Identifying Characteristics of Bridges Vulnerable to Hydraulic Hazards Using Bridge Failure Data

Dominic Wirkijowski[1] and Franklin L. Moon, Ph.D.[2]

Abstract: Hydraulic failure, which encompasses scour, flooding, and overtopping, accounts for most bridge failures in the United States. The research reported herein aimed to evaluate the current practice of bridge hydraulic vulnerability assessment, and to identify characteristics of bridges most prone to hydraulic failure using a historical bridge failure database. Using an inductive, data-driven approach, a bridge failure population was created by synthesizing the National Bridge Inventory (NBI), United States Geological Survey (USGS) station streamflow, and available bridge failure data. Many failures that occurred before mid-2000 had incomplete assessments; this reflected the low share of structures with completed assessments in NBI. In subsequent years, state bridge inventories had completed assessments and hydraulic failures had significantly lower scour critical ratings, which indicated an improvement in identifying structures most vulnerable to failure. The most common bridge types within the overall inventory were generally overrepresented within the bridge failure population (especially those used before the Interstate Era). Differences in structure age, geometry, and condition ratings were most noticeable between failures and the overall waterway bridge inventory. Bridge deck width was narrower for failures and might explain a mechanistic cause of hydraulic failure. Waterway bridges built prior to the Interstate Era resemble those that have historically failed in that era. However, failures of bridges constructed within the Interstate Era did not resemble current waterway structures built in the same time period. **DOI:** 10.1061/(ASCE)CF.1943-5509.0001513.

Author keywords: Bridge failure; Hydraulic hazards; Bridge scour; Resilience; United States.

Introduction

Hydraulic hazards have historically been the leading cause of bridge failures within the United States (Wardhana and Hardipriono 2003; Cook et al. 2015), and it appears such hazards are poised to account for an increasing percentage of bridge failures in the future. While the vulnerability of the current bridge population changes slowly due to its long service life, there is growing evidence that hydraulic hazards are increasing based on the incidence of extreme weather over the last several decades. This trend is expected to continue with the rise of mean global temperatures, and, in turn, increases in mean sea level and precipitation intensity that will endanger many infrastructure sectors (Neumann et al. 2015).

The use of failure data can provide insight into the type of structures most susceptible to failure by observing those that have already failed. However, this is a particularly challenging endeavor considering that bridge failures are infrequent and not formally documented. Despite a dearth of information regarding this important issue, studies have been conducted on the failures of constructed facilities. An early effort to document bridge failures was conducted by Harik et al. (1990) who collected limited information sourced from national news media between 1951 and 1988. Failures caused by accidents and natural causes were most represented in the failure data, but limitations in comprehension prevented a more thorough analysis. The study ultimately recommended a federal division responsible for collecting, analyzing, and identifying potential deficiencies from previous bridge failure data.

The lack of a national bridge failure record represents a key barrier to informative bridge management and engineering improvements. The 1987 collapse of Schoharie Creek Bridge in New York prompted the earliest work to examine scour development and the first national bridge failure data set. The largest known bridge failure database within the United States is maintained by the New York State Department of Transportation (NYSDOT) and relies on voluntary reporting, and thus has two important shortcomings. First, the database is noncomprehensive in nature with many states electing not to report their data. Second, data quality and completeness requirements are difficult to enforce, and many of the failures documented by this database provide few details and are missing information.

The shortcomings of the NYSDOT database notwithstanding, it provides the clearest picture available as to the causes and frequencies of bridge failures within the United States. Previous studies examining this database have concluded that over half of its reported failures were the result of hydraulic hazards (Wardhana and Hardipriono 2003; Cook et al. 2015). Wardhana and Hardipriono (2003) conducted a general analysis of the database and found most failures were caused by external factors during service life. Cook et al. (2015) used the failure database to probabilistically model bridge failure rate in the United States. The study concluded that hydraulic-caused collapses were not age related and collapsed at an annual rate of 1.52×10^{-4}. It was also found that changes in bridge design specifications and maintenance regulations have not had a significant influence on reducing bridge failure rates and that there was no construction era more susceptible to failure (Cook et al. 2015). The primary contribution of these studies has been to alert agencies (and the profession at large) to the primacy of hydraulic-related bridge failures (compared with overload, settlement, fatigue,

[1]Graduate Student, Dept. of Civil and Environmental Engineering, Rutgers Univ., New Brunswick, NJ 08901 (corresponding author). Email: dwirkijowski@gmail.com

[2]Professor, Dept. of Civil and Environmental Engineering, Rutgers Univ., New Brunswick, NJ 08901.

Note. This manuscript was submitted on January 31, 2020; approved on June 2, 2020; published online on August 29, 2020. Discussion period open until January 29, 2021; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Performance of Constructed Facilities***, © ASCE, ISSN 0887-3828.

and impact-related failures). They have not, however, been able to either (1) evaluate the effectiveness of current practice to reduce the number and/or consequences associated with hydraulic-related bridge failures, or (2) identify characteristics commonly associated with bridges that fail due to hydraulic hazards.

The research reported herein aimed to build upon previous bridge failure research and adopted these two goals. The first is aimed to help improve confidence in and/or refine current practice, while the latter aims to provide agencies with a heuristic tool to help prioritize hydraulically vulnerable bridges for interventions. To accomplish these goals, this research combined the NYSDOT database of bridge failures with the National Bridge Inventory (NBI), which provides over 100 data fields for all bridges within the United States. In addition, precipitation and/or stream flow rates in the vicinity of failed bridges during the months leading up to the failure were also recorded and used to quantify the hydraulic demand that resulted in the failure. The result was a database that contained a more complete account of the specific attributes, demands, and conditions associated with bridges that had failed. Through this process, the NBI data of the failed bridge population were compared to those of current structures over waterways to identify potential surrogates that indicate a higher likelihood of hydraulic-related failure.

Data Sources

Available Bridge Failure Data

The NYSDOT bridge failure database was established following the scour-induced collapse of the Interstate 90 New York State Thruway Bridge over Schoharie Creek in 1987. Beginning the collection of data in 1990, the database was intended to document all bridge failures throughout the United States. The NYSDOT database is dependent on questionnaire responses from every state Department of Transportation (DOT) at 4-year intervals. The database currently collects six pieces of information for collapsed bridges: (1) identification (location and features intersecting), (2) year built, (3) year failed, (4) bridge material and type, (5) cause of failure, and (6) type of collapse (total or partial). NYSDOT is the only agency, including federal organizations, that actively collects and maintains a national bridge failure database in the United States (Wardhana and Hardipriono 2003; Cook et al. 2015).

