
	1	1.15E-05	2.80E-06	8.82E-06	2.78E-05	7.11E-04	1.67E-08	3.05E-06	5.77E-04	1.05E-03	1.86E-08	3.17E-06	5.59E-04

Table C-2 contains mean, 5th percentile, median, and 95th percentile predicted leak frequencies for three different (nested) data sets. The first model, Certain LA Assignment, only includes data which had reported leak areas that correspond clearly to a leak size bin. The Certain LA Assignment & Conservative Rupture Assumption includes that data as well as data that could be assigned to ruptures, potentially conservatively. Finally, the All Data model includes both of those data sets as well as data that was assigned to a bin based on analyst assumptions which are not necessarily conservative.

Table C-2 Predicted leak frequencies with leak area assignment assumptions

Component	Leak Size	Certain LA Assignment				Certain LA Assignment & Conservative Rupture Assumption				All Data (certain and assumed)			
		Mean	5th	Median	95th	Mean	5th	Median	95th	Mean	5th	Median	95th
	0.0001	3.73E-05	5.31E-06	2.86E-05	9.50E-05	3.76E-05	4.70E-06	2.87E-05	9.59E-05	5.00E-05	1.13E-05	4.18E-05	1.14E-04

Component	Leak Size	Certain LA Assignment				Certain LA Assignment & Conservative Rupture Assumption				All Data (certain and assumed)			
		Mean	5th	Median	95th	Mean	5th	Median	95th	Mean	5th	Median	95th
Flange and Gasket	0.001	1.88E-05	2.90E-06	1.16E-05	5.53E-05	1.83E-05	2.85E-06	1.13E-05	5.36E-05	5.72E-05	3.54E-06	2.26E-05	1.82E-04
	0.01	8.48E-06	1.29E-06	5.36E-06	2.45E-05	8.17E-05	1.27E-06	5.17E-06	2.35E-05	1.39E-02	2.83E-07	1.40E-05	7.09E-04
	0.1	3.08E-06	8.48E-07	2.52E-06	7.07E-06	3.05E-06	7.20E-07	2.45E-06	7.26E-06	3.62E-03	8.74E-08	8.65E-06	7.25E-04
	1	1.32E-06	4.53E-07	1.15E-06	2.74E-06	7.42E-02	5.15E-09	1.11E-06	2.10E-04	3.62E-03	4.00E-08	5.25E-06	5.45E-04
Heat Exchanger	0.0001	2.06E-03	3.83E-04	1.45E-03	5.58E-03	2.14E-03	3.95E-04	1.50E-03	5.82E-03	2.46E-03	4.29E-04	1.66E-03	6.82E-03
	0.001	6.97E-04	1.90E-04	5.55E-04	1.65E-03	6.97E-04	1.90E-04	5.55E-04	1.64E-03	6.77E-03	2.65E-05	6.71E-04	1.70E-02
	0.01	4.32E-04	3.14E-05	2.15E-04	1.43E-03	4.24E-04	2.95E-05	2.07E-04	1.41E-03	1.14E-03	1.71E-05	2.69E-04	4.07E-03
	0.1	1.03E-04	2.75E-05	8.11E-05	2.48E-04	9.58E-05	2.58E-05	7.57E-05	2.29E-04	4.73E-03	2.00E-06	1.06E-04	5.72E-03
	1	3.90E-05	1.06E-05	3.14E-05	9.12E-05	3.04E-04	9.34E-07	2.82E-05	8.51E-04	6.42E-04	1.30E-06	4.20E-05	1.43E-03
Hose	0.0001	1.73E+04	4.41E+02	5.19E+03	6.07E+04	3.14E+02	1.41E+01	1.25E+02	1.12E+03	1.87E-06	4.49E-07	1.44E-06	4.66E-06
	0.001	6.94E+01	9.34E+00	4.41E+01	2.07E+02	1.01E+01	1.62E+00	6.85E+00	2.89E+01	9.40E-06	3.20E-06	7.95E-06	2.01E-05
	0.01	4.44E-01	1.46E-01	3.74E-01	9.58E-01	4.45E-01	1.46E-01	3.75E-01	9.59E-01	8.39E+01	5.52E-08	4.43E-05	3.45E-02
	0.1	4.48E-03	8.33E-04	3.18E-03	1.22E-02	2.71E-02	6.09E-03	2.05E-02	6.90E-02	3.51E-04	6.12E-05	2.43E-04	9.81E-04
	1	6.53E+02	2.80E-07	2.70E-05	2.65E-03	2.52E+06	3.10E-06	1.12E-03	4.10E-01	1.46E+01	3.80E-06	1.33E-03	4.93E-01
Joint	0.0001	NA	NA	NA	NA	6.55E+04	0.00E+00	1.56E-03	7.90E+204	8.61E+18	9.92E+02	3.52E+04	1.25E+06
	0.001	NA	NA	NA	NA	6.55E+04	0.00E+00	1.44E-03	8.63E+152	1.34E+12	3.20E+01	4.77E+02	7.10E+03
	0.01	NA	NA	NA	NA	6.55E+04	0.00E+00	1.35E-03	9.78E+100	5.97E+04	1.00E+00	6.47E+00	4.18E+01
	0.1	NA	NA	NA	NA	1.8E+157	0.00E+00	1.27E-03	1.07E+49	1.31E-01	2.78E-02	8.77E-02	2.76E-01
	1	NA	NA	NA	NA	1.44E-03	4.32E-04	1.19E-03	3.26E-03	1.44E-03	4.32E-04	1.19E-03	3.26E-03
Loading Arm	0.0001	3.46E+01	5.24E-01	7.36E+00	1.03E+02	3.96E-01	1.15E-03	7.35E-02	1.53E+00	1.63E+00	2.57E-04	1.17E-01	5.25E+00
	0.001	1.77E-01	1.51E-02	9.24E-02	5.66E-01	1.28E-02	3.16E-04	6.23E-03	4.52E-02	2.82E-02	1.10E-04	8.96E-03	1.11E-01
	0.01	1.87E-03	2.51E-04	1.16E-03	5.38E-03	9.07E-04	4.53E-05	5.24E-04	2.75E-03	3.09E-03	1.43E-05	6.79E-04	7.65E-03
	0.1	1.74E-05	5.57E-06	1.46E-05	3.81E-05	7.49E-05	1.03E-05	3.50E-05	2.20E-04	1.11E-04	1.02E-05	3.67E-05	3.19E-04
	1	2.79E-07	4.14E-08	1.83E-07	8.10E-07	5.72E-03	1.71E-08	3.86E-06	4.22E-04	1.28E-01	1.61E-08	4.19E-06	4.50E-04
Pipe	0.0001	7.54E-05	2.35E-08	1.41E-06	9.08E-05	7.43E-05	2.31E-08	1.36E-06	8.49E-05	4.75E-05	2.07E-08	1.11E-06	6.32E-05
	0.001	3.13E-06	5.95E-08	8.10E-07	1.10E-05	3.07E-06	5.86E-08	8.01E-07	1.09E-05	2.95E-06	4.51E-08	6.79E-07	1.03E-05
	0.01	1.54E-06	3.80E-08	4.58E-07	5.49E-06	1.55E-06	3.87E-08	4.66E-07	5.57E-06	1.42E-06	3.29E-08	4.12E-07	5.09E-06
	0.1	7.09E-07	2.57E-08	2.63E-07	2.49E-06	7.17E-07	2.83E-08	2.73E-07	2.52E-06	6.82E-07	2.49E-08	2.50E-07	2.42E-06

Component	Leak Size	Certain LA Assignment				Certain LA Assignment & Conservative Rupture Assumption				All Data (certain and assumed)			
		Mean	5th	Median	95th	Mean	5th	Median	95th	Mean	5th	Median	95th
	1	1.17E-06	7.54E-09	1.43E-07	3.14E-06	2.64E-06	3.74E-09	1.56E-07	6.81E-06	2.46E-06	3.52E-09	1.49E-07	6.56E-06
Valve	0.0001	1.05E-04	2.78E-05	8.63E-05	2.43E-04	9.93E-05	2.34E-05	8.02E-05	2.35E-04	1.20E-04	3.37E-05	9.98E-05	2.71E-04
	0.001	5.17E-05	8.03E-06	3.40E-05	1.49E-04	5.20E-05	7.97E-06	3.40E-05	1.50E-04	1.72E-04	9.83E-06	7.31E-05	5.81E-04
	0.01	2.13E-05	3.10E-06	1.37E-05	6.23E-05	2.43E-05	3.21E-06	1.49E-05	7.26E-05	1.95E-03	8.63E-07	5.59E-05	3.57E-03
	0.1	7.56E-06	1.56E-06	5.57E-06	1.98E-05	9.31E-06	1.78E-06	6.54E-06	2.52E-05	2.02E-04	2.60E-06	4.17E-05	6.98E-04
	1	3.08E-06	5.88E-07	2.30E-06	7.94E-06	9.55E-04	1.83E-08	2.99E-06	4.11E-04	1.32E-04	2.03E-06	3.22E-05	4.72E-04
Vaporizer	0.0001	NA	NA	NA	NA	6.55E+04	0.00E+00	1.12E+00	8.39E+207	1.67E+01	1.26E-04	8.08E-03	5.22E-01
	0.001	NA	NA	NA	NA	6.55E+04	0.00E+00	1.06E+00	8.23E+155	4.49E-01	1.23E-03	2.61E-02	5.54E-01
	0.01	NA	NA	NA	NA	6.55E+04	0.00E+00	1.01E+00	8.41E+103	1.92E-01	1.14E-02	8.42E-02	6.22E-01
	0.1	NA	NA	NA	NA	7.50E+163	0.00E+00	9.45E-01	8.77E+51	3.53E-01	8.63E-02	2.71E-01	8.54E-01
	1	NA	NA	NA	NA	1.14E+00	2.78E-01	8.76E-01	2.76E+00	1.15E+00	2.78E-01	8.76E-01	2.76E+00
Vessel	0.0001	6.90E-04	2.42E-04	5.94E-04	1.44E-03	1.26E-03	3.82E-04	9.68E-04	2.97E-03	1.28E-03	8.83E-05	5.31E-04	3.91E-03
	0.001	3.20E-04	1.22E-04	2.80E-04	6.46E-04	3.14E-04	1.13E-04	2.70E-04	6.58E-04	2.89E-03	4.21E-06	1.51E-04	5.58E-03
	0.01	1.49E-04	6.03E-05	1.32E-04	2.89E-04	8.65E-05	2.71E-05	7.44E-05	1.84E-04	2.91E-04	1.71E-06	4.20E-05	1.01E-03
	0.1	7.22E-05	2.66E-05	6.23E-05	1.49E-04	2.53E-05	5.81E-06	2.03E-05	6.01E-05	2.02E-04	2.52E-07	1.16E-05	5.20E-04
	1	3.48E-05	1.15E-05	2.97E-05	7.44E-05	3.06E-03	2.33E-08	5.25E-06	1.28E-03	5.62E-04	1.87E-08	3.16E-06	5.56E-04

APPENDIX D. PLOTS OF PREDICTED LEAK FREQUENCIES

This appendix contains plots of the predicted leak frequencies for all components in the two sensitivity analyses outlined in Section 4.1. Plots included earlier in the report for analysis are also included here for completeness. Black dots on the plots indicate the data points used to fit each model.

