Documentation for Collision Probability Computation

Promit Chakroborty

March 2025

1 Introduction

On March 26, 2024, the Francis Scott Key Bridge in Baltimore, MD, collapsed when the container ship Dali collided with one of its piers. To investigate whether this event was indicative of a broader issue with the degree of vulnerability of American bridge infrastructure, our team at Johns Hopkins University embarked on an NSF-funded research project to accurately estimate the likelihood of ship-bridge collisions based on real-world shipping data.

For the one-year anniversary of the Key Bridge's collapse, we released estimated return periods of collisions involving large ships ($\geq 150\,\mathrm{m}$) for some of the most important or highly trafficked bridges in the US. These estimates follow the AASHTO methodology (as prescribed in Section 3.14 in [1]), using estimates of ship traffic computed from our real-world ship data collection effort. The estimates can be found on our website.

The computer program to generate these estimates is written in Python and is available on our public GitHub repository. This document serves as a thorough, in-depth documentation of the procedure utilized in the computer program, and details all assumptions and decisions involved in calculating our numbers.

The AIS data was collected from here.

2 Modeling Assumptions and Choices

The following assumptions were made in the procedure and computer program used to compute the expected collision probability/return period.

1. The collision probability calculation has been conducted only for those bridges that we earlier highlighted as having some of the highest traffic from large ($\geq 150\,\mathrm{m}$) or mega ($\geq 300\,\mathrm{m}$) ships, as presented on our website here and here. A total of 23 bridges have been analyzed, including the Francis Scott Key Bridge. As mentioned earlier, the AASHTO methodology is used to compute the collision probability, using traffic computed using

historical AIS data. Equation 3.14.5-1 is used, which is reproduced below along with the associated term descriptions.

$$AF = (N)(PA)(PG)(PF) \tag{1}$$

where AF is the annual frequency of bridge component collapse due to vessel collision, N is the annual number of vessels, classified by type, size, and loading condition, that utilize the channel, PA the probability of vessel aberrancy, PG the geometric probability of a collision between an aberrant vessel and a bridge pier or span, and PF adjustment factor to account for potential protection of the piers from vessel collision due to upstream or downstream land masses or other structures that block the vessel.

2. Collision is only considered to occur with the pier of a bridge, and the dimensions used to compute the collision probability correspond to the base footprint of the pier at the waterline. Collision at different heights above the waterline is not considered. Further, the measurements of the pier used are collected from design drawings where available and from satellite images otherwise.

3. Only the immediate head-on path to collision is considered; thus, bends in the waterway or ship travel path are not incorporated.

4. Dolphins are the only constructions assumed to provide protection against collisions. Since fenders and other constructions are attached to the pier, they are assumed to reduce the probability of collapse given a collision occurs, but not reduce the probability of collision itself.

5. Ship traffic for each bridge is collected between January 2018 and December 2023. Traffic is separated into six classes by length: $150 - 180\text{m}, 180 - 215\text{m}, 215 - 250\text{m}, 250 - 275\text{m}, 275 - 300\text{m}, \geq 300\,\text{m}$, and the annual average of each size class is used for N (Eq. (1)). The average width of ships belonging to a given size class (for a specific bridge) is used as the representative width for all ships belonging to that size class. Similarly, the middle value of the interval of possible lengths is used as the representative length for all ships within a given size class (eg., all ships belonging to the $150 - 180\text{m}$ class are considered to have a size of $165\,\text{m}$). For the final size class, the nominal length is considered to be $300\,\text{m}$ as a conservative case.

6. Ship traffic is assumed to be symmetric in the upstream and downstream directions, which facilitates the computation of PG and PF. Further, assuming symmetry allows us to simplify the analysis for bridges that have parallel spans and piers stacked behind one another.

7. To compute PG, the Normal distribution is required to be centered at the centerline of the vessel sailing path. This is estimated by considering the midpoint between the ship tracks closest to the two banks of the waterway that is spanned by the bridge under consideration.

8. The base aberrancy rate specified in Section 3.14.5.2.3 is used as PA in all cases, without any further modifications.

3 Summary of the computer program

The file 'collision_probability_calculator.py' contains the program to compute the collision probabilities. Comments included in the file explain the procedure line-by-line. Here, a summary of the procedure is provided.

The program reads all the relevant data from multiple input files (see Section 4), then computes the collision probability of each bridge as iterations of a loop. For each bridge, there are further loops for the different size classes and the piers for each bridge. In particular, Eq. (1) must be calculated for each combination of ship size class, bridge pier, and dolphin protecting bridge pier. The procedure results in a separate AF number for each pier for each ship size, with the protection from each dolphin being summed up to give PF if multiple dolphins protect a single pier. The final probability of collision of a ship with a given bridge is the sum of all the AF numbers corresponding to the different ship size and pier combinations.