NYSDOT has categorized bridge failures as either total or partial collapse (NYSDOT, personal communication, 2018). A total collapse is defined as "structures on which all primary members of a span or several spans have undergone severe deformation such that no travel lanes are passable." In contrast, a partial collapse is defined as "structures on which all or some of the primary structural members of a span or multiple spans have undergone severe deformation such that the lives of those traveling on or under the structure would be in danger." Hydraulic bridge failures will encompass both collapse types in this study.

Despite the value of data compiled by NYSDOT, the failure database is inconsistent as it contains substantial information from some states and little to no information from others. The deficiency of the failure database is evident by its incomplete, uncounted, and limited failure records. This may be due to various state DOTs' reluctance to allocate resources to reporting the requested data and particularly anything beyond the NBI Standards stipulations (Cook et al. 2015). At the DOT level, bridge failure records are rarely compiled, stored, or allocated resources for time to respond to an NYSDOT questionnaire (Cook et al. 2015). On more than one occasion, DOTs that did respond to the questionnaire reported no updates to the existing failure database, when in fact bridge failures did occur in these DOT regions and were documented in other published works (Cook et al. 2015). As a result, it is important to stress that the database does not constitute a complete listing of all failures that have occurred in the United States and, therefore, should not be used to compare bridge failure rates among states.

National Bridge Inventory

The Federal Highway Administration's (FHWA) NBI contains information of all publicly owned bridges in the United States with a length of 6.1 m (20 ft) or more. NBI records have been published annually since 1992 and are in the public domain through the FHWA. For the purposes of this study, the NBI was used to cross-identify bridges listed in the NYSDOT bridge failure database and for comparison of the bridge inventory in 2018.

The FHWA has designated NBI fields to describe structural vulnerability to hydraulic hazards. Items 60, 61, 71, 111, and 113 have been used to describe the substructure condition, channel and embankment protection condition, waterway adequacy to bridge overtopping, pier and abutment protection for navigation, and scour vulnerability ("scour critical" rating), respectively. Among these records, substructure condition and scour critical ratings are the only fields that provide inspection data describing structural condition and vulnerability to scour. Substructure condition describes the physical condition of piers, abutments, piles, fenders and footings that should be inspected for visible signs of distress including evidence of cracking, section loss, settlement, misalignment, scour, collision damage, and corrosion. The NBI does not provide information on substructure type, geometry, or depth.

A bridge is deemed "scour critical" when its abutment or pier foundation is deemed unstable due to observed scour at the bridge site (rating factor of 2, 1, or 0) or a scour potential as determined from a scour evaluation study (rating factor of 3). Code rating 4 indicates that, although the foundations are deemed stable, action is required to protect exposed areas from the effects of additional erosion or scour. It is assumed that the scour critical rating has been based on an engineering evaluation, which includes consultation of the NBIS field inspection finding (FHWA 2001). If a bridge is found to be scour critical, a plan of action is developed and implemented to address the deficiencies that have made the foundation unstable for an observed or calculated scour condition.

USGS Streamflow Data

The synthesis of failure dates and site geography permitted the analysis of reported hydraulic demands for the failure population. Failures containing complete failure dates were considered; of these, approximately two-thirds of the cases were total collapses and one-third were partial collapses. Since no failure site had streamflow data collection, United States Geographical Survey (USGS) streamflow stations within the same watershed and county were studied to provide a relative hydraulic demand at failure. The maximum peak and average daily flowrates were obtained for each USGS station within a week period of each failure date and compared to peak flow flood scenarios (e.g., 100, 50, 25, 10-year) determined by each station. The largest peak failure scenario exceeded was documented for each reported failure for several reasons: (1) the number of USGS stations varied by failure site based on location; (2) many stations did not provide peak flow events more frequent than 50- or 100-year, so stations that provided 25- or 10-year peak flow events could characterize lower hydraulic demands; and (3) the largest peak event would likely best represent the hydraulic demands experienced for bridge failure.

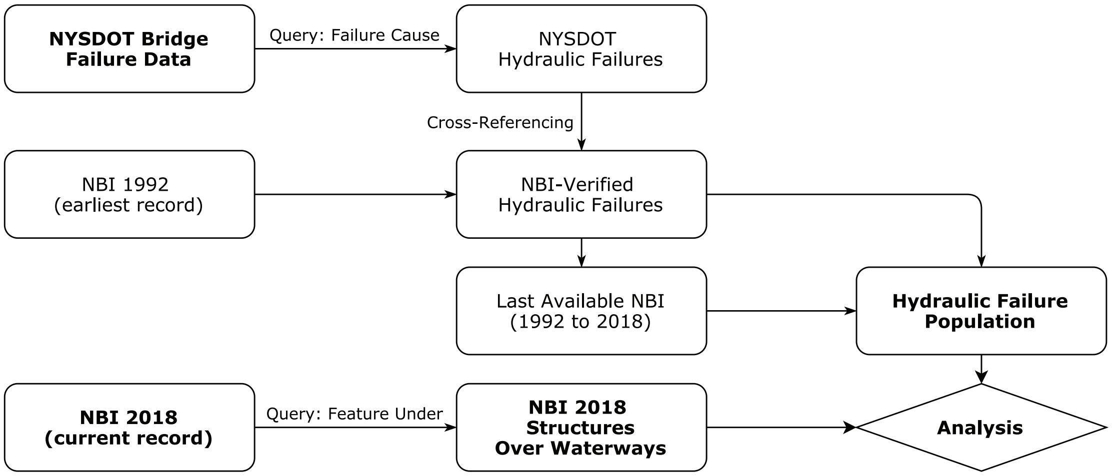

Fig. 1. Failure population formation.

Cross-Referencing of Available Sources

Bridge Failure Cross-Identification

The largest shortcoming of the NYSDOT bridge failure database was the lack of descriptive information, notably condition and geometry. To address this deficiency, a descriptive hydraulic failure population was assembled combining both NBI and NYSDOT bridge data sets. The formation of this descriptive hydraulic failure population is shown in Fig. 1. First, the NYSDOT bridge failure data set was queried for failures caused by hydraulic hazards between 1992 and 2018; these were years for which NBI records were available to describe reported failures. Keywords used to select bridges by failure cause included "hydraulic," "flood," "scour," "ice," "storm," "hurricane," and "erosion." The query found 439 self-reported hydraulic failures among 30 states. States with a substantial representation of hydraulic failures were considered to observe regional bridge design and inspection practices. The largest state subpopulations included 119, 66, 33, and 32 hydraulic failures.