D.1. Effect of Generic Data on Leak Frequency Predictions

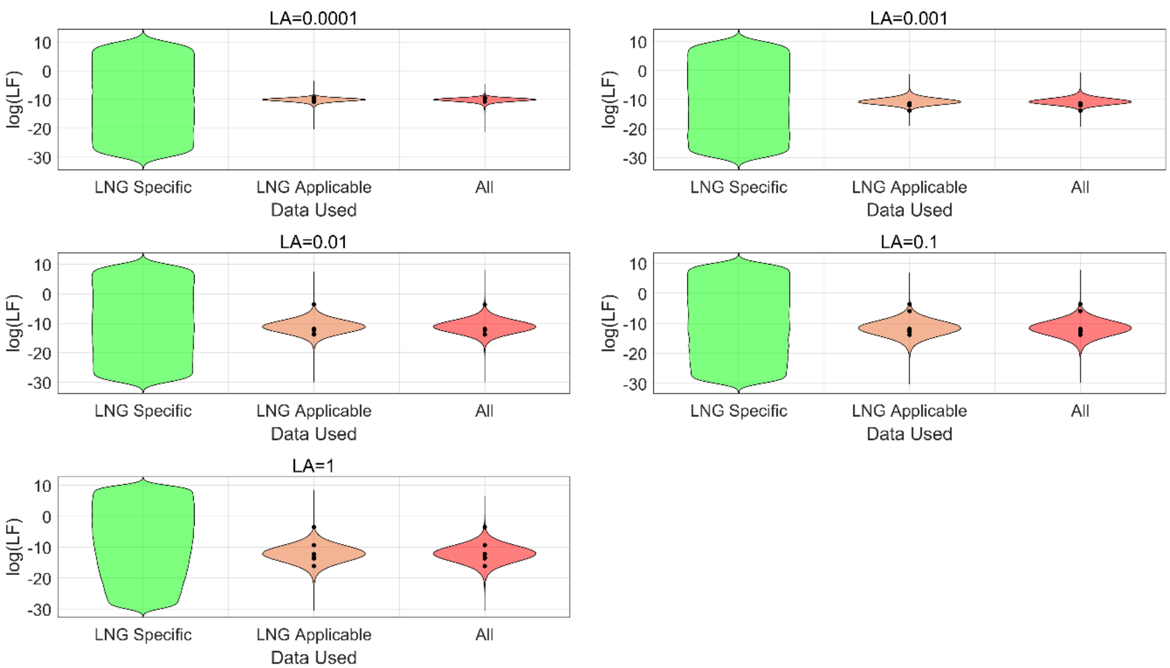

Figure D-1 Flange and gasket sensitivity analysis models based on LNG applicability