The value of N in each case is as explained in Assumption 4. PG is calculated using the function 'geometric_probability', in accordance with the procedure described in section 3.14.5.3 in the AASHTO design code. The procedure is slightly different for the Bayonne Bridge (compared to all the other bridges), as explained within comments in the code. It is called once for each pier of each bridge.

PF is computed by the function 'protection_factor', which iterates over all the dolphins for a given pier of a bridge, and adds the protections provided by them. It is called once for each pier of the bridge.

Finally, all the above computations make use of normalized locations of the piers and ship travel lines. The normalized location of any quantity is a number between 0 and 1, and signifies a location on the line joining the 'start' and 'end' points of a bridge. The value of the normalized location is computed as the ratio of the distance between the point of interest and the 'start' point of the bridge to the distance between the 'start' and 'end' points of a bridge. The normalized locations of the vessel sailing paths have been computed in advance and are provided in the input file 'ship_travel_lines.csv' (see Section 4). The normalized locations of the piers are computed by the function 'normalized_pier_center_location'. this function first computes the latitude and longitude of the base of the perpendicular dropped from the center of the pier onto the line segment joining the 'start' and 'end' points of the bridge. This is done to account for any bends or curves in the bridge. Once the foot of the perpendicular has been located, its normalized location is computed and assigned to be the normalized location of the pier.

4 Summary of input files

There are several input files required by 'collision_probability_calculator.py' to compute the collision probabilities. The contents of each file are summarized below:

1. **'bridge_parameters.csv':** This file lists many of the relevant general parameters of each bridge. The first column of the file lists the names of each bridge. The second column lists the length of the bridge, which is measured as the great circle distance between the 'start' and 'end' points of the bridge. The last four columns list the longitude (X coordinate) and latitude (Y coordinate) of the 'start' and 'end' points of the bridge. These points are located by eye to lie on the banks of the waterway that the bridge spans. Thus, the length of the bridge is not the length of the span between any piers, but rather the length of the bridge over water. The third column of the file lists the number of piers of the bridge that are susceptible to collision with ships, while the fourth column lists the number of dolphins protecting these vulnerable piers.

2. **'bridge_piers.csv':** This file contains information on the piers of the bridge susceptible to collision by a ship. The first column once again lists the name of the bridge, while the second column assigns an index number to each pier of the bridge. The third and fourth columns list the longitude (X coordinate) and latitude (Y coordinate) of the center of the bridge, while the fifth and sixth columns list the dimensions of the pier. The fifth column provides the length, i.e., the dimension perpendicular to the deck of the bridge (perpendicular to the long axis of the bridge). The sixth column provides the width, i.e., the dimension parallel to the axis of the bridge. This is the dimension of the pier that faces oncoming traffic.

3. **'bridge_protections.csv':** This file lists the parameters of the dolphins of the ship. First column lists the names. The second column assigns an index number to each dolphin, while the third column lists the corresponding index number of the pier protected by the dolphin. (These pier index numbers are the same as the pier index numbers from the 'bridge_piers.csv' file.) The fourth column lists the traffic factor; this is the factor that should be multiplied with the traffic N to give consideration to the fact that a dolphin only protects a pier from traffic from one direction. The symmetry assumption on the traffic is used to compute these parameters. The fifth and sixth columns list the Longitude (X coordinate) and the Latitude (Y coordinate) of the center of the dolphin. The seventh column provides the distance between the center of the dolphin to the center of the corresponding pier. The eighth and final column lists the diameter of the dolphin.

4. **'ship_travel_lines.csv':** This file has only two columns; the first lists the name of each bridge, while the second lists the location of the centerline of the vessel sailing path (as discussed in assumption 6). The value of this

location is a number between 0 and 1 corresponding to the ratio between the distance of the centerline from the 'start' point of the bridge (as listed in 'bridge_parameters.csv') to the distance between the 'start' and 'end' points of the bridge.

5. **'Average_Widths.csv':** This file provides the average width of ships belonging to each of the six size classes considered, for each bridge individually. The average is computed over 6 years worth of shipping data (2018-2023), only considering ships that pass underneath the given bridge. The first column notes the name of the bridge, and the subsequent columns list the average width for each size class in increasing order.

6. **'traffic' folder:** This folder includes 23 csv files, following the naming convention '[bridge name] Counts.csv'. Within each file, traffic is divided by year and size class. Each row corresponds to a separate year, while each column lists the total number of ships belonging to a given size class that crossed the bridge that year. The first row lists the column titles, including the ship size classes, while the first column lists the corresponding year.

References

[1] AASHTO (American Association of State Highway and Transportation Officials). 2017. *AASHTO LRFD Bridge Design Specifications.* 8th ed.Washington, DC: AASHTO.