An attempt was made to cross-identify each bridge failure with state NBI records in 1992, the earliest available year in which a structure could still be in service. Without structure number identifiers, the process was not automated and relied on manually pairing similar descriptors. Datapoints shared between NBI and NYSDOT data sets included "route carried," "feature intersected," "year built," and "bridge design." The cross-identification process was presented with challenges from many sources of error and data-reporting practices. There were 35 NYSDOT failure entries that noted that a structure did not have a total length of at least 6.1 m (20 ft) and, as a result, would not be included within the NBI database. Other entries had incomprehensive, varying, or misspelled descriptors or missing year built information. Incorporating structure numbers with reported failures would have greatly expedited and helped form a more complete failure population.

Each bridge within the failure population was then traced to its last NBI record available prior to its year of failure. This record provided vital information on condition ratings and evaluations prior to hydraulic failure. In 80% of cases, an NBI record was in the year prior to hydraulic failure, but 10% of the failure population had its last available entry recorded over a decade prior.

Hydraulic Failure Population

The cross-identification process yielded sufficiently large "matched" failure subsets in three states; the largest state failure subpopulations contained 63, 32, and 30 structures and will herein be referenced as States A, B, and C, respectively. The three states maintained relatively large bridge populations in the United States greater than 10,000 structures in the NBI 2018. Referred herein as the failure population, the total set was composed of 125 bridges confidently cross-identified out of 217 hydraulic failures reported by the three state DOTs. Of the 125 failures, 81 were classified as a total collapse, 35 were a partial collapse, and 9 did not specify. A total of 94 failure entries specified a complete date of failure; the remaining 31 only specified the failure year.

The following analysis assumed that, even if the reporting was voluntary and thus not comprehensive, the failures reported were representative of the total number of failures that have occurred. For this reason, states with fewer than 30 failures (that were able to be confidently cross-referenced with NBI) were not considered. Like the NYSDOT failure database, the failure population could not reliably be used to compare the total number of hydraulic failures among states. The primary reason was that the failure reporting was voluntary and that differences may simply reflect a difference in the completeness of state DOT reporting. With those shortcomings understood, the failure population still provides the most comprehensive collection of bridge failures reported by state DOTs.

Data Synthesis

Failure Correlation with Hydraulic Demand

Because flow measurements were not directly available at bridge sites, a limited analysis was performed examining the effect of flow measurements to failure type (e.g., total or partial collapse). All 94 failures that reported a precise date noted "scour" or "flood" as the failure cause; since both causes are considered equivalent within this research, this field was not considered. Fig. 2 presents the hydraulic demand for the failure population with respect to average daily flow and peak flow at failure. A distinction between total and partial collapse was noted for failures in all peak flood scenarios. Observing the flow measurements relative to peak flow scenarios uncovered the following:

- Total collapses were characterized by larger peak flood events for both flow measurements. This result was intuitively expected

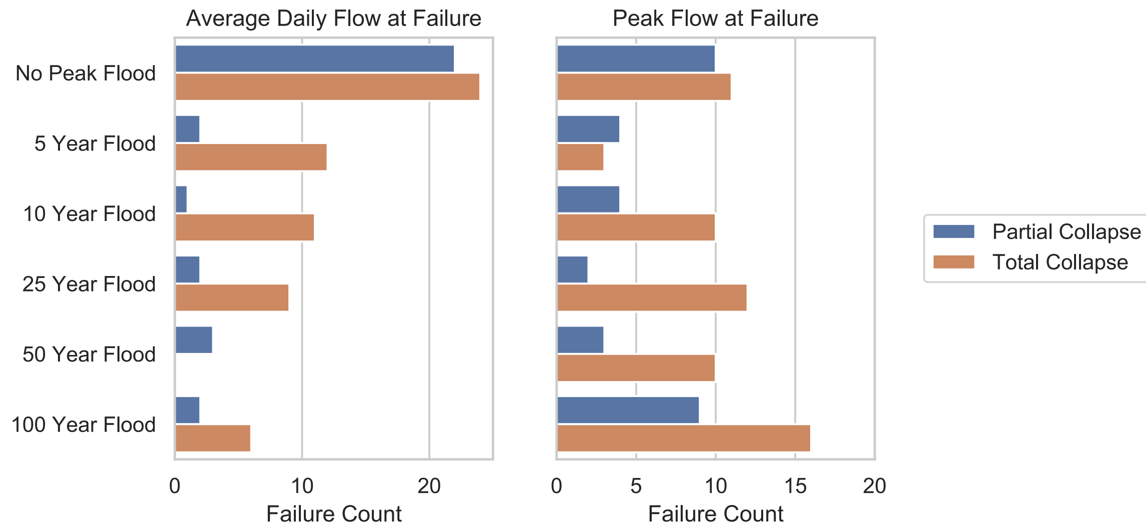

Fig. 2. Average daily flow and peak flow at failure.

as total collapses would generally require substantially larger hydraulic demand than partial collapses.

- The average daily flow for failure events only exceeded peak flood scenarios for 51% (48/96) of failures; 79% (38/48) of these failures were classified as a total collapse. Despite only capturing 61% (38/62) of total collapse failures beyond a flood scenario, the average daily flow metric was able to "filter out" partial collapses as only 31% (10/32) exceeded a flood scenario.
- In contrast, 78% (73/96) of peak flow measurements exceeded a station peak flood scenario; 69% (22/32) of partial collapses and 82% (51/62) of total collapses registered above a station scenario. The 100-year flood was the most common flood scenario observed by both collapse types. The larger (100-, 50-, and 25-year) flood scenarios (52/94) were more represented than smaller (5- and 10-year) flood scenarios (21/94) by peak flow, indicating that large and most infrequent hydraulic demands were required to cause bridge failure.