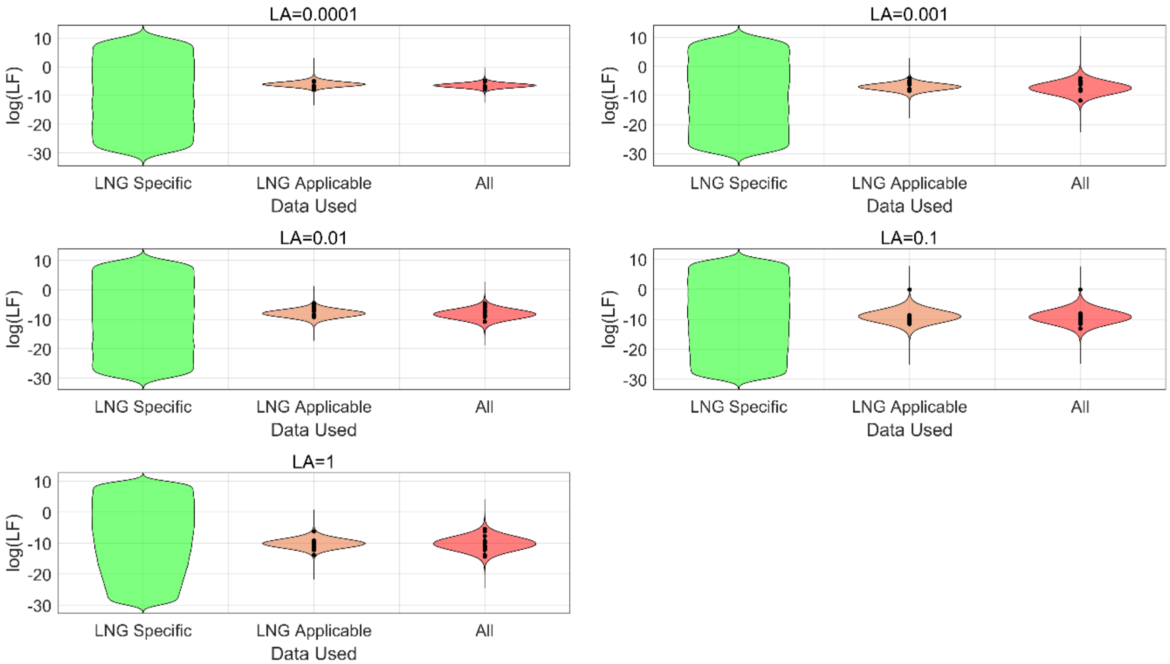

Figure D-2 Heat exchanger sensitivity analysis models based on LNG applicability

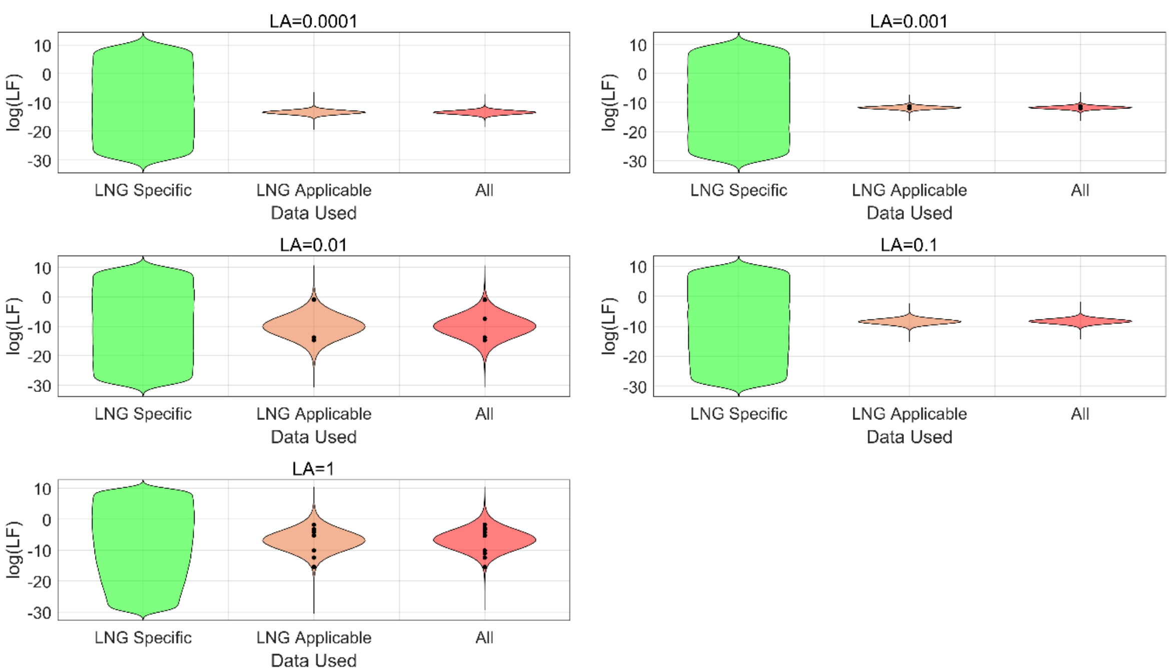

Figure D-3 Hose sensitivity analysis models based on LNG applicability

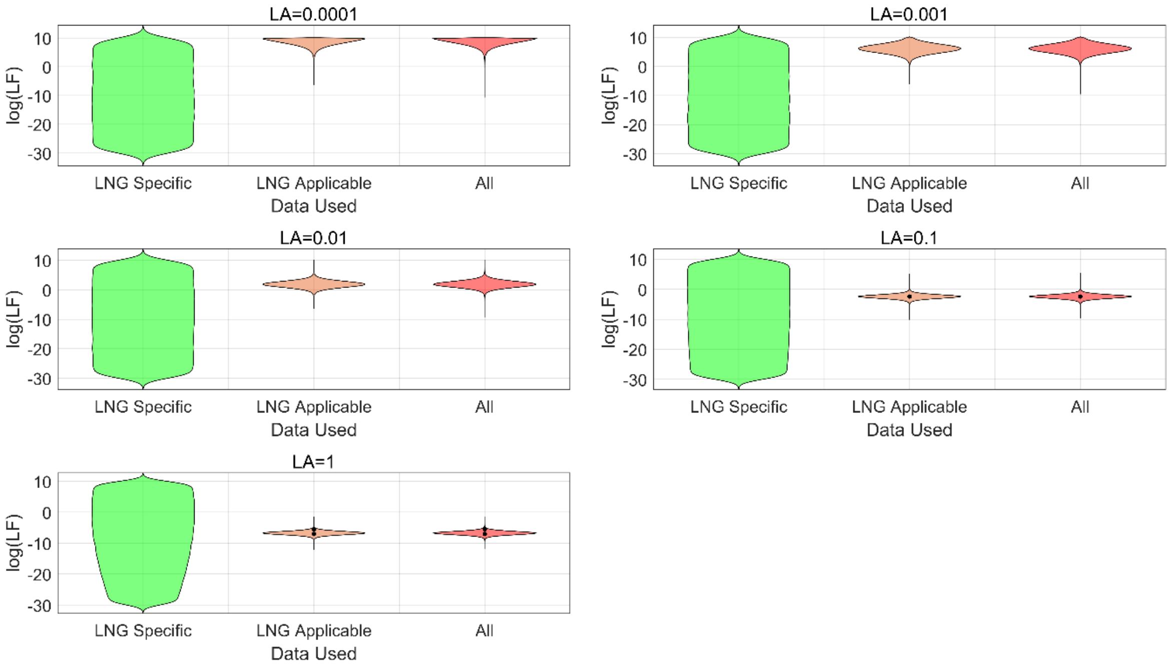

Figure D-4 Joint sensitivity analysis models based on LNG applicability

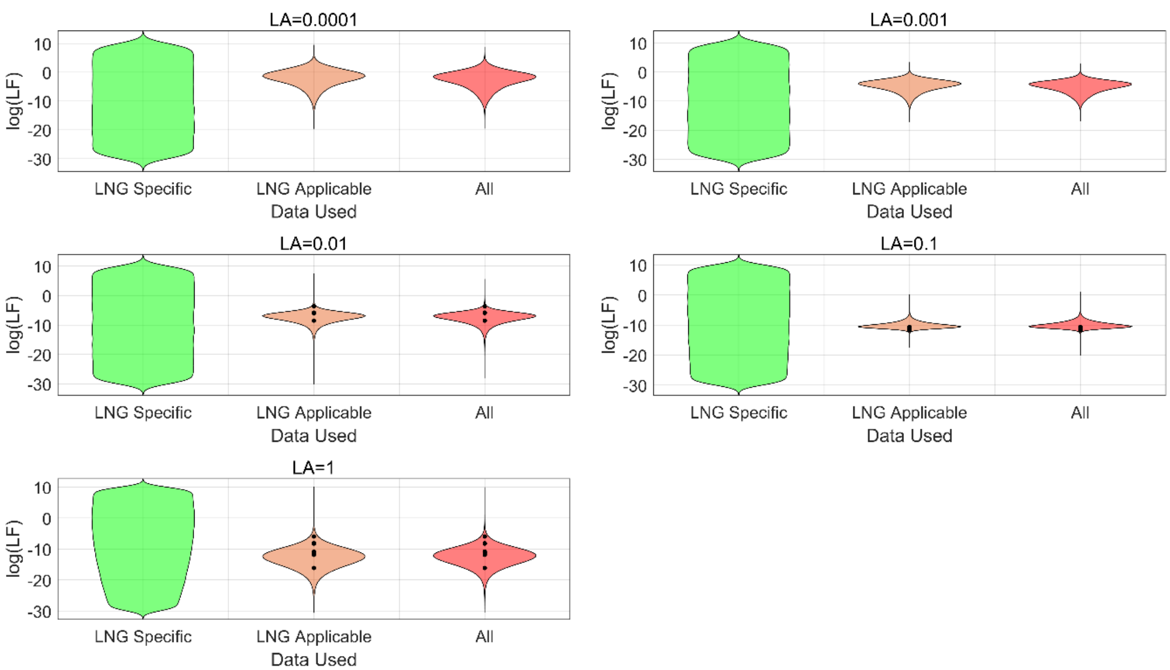

Figure D-5 Loading arm sensitivity analysis models based on LNG applicability

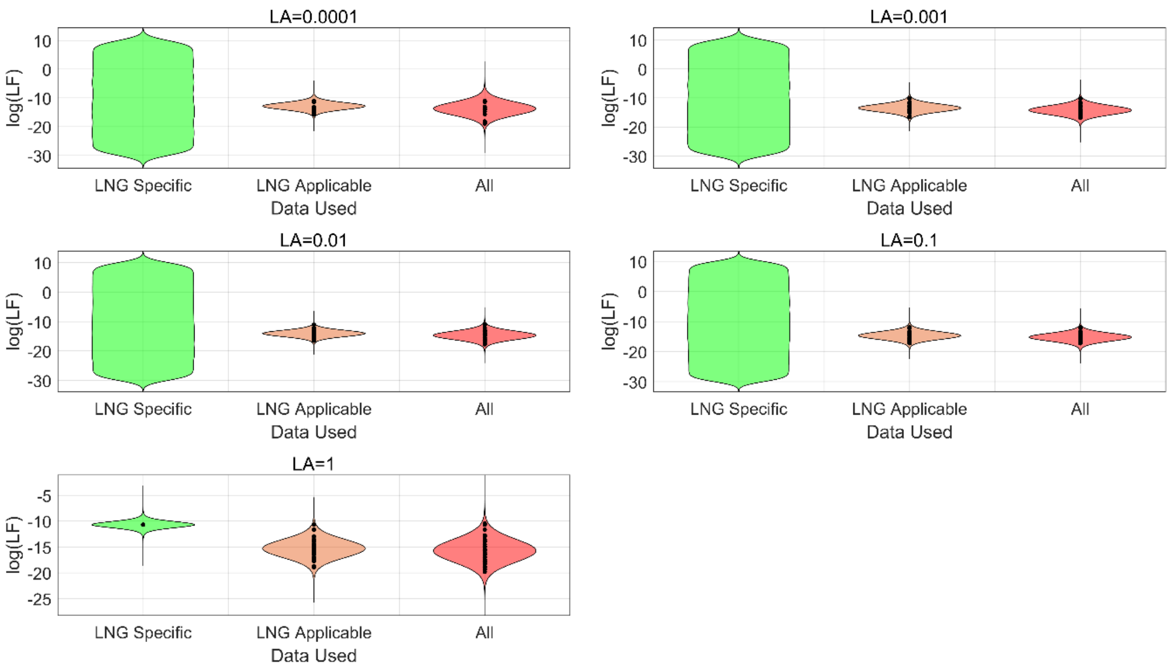

Figure D-6 Pipe sensitivity analysis models based on LNG applicability

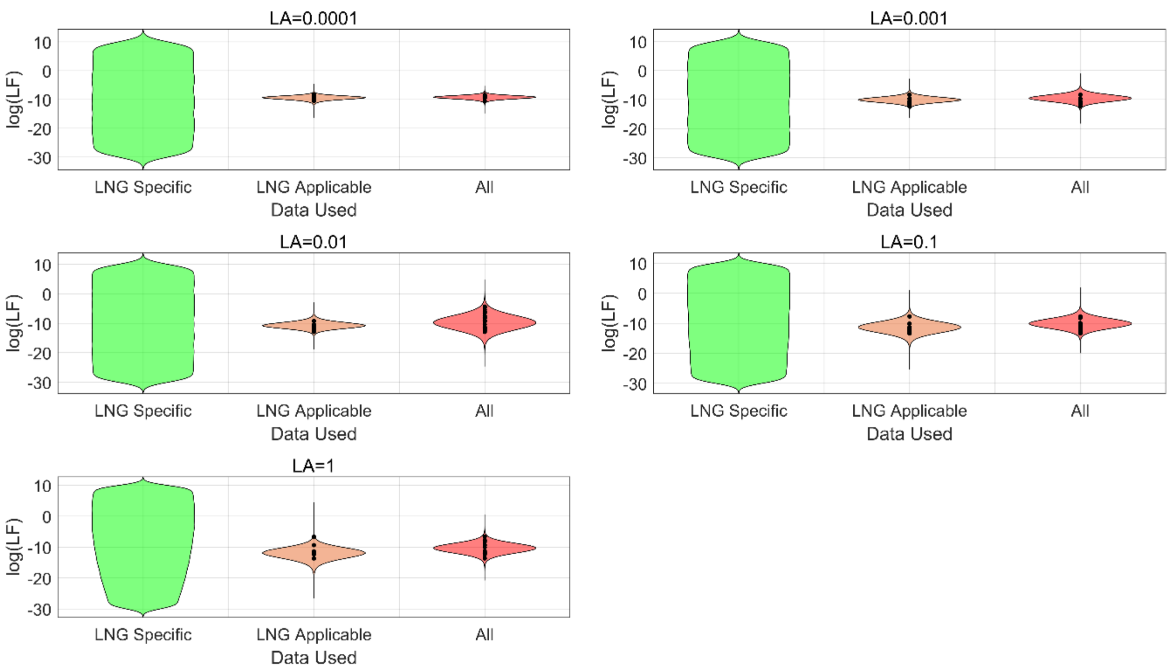

Figure D-7 Valve sensitivity analysis models based on LNG applicability

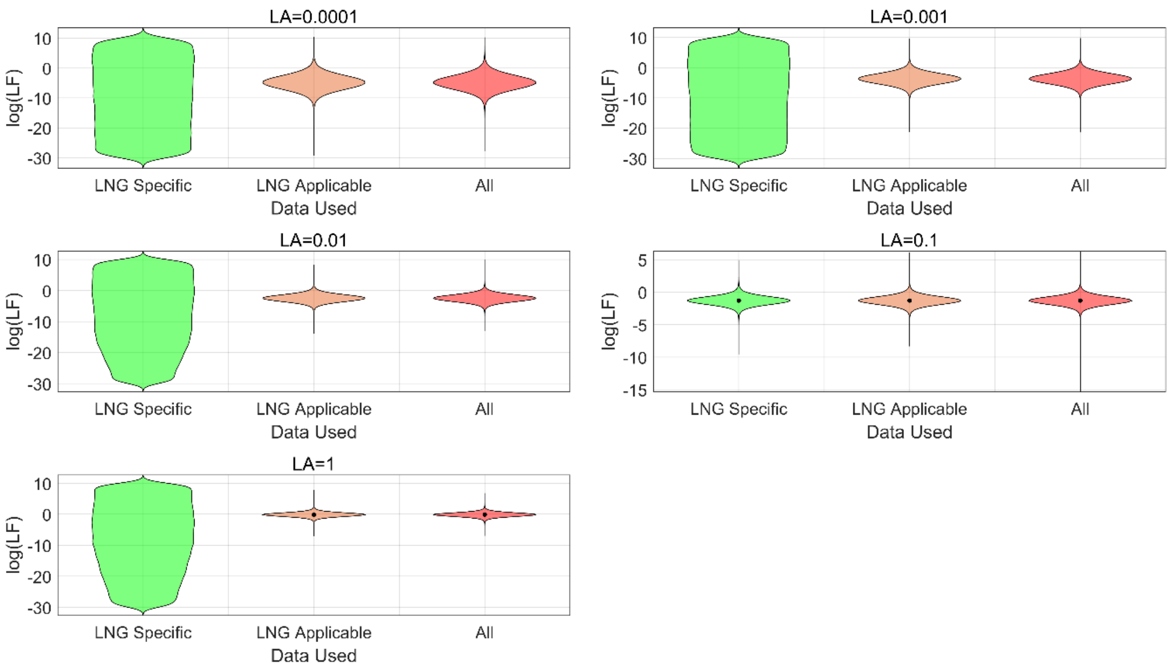

Figure D-8 Vaporizer sensitivity analysis models based on LNG applicability

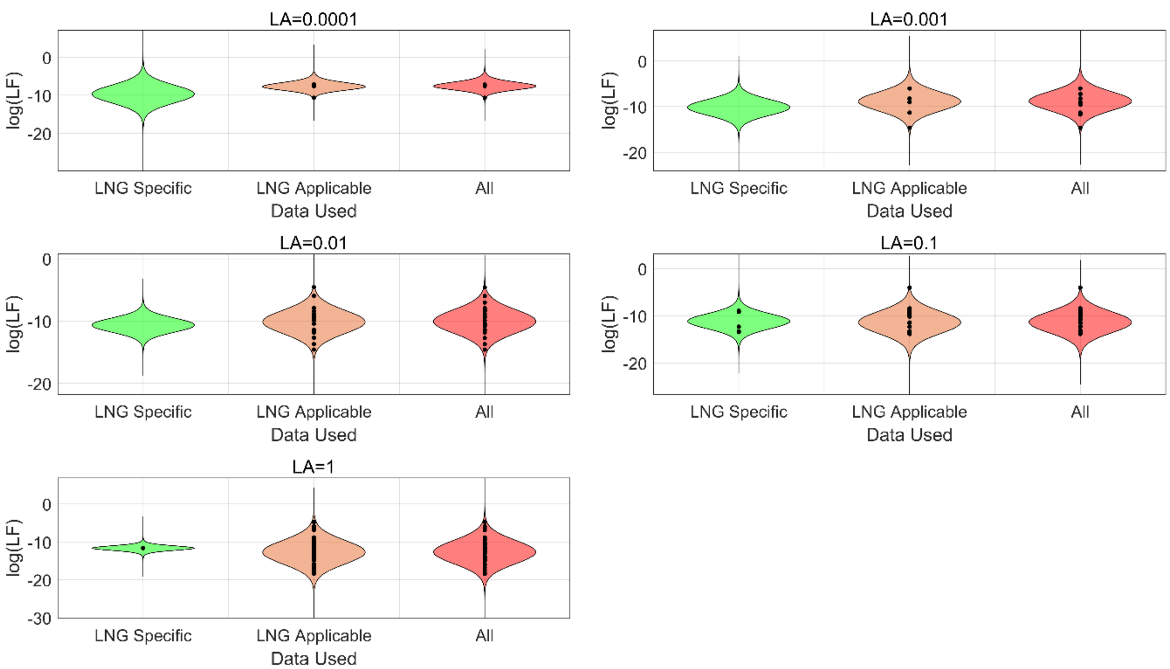

Figure D-9 Vessel sensitivity analysis models based on LNG applicability

D.2. Effect of Leak Area Assumptions on Leak Frequency Predictions

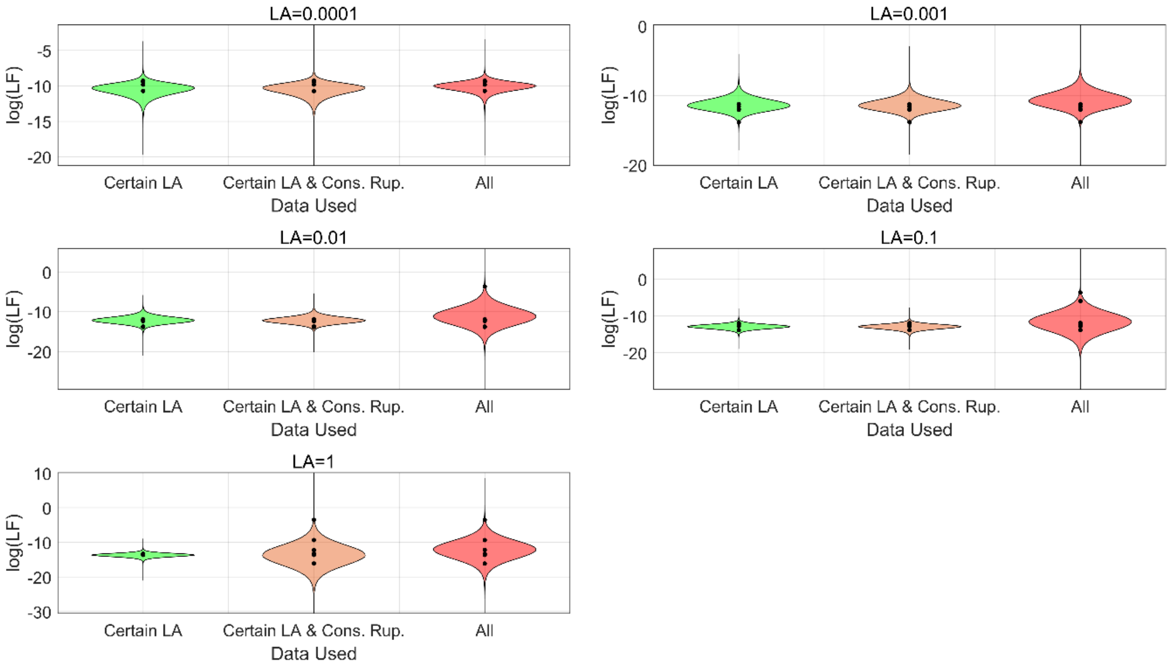

Figure D-10 Flange and gasket sensitivity analysis models based on leak area assumptions

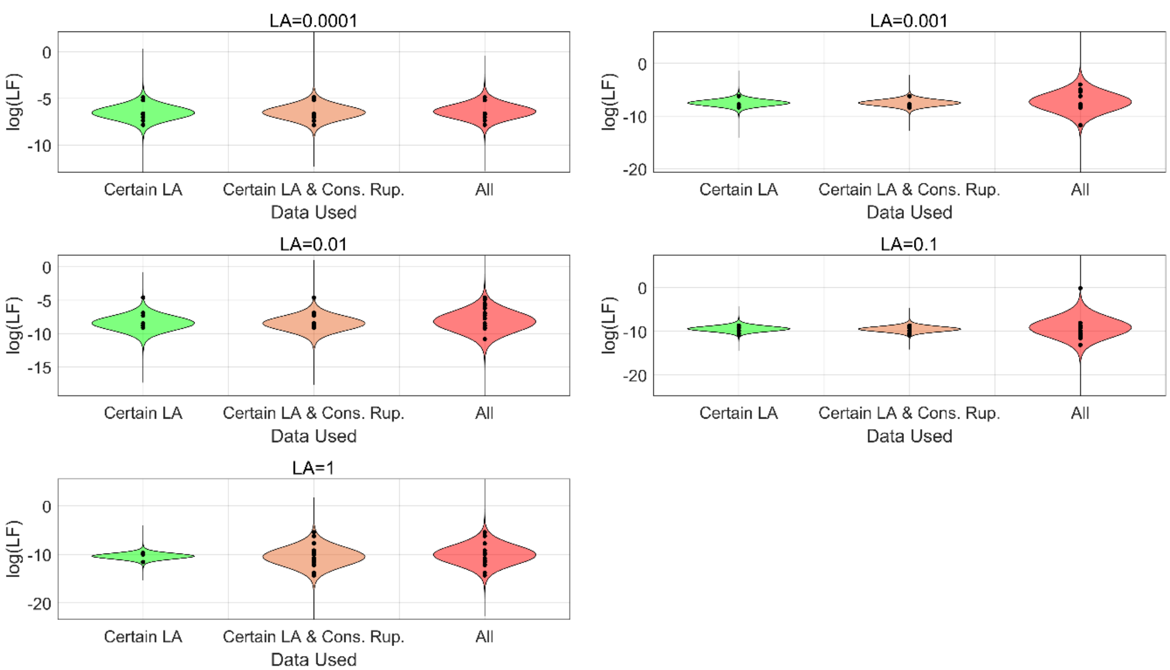

Figure D-11 Heat exchanger sensitivity analysis models based on leak area assumptions