Evaluation of Scour Assessment Practices

To observe the effectiveness of previous hydraulic hazard assessments, scour critical ratings of the failure population and current waterway bridge populations were compared and presented in Table 1. Bold failure percentages (%) represent the percent per state of the observed failures for the specified rating level. The subsequent estimated NBI 2018 percentages represent the total number of waterway bridges in a state with specified rating levels divided by total number of state waterway bridges in the NBI 2018. Approximate NBI 2018 values were used to withhold the identity of states. In comparing the distribution of scour critical ratings and failures for each state shown in Table 1, several observations and potential explanations were apparent

- Only 18% of reported failures were designated as scour critical, and 45% (57/125) were given a scour critical rating between 0 and 5 (right column in Table 1). Considering that less than half of failures were rated in this "vulnerable" range indicates a very low correlation between scour critical rating and failure. When only considering structures that had a scour assessments conduct, 61% (57/93) of the failure population was designated with a rating between 0 and 5.
- Approximately 25% (31/125) of the failure population was assessed as stable with a calculated scour above footing, which corresponds to the lowest level of hydraulic vulnerability. This could be due to two possible reasons: (1) the bridges were exposed to extreme hydraulic events that exceeded initial hazard assessments, or (2) the assessment process underestimated the vulnerability of those bridges to hydraulic hazards. In either case, the scour assessment process failed to identify these vulnerable structures prior to failure.
- Structures with countermeasures installed had the fewest reported failures (4%) and were similarly represented in the 2018 waterway bridge inventory (5%). As a result, it appears that the installation of scour countermeasures did not drastically reduce a bridge's likelihood of failure (as the percentages within the failed population and total population are approximately the same).
- State A had the largest share of failures that were rated as scour critical (32%) as well as the largest share of its total population rate scour critical bridges in the NBI 2018 (15%). In contrast, State C did not have any scour critical failures and an insignificant (approximate 0%) population of scour critical bridges in the 2018 waterway inventory.

Table 1. Scour critical rating distribution for hydraulic failure population

Scour critical rating	Percent of total observed failures (count), ~percent of NBI 2018 per state			
	State A	State B	State C	Total
0–3 Scour critical	**32% (20)** ~15%	**9% (3)** ~5%	**0% (0)** ~0%	**18% (23)** ~10%
4–5 Stable within limits	**40% (25)** ~35%	**0% (0)** ~35%	**30% (9)** ~40%	**27% (34)** ~35%
6 No evaluation	**14% (9)** ~0%	**50% (16)** ~0%	**23% (7)** ~0%	**26% (32)** ~0%
7 Countermeasures	**3% (2)** ~10%	**0% (0)** ~0%	**10% (3)** ~5%	**4% (5)** ~5%
8–9 Stable above footing	**11% (7)** ~40%	**41% (13)** ~60%	**37% (11)** ~55%	**25% (31)** ~50%
Total failures	**100% (63)**	**100% (32)**	**100% (30)**	**100% (125)**

Note: The bold numbers represent observed failures for the specified rating level. NBI 2018 percentages were rounded to the nearest 5%.

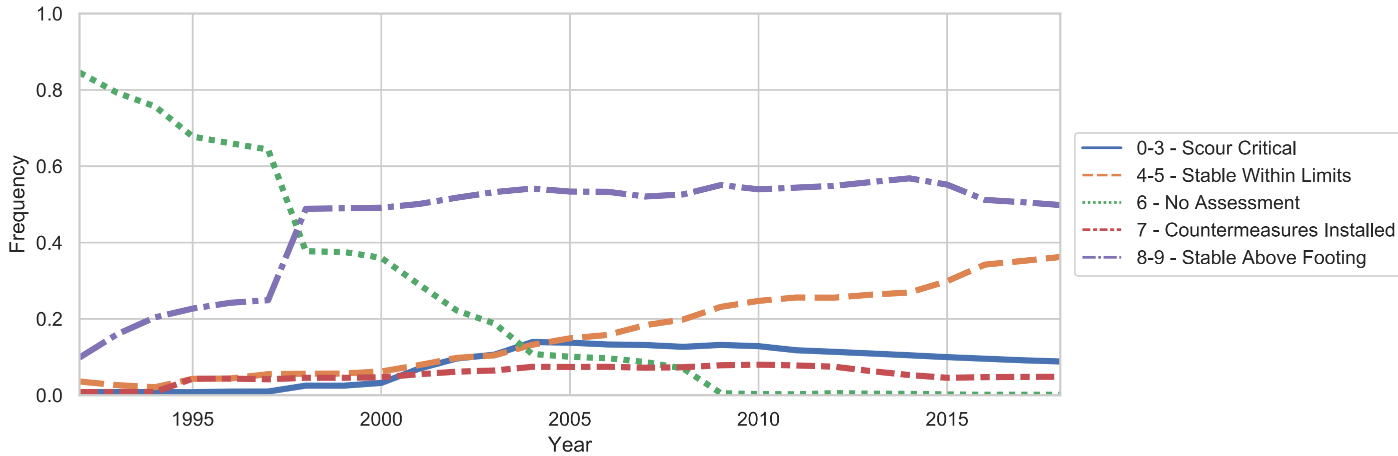

Fig. 3. Frequency of scour critical ratings per NBI year.

- Half of State B failures did not have a scour critical assessment conducted. Most of the remaining failures deemed the scour condition stable above footing (41%) with only 9% corresponding to bridges identified as scour critical.
- 71% (45/63) of failures in State A were rated between 0 and 5, which is significantly greater than States B and C (9% and 30%, respectively). This may indicate that State A, despite reporting the most vulnerable bridges in NBI (which may appear unfavorable), conducted the most accurate scour assessments out of the three states.

One key finding was that 26% (32/125) of the failure population had not undergone a scour vulnerability assessment. Since the initiation of the National Bridge Scour Evaluation Program in 1988, the FHWA has mandated state DOTs to conduct scour evaluations for bridges over waterways. Observing the timeline of NBI scour critical ratings in Fig. 3, the number of bridges without assessments dropped from over 80% in 1992 to under 40% by the late 1990s. By mid-2000, FHWA reported that most state DOTs had completed over 90% of their bridges needing a scour evaluation (FHWA 2006). The analysis found that through the completion of assessments between 1992 and 2005, the frequency of scour critical bridges in NBI surged from 0% to a peak of 14%. State DOTs have since addressed this sudden increase by a steadily decreasing share of scour critical bridges; 10% of bridges were deemed scour critical in 2018. The eventual completion of scour assessments in the three study states would occur in the late 2000s. Failures lacking scour assessments thus serve to highlight the importance of performing hydraulic evaluations on a regular basis.

To further examine the effectiveness of scour assessments, the scour critical ratings of the failure population were observed with respect to failure year. A boxplot timeline of scour critical ratings is presented in Fig. 4. Akin to the changes seen in the NBI inventory, the scour critical ratings for the failure population suggested changes with hydraulic assessment practices had occurred within three decades. In the first half of the study period (1992–2005), the failure population had scour critical ratings mostly incomplete (33%) or ratings that appeared to greatly underestimate the vulnerability (i.e., 29% were deemed stable above footing, 20% were deemed stable within calculated scour limits, and only 13% were rated scour critical). In contrast, during the second half of the reporting period (2006–2018), most bridge failures were associated with either scour critical bridges (38%) or bridges deemed stable within limits (49%). This observed refinement in failure scour critical ratings showcased the improvements that state DOTs have recently identified for structures most vulnerable to hydraulic failure.

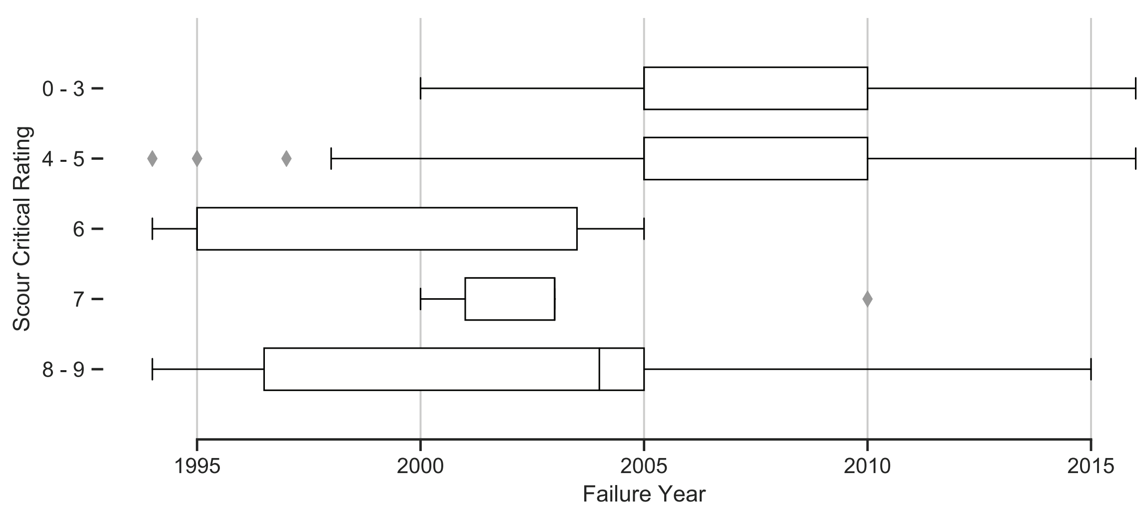

Fig. 4. Scour critical ratings with respect to failure year.

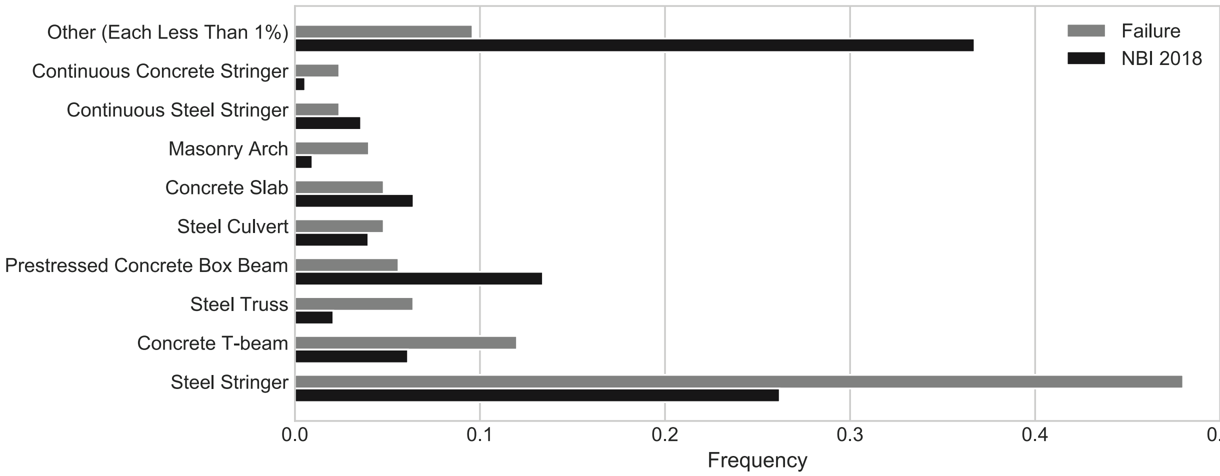

Fig. 5. Frequency of bridge design types.

Comparison with the Current State Waterway Inventory

The failure population was compared to current waterway bridges to identify key differing characteristics. The current waterway bridge population is comprised of State A, B, and C bridges listed in the NBI 2018 with Item 42B "Type of Service Under Bridge" marked as "waterway" (code 5). This population consisted of approximately 40,000 bridges and served as a comparative population exposed to hydraulic hazards. By discovering differences between bridges that failed and current structures, one can attempt to describe hydraulic failure mechanisms and identify structures most vulnerable to hydraulic failure in the future.

Bridge design was observed to determine which types failed most frequently relative to their representation in the NBI 2018. NBI Items 43A "Structure Type" and 43B "Structure Kind" were combined to characterize "designs" for failed bridges and waterway bridges within the NBI 2018. Fig. 5 presents the results for designs that contribute more than 1% of either population. From this figure, the following observations are drawn:

- Nearly half (48%) of the failure population was comprised of simply supported steel stringer bridges, which were by far the most common followed by concrete T-beams (12%). Even as these two bridge types were among the most prevalent in the current waterway inventory, they only accounted for 26% and 6%, respectively, and thus were overrepresented within the failed population by a factor of 2. Steel trusses (6%) and masonry arches (5%) were among the oldest bridge designs and were similarly overrepresented in the failure population.
- Steel stringer, concrete T-beam, steel truss, masonry arch, and concrete continuous stringer were all overrepresented within the failed population compared to the waterway inventory. However, these designs inherently possessed the greatest sensitivity to hydraulic failure with older age and design standards. While masonry arch and concrete continuous stringer were the most overrepresented within the failed population by percentage, they accounted for such a small portion of the total population (less than 1%) that these observations may not necessarily be reliable.
- Bridge types that were underrepresented within the failed population included continuous steel, prestressed concrete, and all remaining design types grouped into "other." These designs are more modern and thus their underrepresentation may be due more to improved engineering practices than anything inherent within these bridge types.