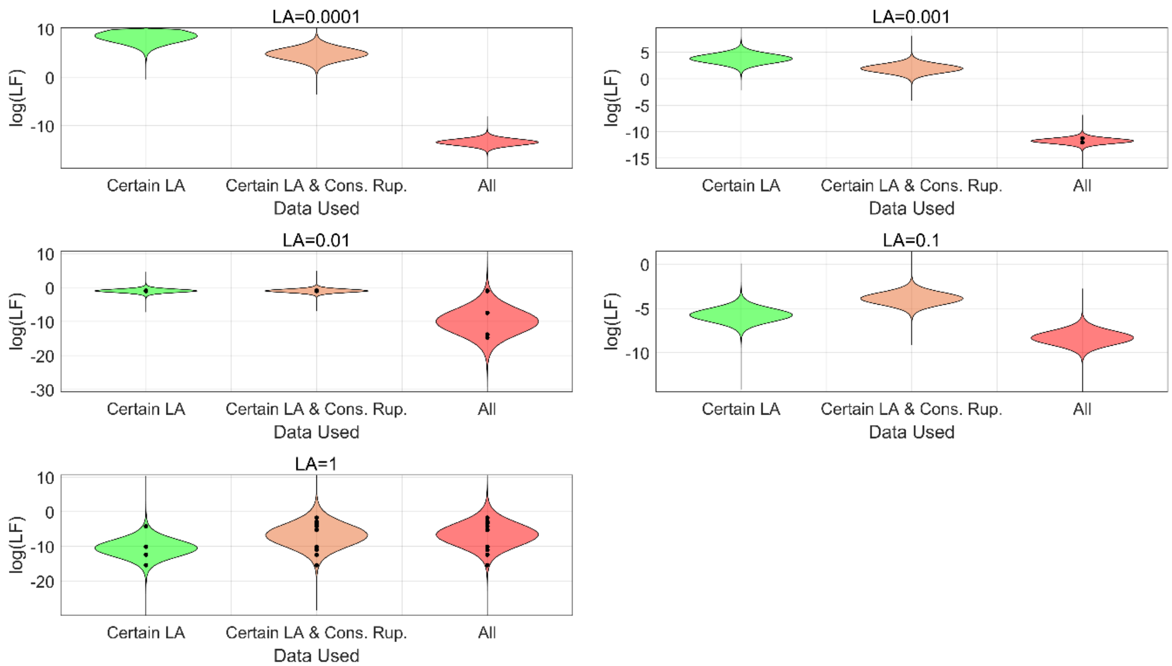

Figure D-12 Hose sensitivity analysis models based on leak area assumptions

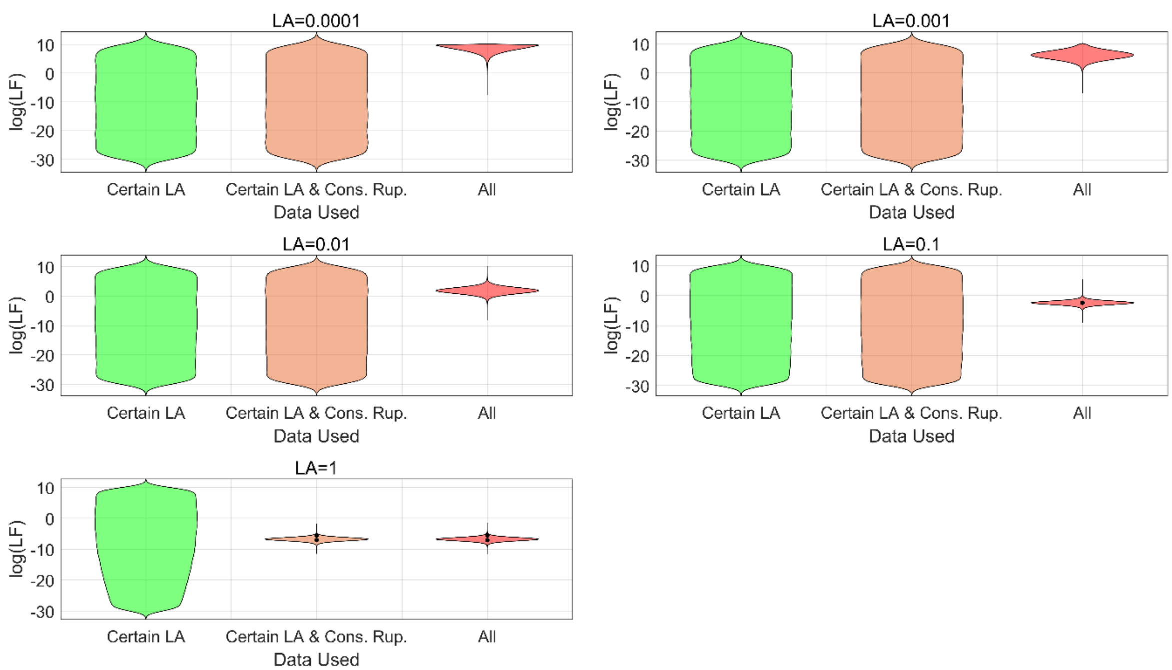

Figure D-13 Joint sensitivity analysis models based on leak area assumptions

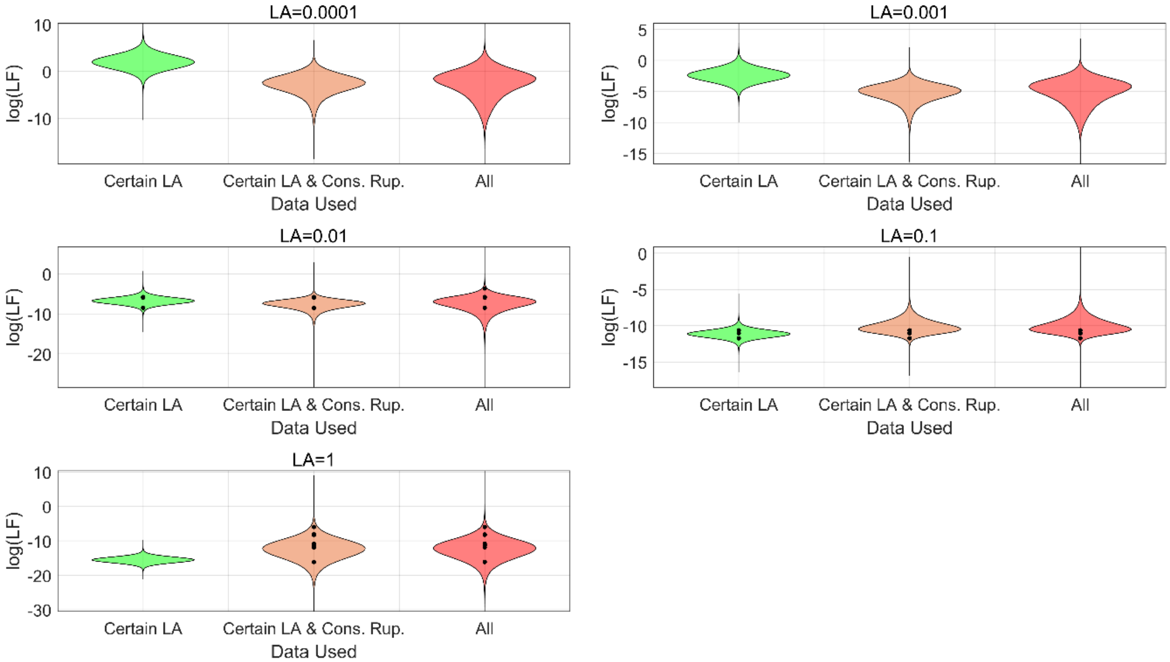

Figure D-14 Loading arm sensitivity analysis models based on leak area assumptions