A second analysis was conducted by observing the relationship between the prevalence of each structure type and year built. Fig. 6 presents kernel densities (or histograms) of both failure population and waterway population. In conjunction with the previous findings, two observations are apparent:

- A large share of the failure population was built before the mid-twentieth century; 60% were built before 1940 and 80% were built before 1960. Nearly half (46%) of the failure population was constructed between 1920 and 1940, a time when US bridge production was amplified through publicly funded economic revitalization projects. It was in this period that state DOTs created standard bridge plans to be used on state highways and made them available for local government engineers (NCHRP 2005).
- Failures in each bridge design category appear to have occurred in the earlier years of practice. Design practices for each bridge type have since been refined and failures have diminished with increased experience and improved design criteria. This highlights the evolution of bridge engineering and how failures are less frequent with modern bridge design practices.

NBI Bivariate Correlation Matrix

A Pearson parametric correlation test was used to investigate the dependence between NBI fields within the waterway bridge population as well as the bridge failure population. Considering the differences in population size, the tests were conducted separately for these two populations. The results for each correlation matrix were compared to identify initial differences between the two bridge populations. In order to compare superstructure and substructure condition ratings, nonnumerical values (designated as "N" rather than a value ranging from 0 to 9) were omitted from the analysis. These entries corresponded to culverts that constituted approximately 10% of failures and 20% of the current waterway inventory.

Positive correlation coefficients indicated that the two NBI parameters being compared were "in phase" (i.e., an increase in one parameter corresponded to an increase in the second parameter). Fig. 7 summarizes the results for the failure population (lower triangle) and the NBI 2018 waterway bridges (upper triangle). The correlation matrix generated shows the following observations:

- Structure year built (or the inverse of structure age) had significant positive correlations with superstructure (0.54), substructure (0.61), and sufficiency rating (0.54) for the current waterway

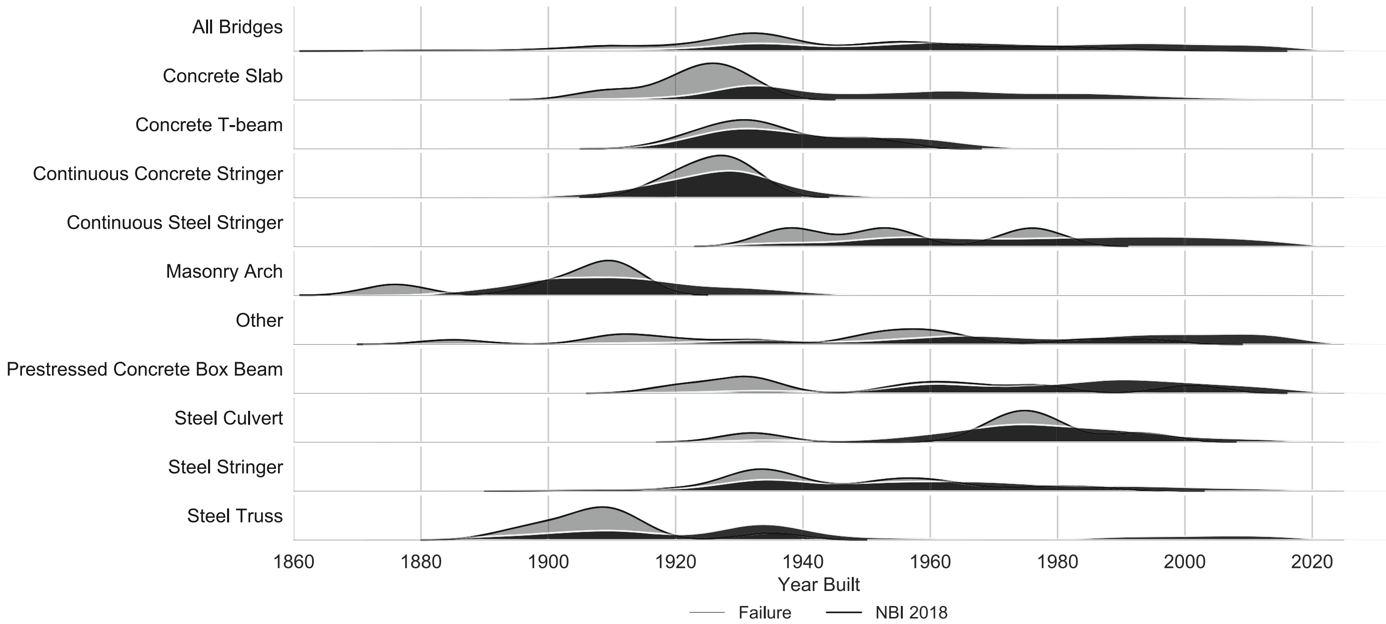

Fig. 6. Kernel densities of bridge design types.

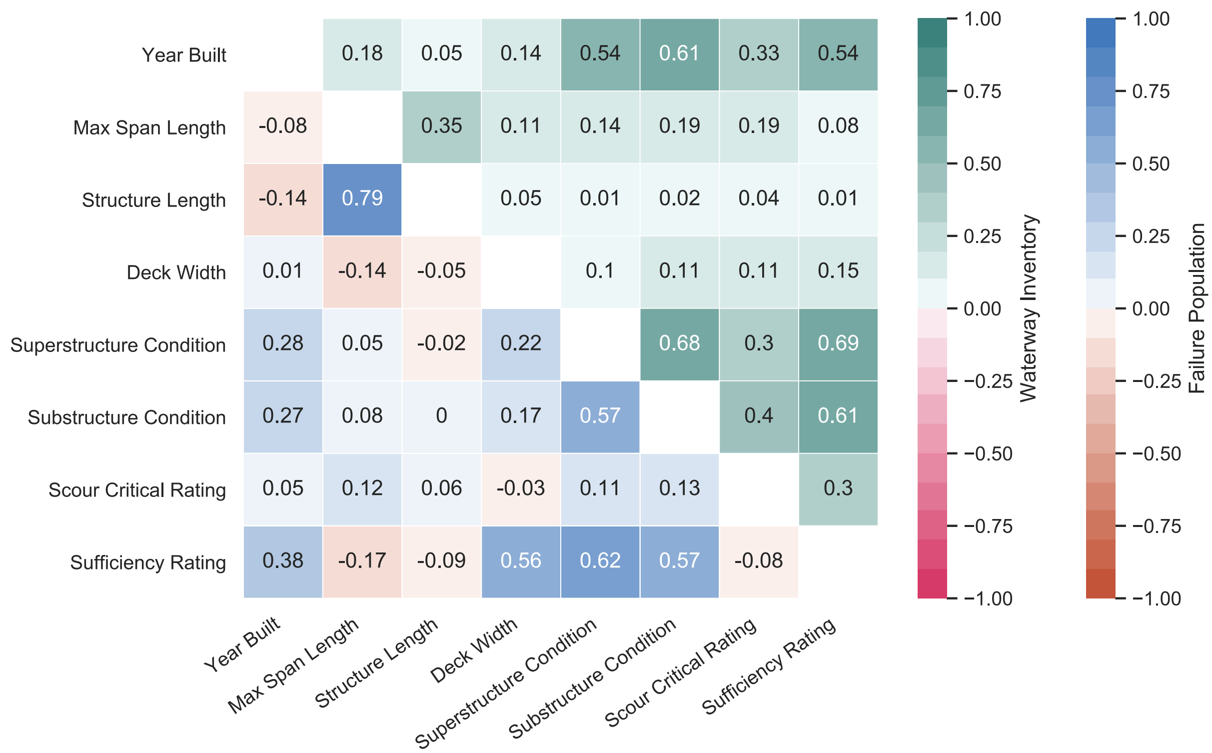

Fig. 7. Correlation matrix of NBI fields for failure population and the NBI 2018.

inventory. The failure population found a weaker correlation for each of these comparisons (0.28, 0.27, and 0.39, respectively). For the current inventory, a positive correlation between year built and condition ratings was intuitively expected as the parameters reflect structural deterioration.

- Correlation tests with scour critical rating yielded no significant relationships for the failure population. In contrast, the scour critical rating was positively correlated with superstructure (0.30), substructure (0.40), and year built for the waterway inventory (0.33).
- Structure length and max span length were more positively correlated for the failure population (0.79) than current inventory (0.35), while number of main spans was less correlated with structure length (0.54 and 0.78, respectively). This indicated that the failure population comprised of structures with fewer spans, mostly single-span structures.
- Deck width was more positively correlated with sufficiency rating for the failure population than the waterway inventory (0.56 and 0.15, respectively). Sufficiency rating rewards structures for increased structural adequacy and traffic serviceability.

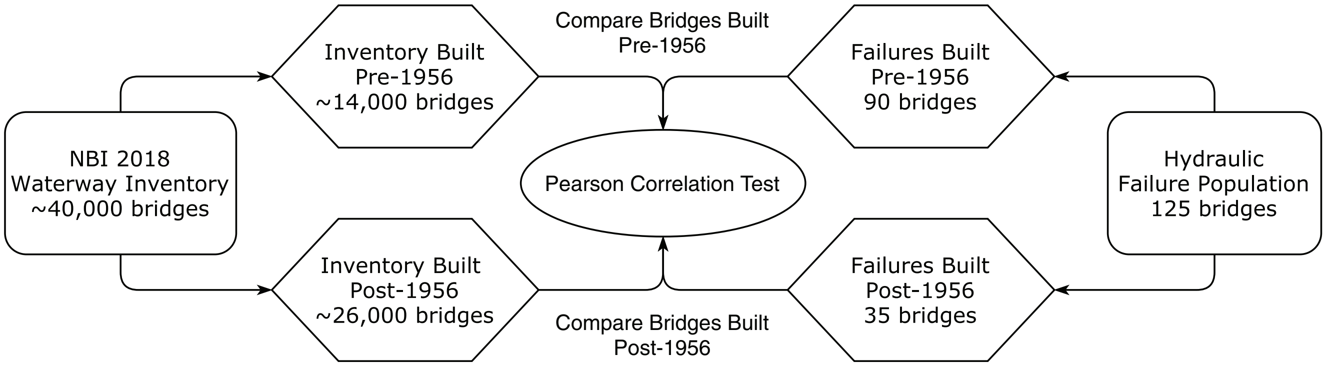

Fig. 8. Subdivision of failure population and waterway inventory by year built.

Therefore, this comparison indicated the penalty incurred on the failure population sufficiency ratings based on narrowness. It was also intuitively expected that deck width correlated with year built by accommodating larger vehicles and roadway shoulders. The analysis found this correlation weak for the waterway inventory (0.14) but not present in the failure population (0.01). Newer bridges had slight increases in deck width within the waterway bridge population but not within the failure population.

Comparison of Bridge Built before and after the Interstate Era

Subdividing Bridge Populations

Considering the influence of year built for many data fields, further analysis was conducted after subdividing the failure population and current inventory into two time periods. The date of 1956 was selected as the cut-off year in which the Federal Highway Aid Act of 1956 was passed, establishing the Interstate Highway System and permanently changing highway design (NCHRP 2005). Both failure population and the NBI 2018 waterway inventory were subdivided by year built. Structures built before 1956 were deemed "Pre-1956" and those remaining referred to as "Post-1956." Fig. 8 provides a schematic illustrating this subdivision process. An important observation from this process was the proportion of bridges built before and after 1956 in both bridge sets. Most of the failure population was built before highway standardization, while the 2018 waterway inventory had more structures built after.