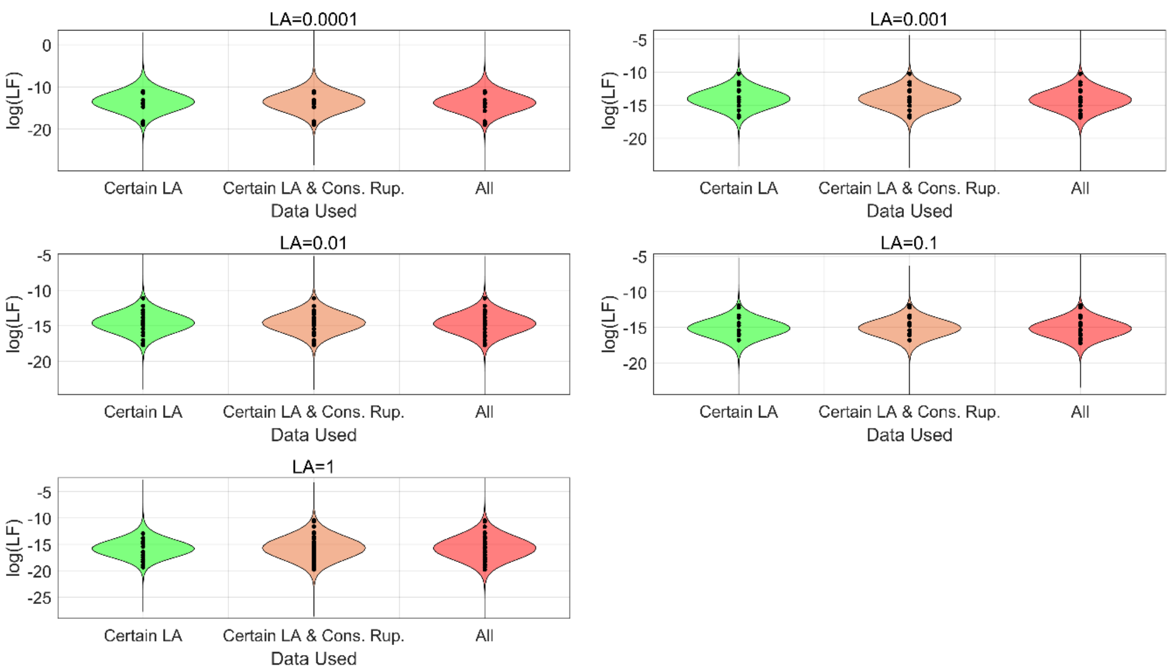

Figure D-15 Pipe sensitivity analysis models based on leak area assumptions

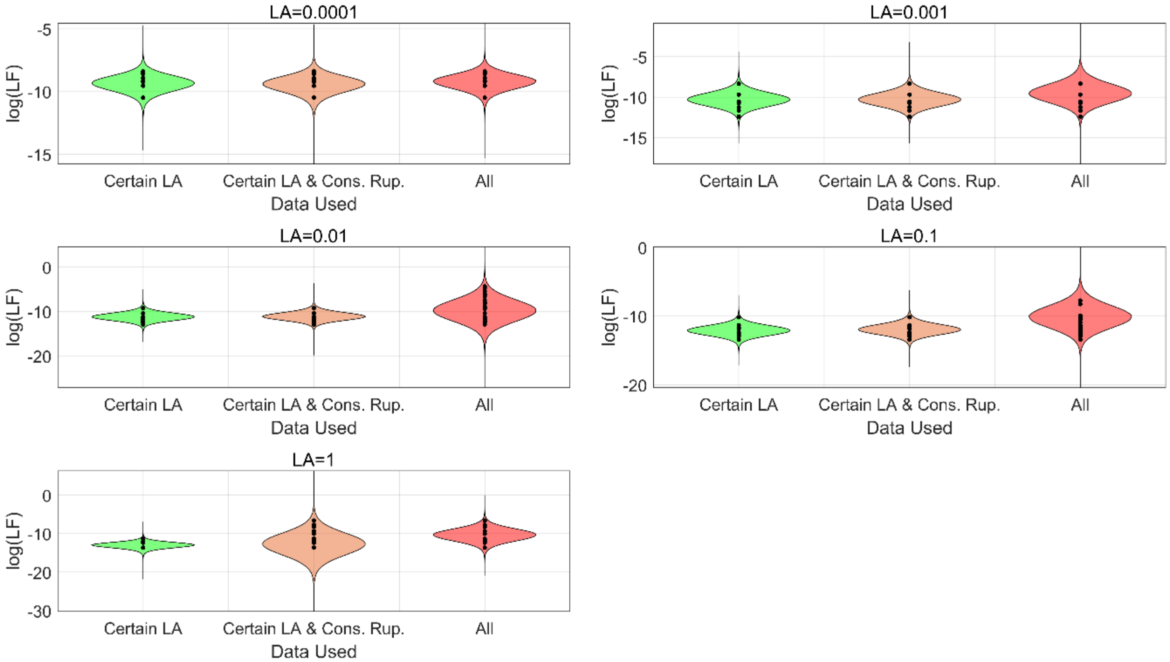

Figure D-16 Valve sensitivity analysis models based on leak area assumptions

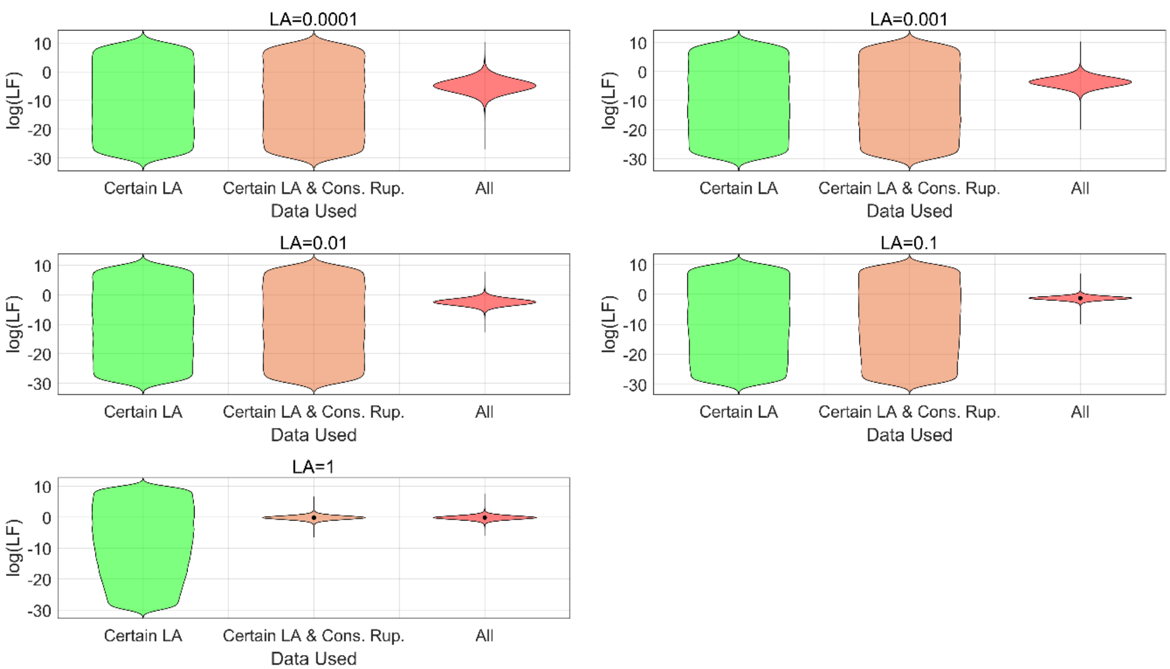

Figure D-17 Vaporizer sensitivity analysis models based on leak area assumptions

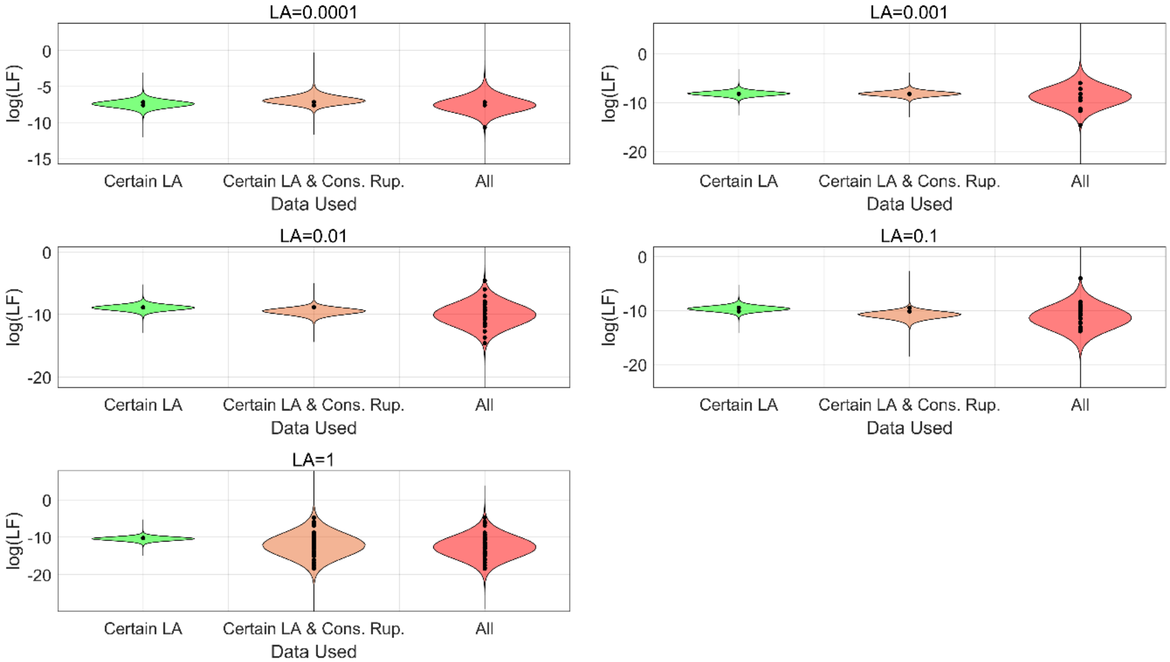

Figure D-18 Vessel sensitivity analysis models based on leak area assumptions

APPENDIX E. TESTING FOR CURVATURE

This appendix includes the numerical results of the test for curvature outlined in Section 4.2.4. Recall that the hypotheses under investigation were

$$H_0: \alpha_3 = 0 \leftrightarrow H_1: \alpha_3 \neq 0$$

Thus, if 0 is not included in the 95% interval, H_0 is rejected and we conclude that there is significant curvature in the mean function. The 95% intervals for each component are provided in Table E-1; intervals containing 0, which suggest significant curvature, are indicated in bold.

Table E-1 Quantile interval for α_3

Component	2.5th Percentile	Median	97.5th Percentile	Mean
Flange & Gasket	0.067	0.124	0.174	0.123
Heat Exchanger	-0.074	-0.037	0.000	-0.037
Hose	**-0.403**	-0.141	**0.117**	-0.143
Joint	**-25.970**	-4.492	**3.845**	-6.889
Loading Arm	0.346	0.539	0.744	0.540
Pipe	-0.086	-0.047	-0.005	-0.046
Valve	0.032	0.067	0.100	0.067
Vaporizer	**-5.839**	7.446	**24.227**	7.891
Vessel	-0.103	-0.064	-0.025	-0.064

APPENDIX F. CONVERGENCE ANALYSIS PLOTS

This appendix contains plots from the convergence analysis in Section 4.3. Convergence in this instance is with respect to sampling uncertainty; converged results have negligible uncertainty due to sample size.

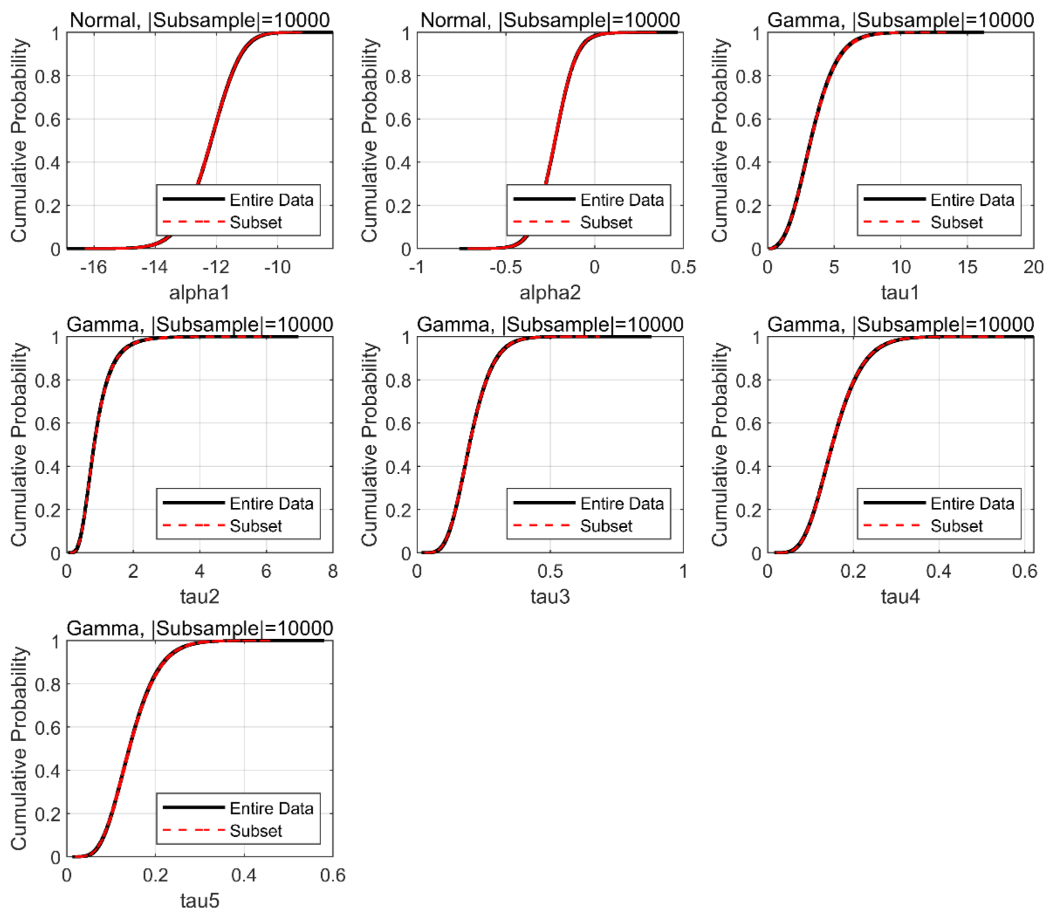

Figure F-1 Plot showing convergence of the posterior distributions for the flange and gasket leak model

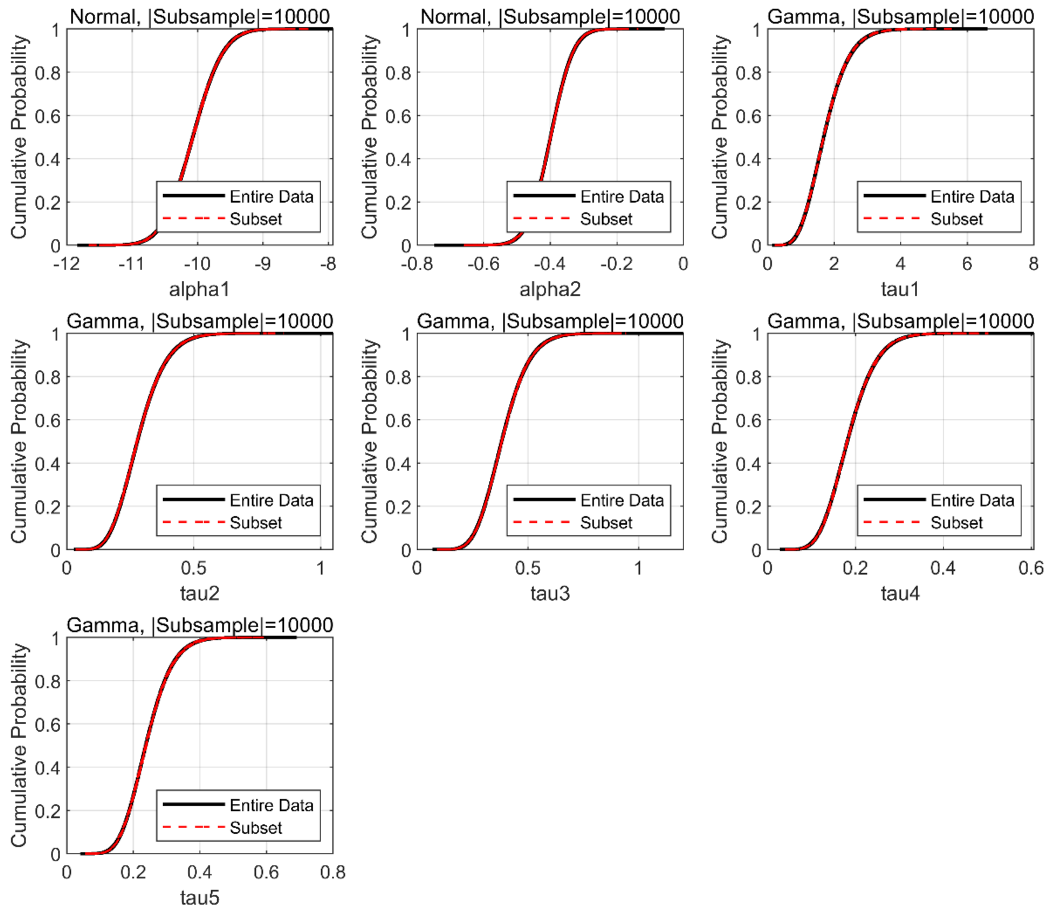

Figure F-2 Plot showing convergence of the posterior distributions for the heat exchanger leak model

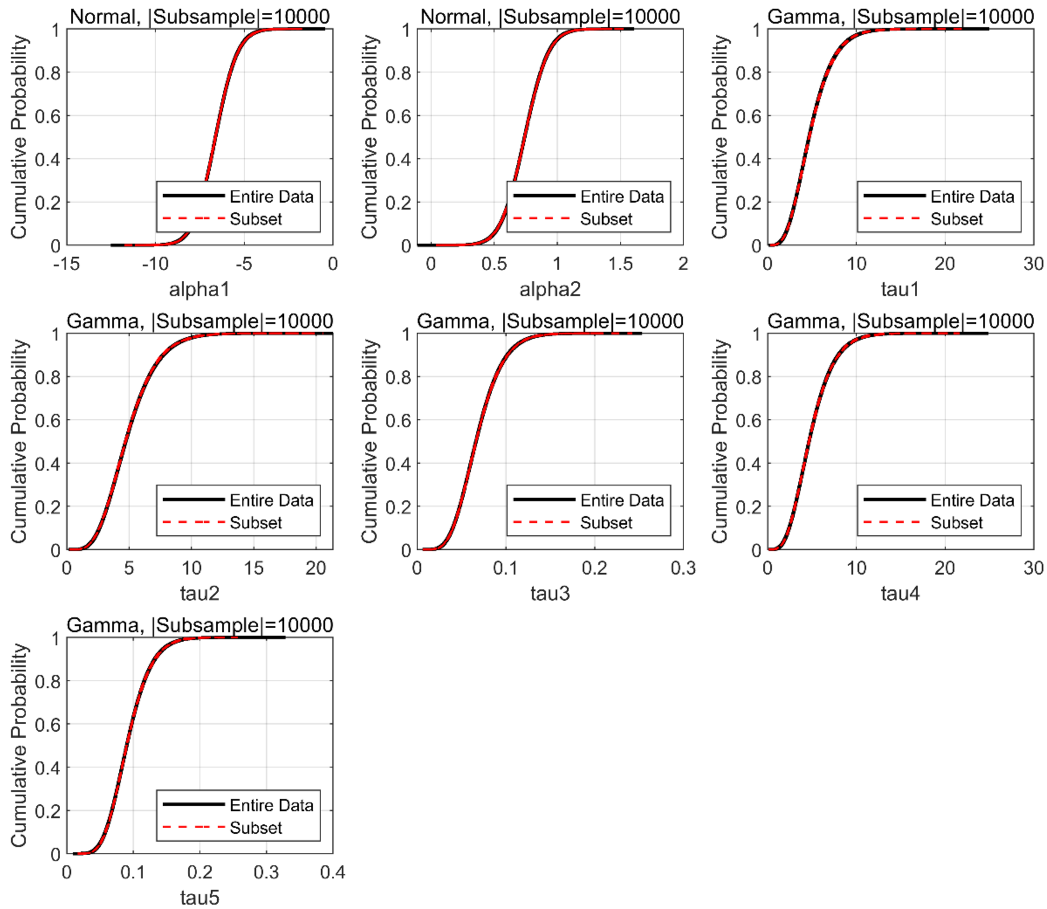

Figure F-3 Plot showing convergence of the posterior distributions for the hose leak model

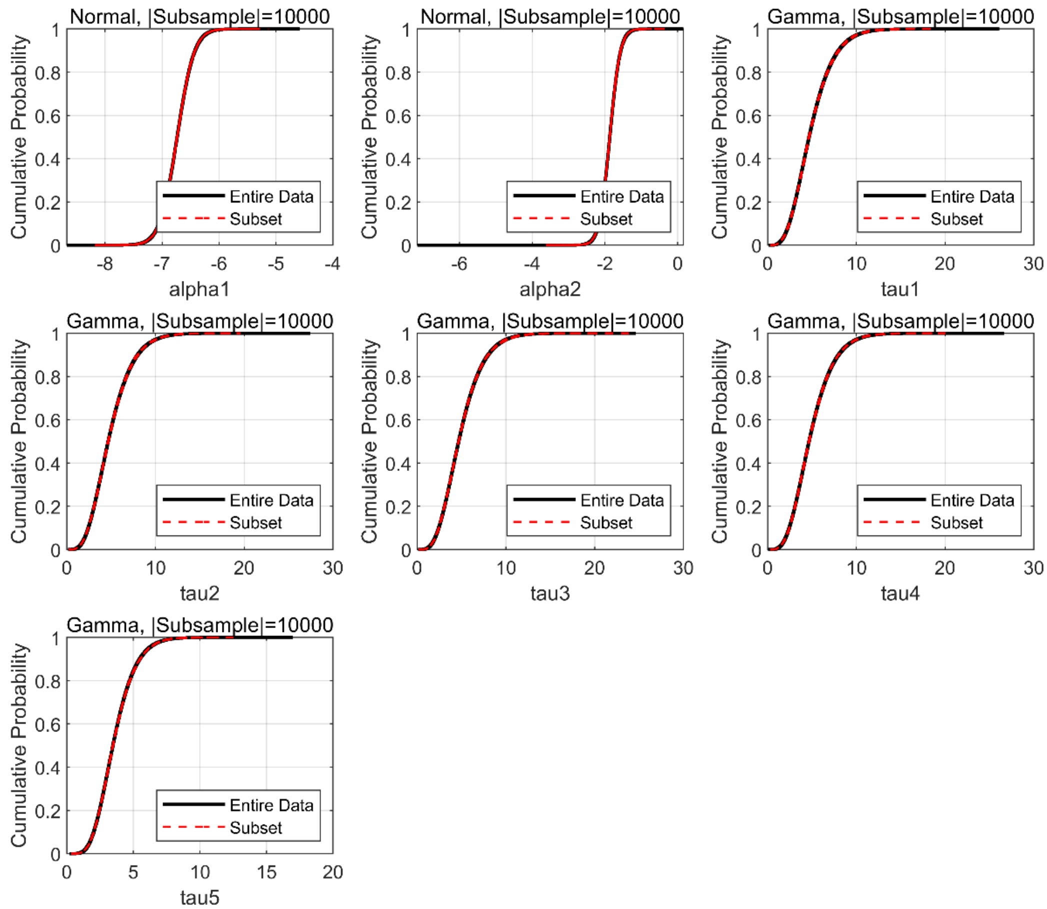

Figure F-4 Plot showing convergence of the posterior distributions for the joint leak model

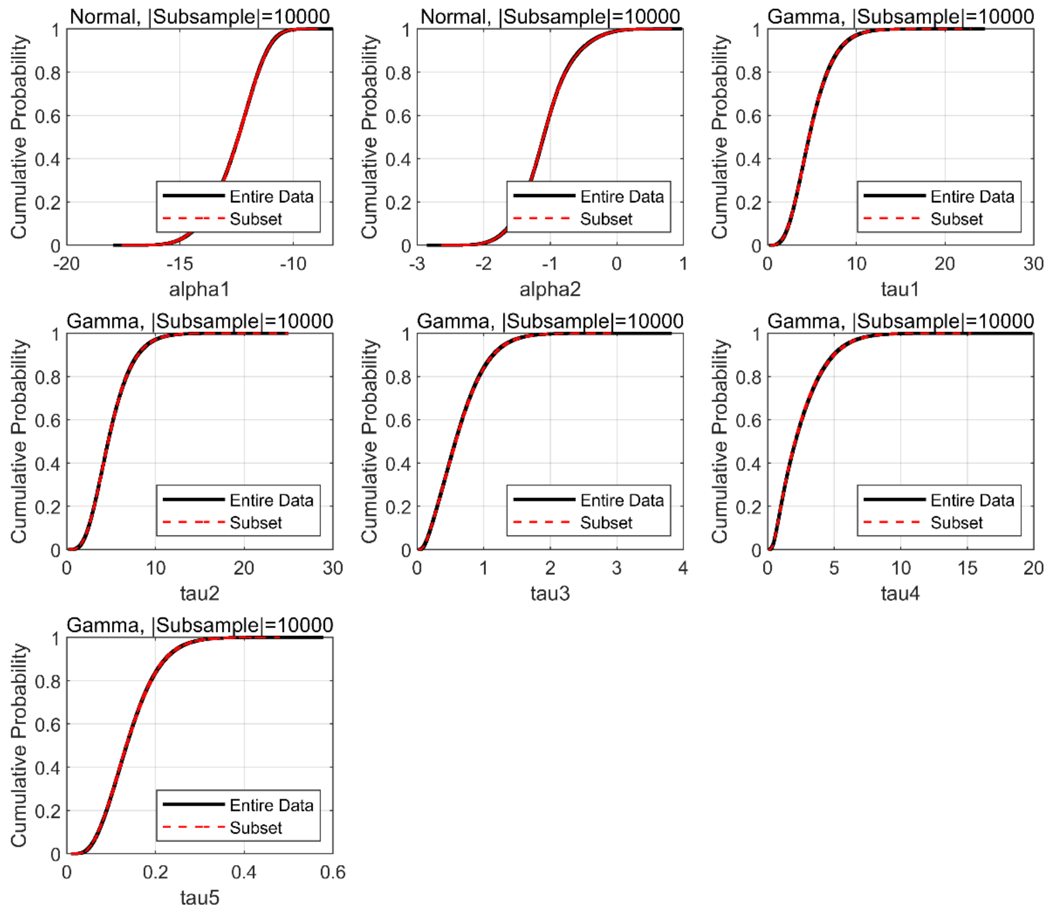

Figure F-5 Plot showing convergence of the posterior distributions for the loading arm leak model

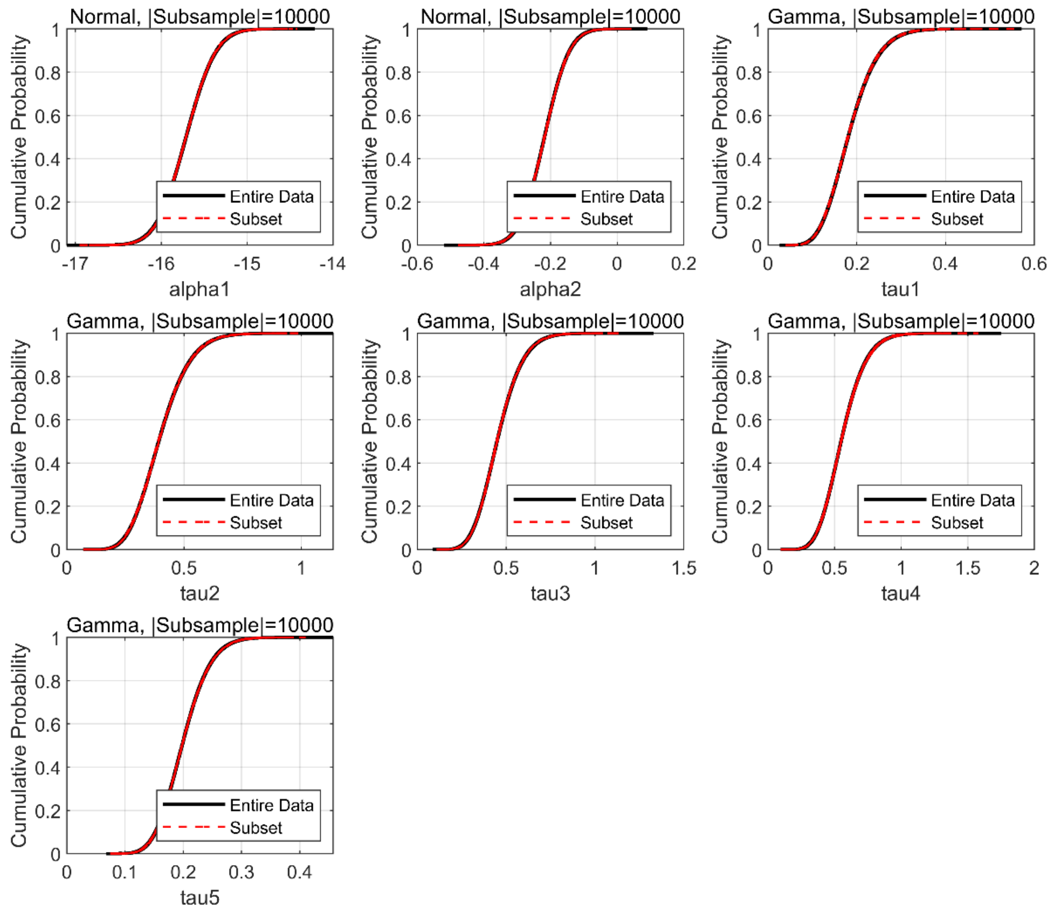

Figure F-6 Plot showing convergence of the posterior distributions for the pipe leak model

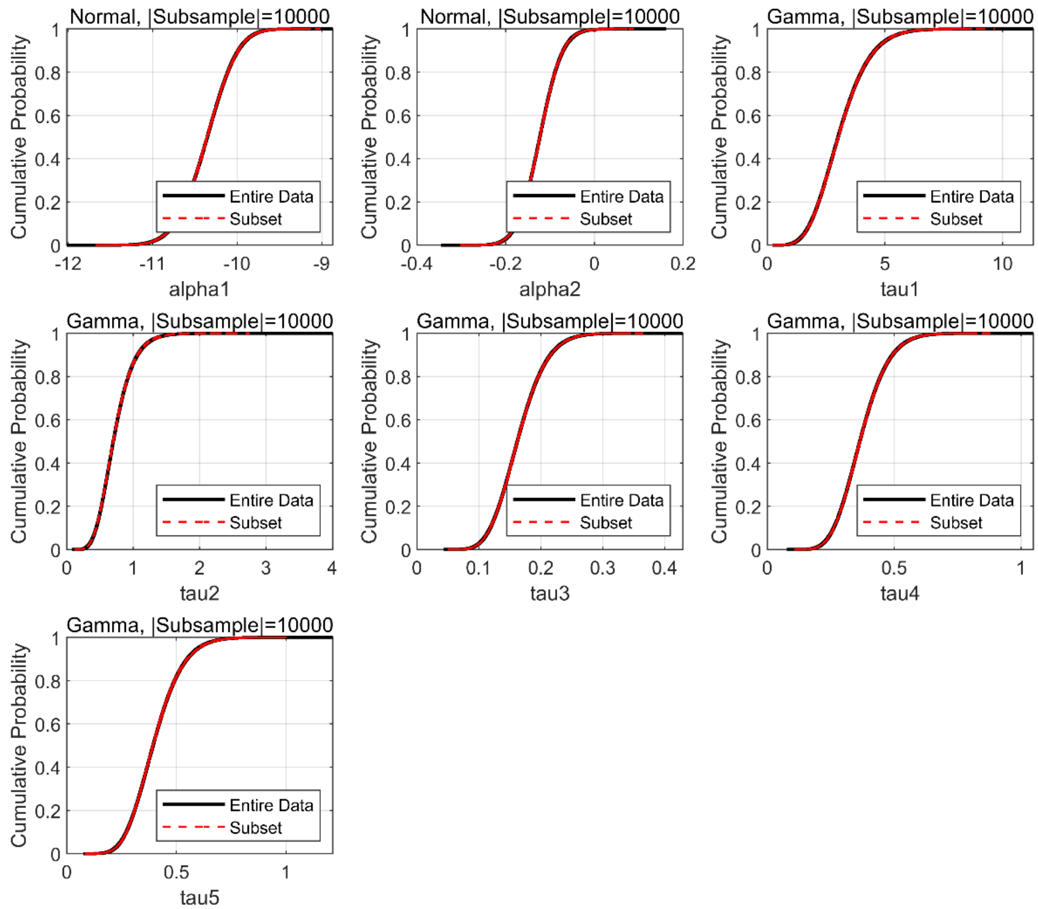

Figure F-7 Plot showing convergence of the posterior distributions for the valve leak model

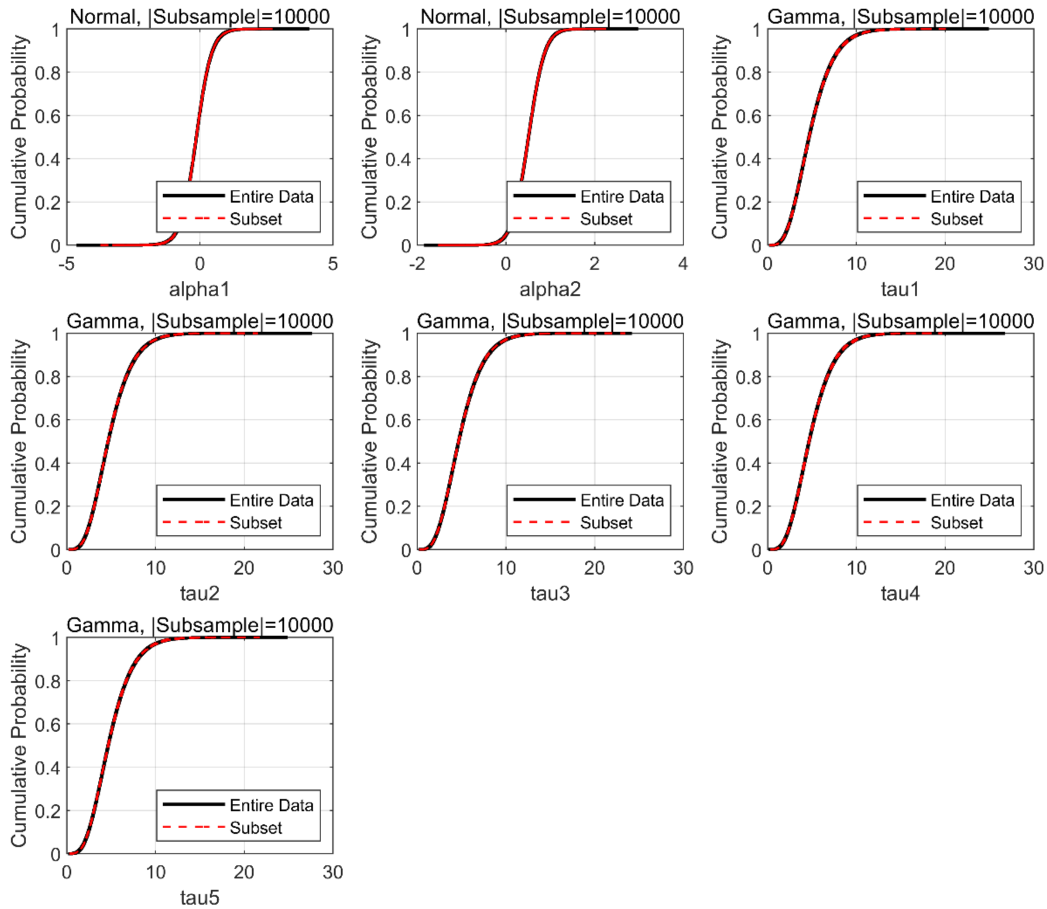

Figure F-8 Plot showing convergence of the posterior distributions for the vaporizer leak model

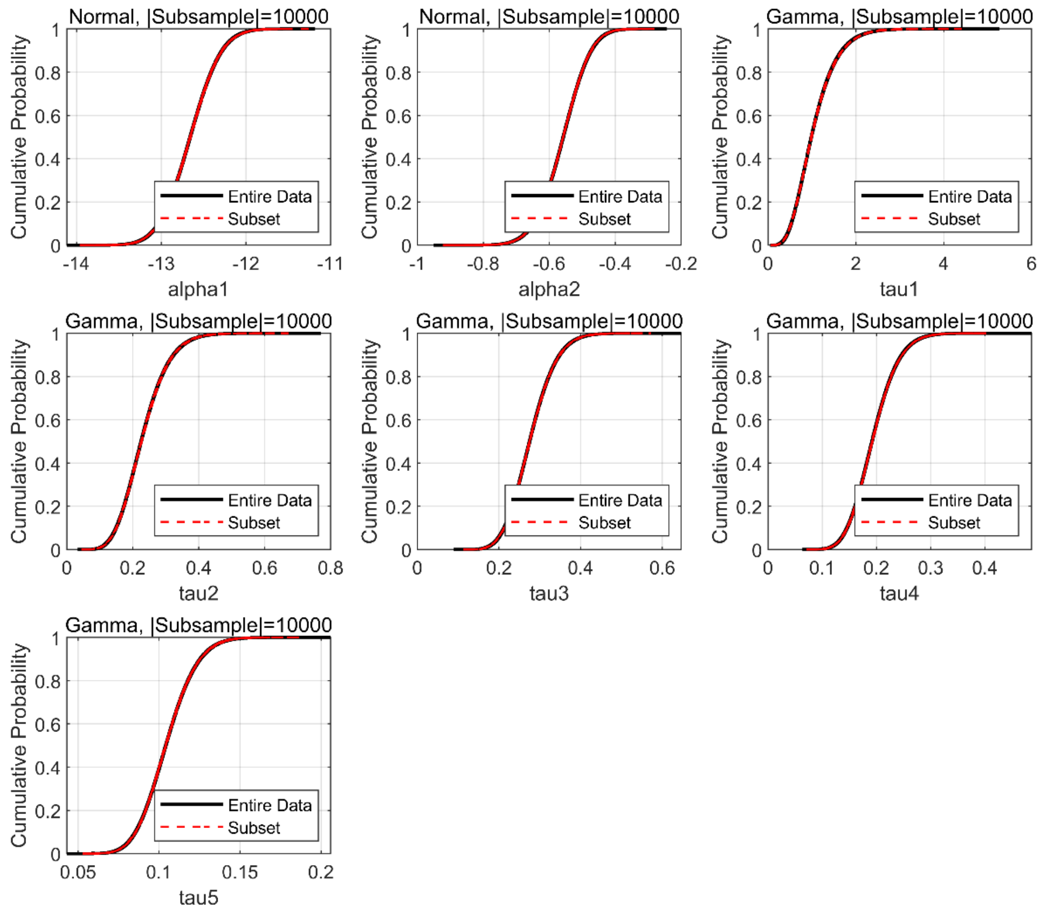

Figure F-9 Plot showing convergence of the posterior distributions for the vessel leak model

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Dusty Brooks	08853	dbrooks@sandia.gov
Michael Starr	08853	mjstarr@sandia.gov
Brian Ehrhart	08854	bdehrha@sandia.gov
Scott Sanborn	08854	sesanbo@sandia.gov
Technical Library	01911	sanddocs@sandia.gov

Email—External

Name	Company Email Address	Company Name
Garrett Mulcahy	gmulcahy@purdue.edu	Purdue University
Thach Nguyen	thach.d.nguyen@dot.gov	DOT PHMSA
Buddy Secor	meredith.secor@dot.gov	DOT PHMSA
Sentho White	sentho.white@dot.gov	DOT PHMSA

This page left blank

Sandia National Laboratories