Correlations between Pre- and Post-1956 Bridge Populations (Failure Population versus Waterway Bridge Population)

To further investigate significant differences between the failure population and NBI 2018 waterway structures, a bivariate Pearson correlation test was conducted between each subdivided bridge population per state. For example, the State A Pre-1956 failure population was compared to its Pre-1956 waterway inventory. The correlation results of States A, B, and C were then used to compute weighted averages for each NBI value. Positive correlation coefficients indicated that the NBI parameter distributions are similar; negative correlations indicated distributions that were inversely related. Categorical (qualitative) NBI data, such as design load and condition ratings, were binned by unique values. Continuous (quantitative) data, such as year built and geometry, were binned at geometric intervals to provide an approximately equal class width that adjusted for nonnormal distributions and extreme outliers. If a response bin was not represented by both failure and current state subpopulations, it was omitted from the correlation calculation to avoid overfitting-bias. Table 2 summarizes the findings using the three-state average per subpopulation. Fig. 9 contains kernel densities of each field to provide context to each subpopulation distribution and similitude. The following observations can be made from the analysis:

- Structure year built showed the greatest polarity between subpopulations. The correlation between all failure populations and current waterway inventory was not significant (0.160). The correlation of Pre-1956 was quite strong (0.808), indicating that failures built before 1956 were representative of the bridge inventory built at that time. In contrast, the Post-1956 correlation was weak and negative (−0.200) and best characterized by the distribution shapes; the Post-1956 waterway inventory is relatively uniform while its failure population is skewed toward older structures.
- Structure length had consistently high correlations for all populations considered. This indicated that, regardless of year built, structure length was not a significant characteristic of hydraulic failures.
- The correlations for deck width varied by bridge population. While a moderate correlation was seen between the overall populations (0.548), the relationship was stronger for the Pre-1956 analysis (0.772) and weaker for Post-1956 (0.361). The overall failure population had narrower deck widths than the waterway bridge population. The average deck width for all failures was 6.7 m (22 ft) compared to 9.1 m (30 ft) and 9.7 m (32 ft) for Pre and Post-1956 waterway bridges, respectively. The Post-1956 population had 34% of structures built with a deck width less

Table 2. Pearson correlation coefficient between failure and NBI 2018 waterway bridges

NBI field	Pearson correlation coefficient: failure versus NBI 2018 waterway: Total	Pre-1956	Post-1956
Year built	0.160	0.808	−0.200
Structure length	0.945	0.925	0.911
Deck width	0.548	0.772	0.361
Sufficiency rating	−0.055	0.124	0.183
Design load	0.480	0.779	0.378
Scour critical rating	0.609	0.698	0.355
Superstructure rating	0.486	0.790	0.390
Substructure rating	0.423	0.598	0.494

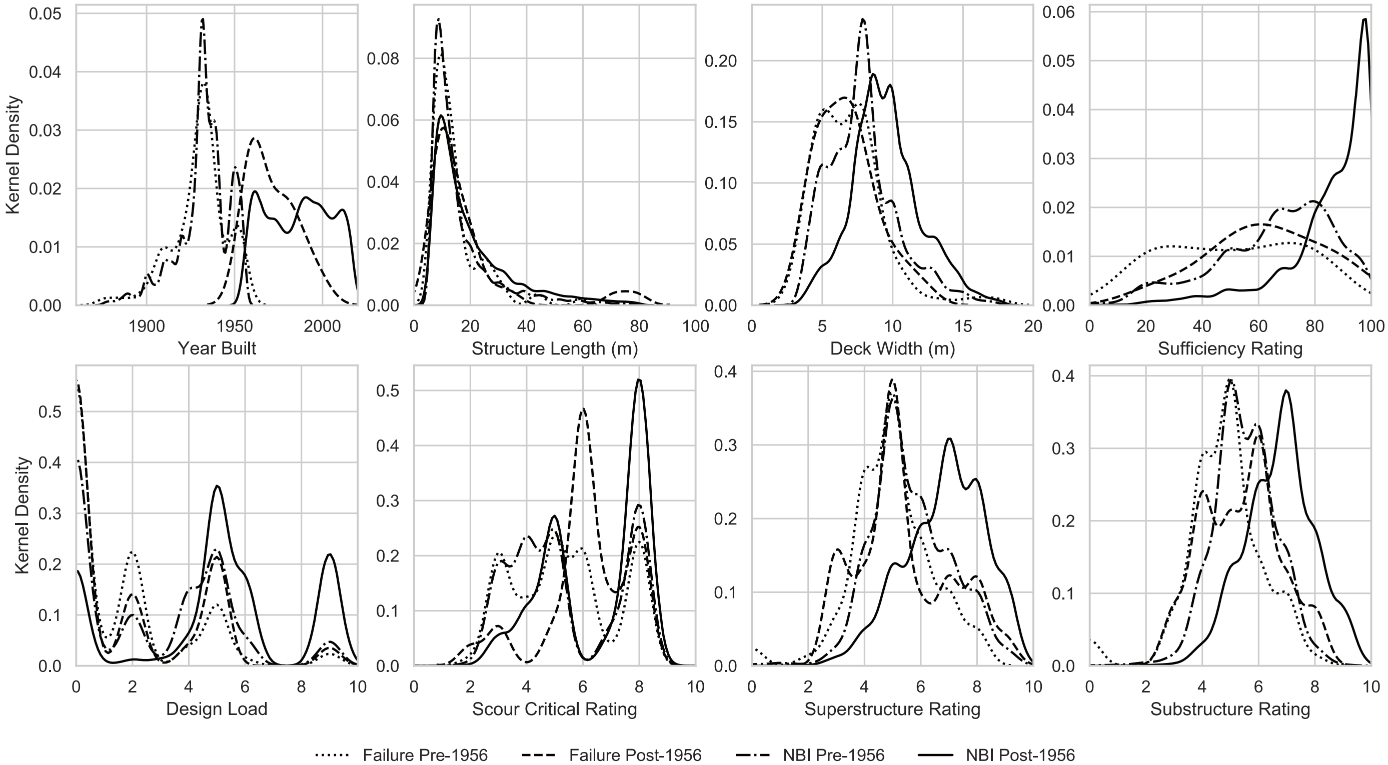

Fig. 9. Kernel densities for selected NBI fields.

than 8 m (26 ft), which characterized half of Pre-1956 structures (52%) and most failures (78%).

- Sufficiency rating was poorly correlated for all populations considered. The distributions for both failure populations (Pre- and Post-) were skewed to lower ratings. The waterway bridge population had significantly higher ratings; this was especially evident with Post-1956 bridges with improved structural adequacy and traffic serviceability associated with year built and deck width data. Mean sufficiency ratings for the failure population, Pre-1956 NBI, and Post-1956 NBI waterway structures were 54, 67, and 86, respectively.
- Structure design load was moderately correlated (0.480) between all failures and the total waterway bridge population, but it displayed a stronger (0.779) correlation for Pre-1956, and slightly weaker (0.378) for Post-1956 waterway bridge populations. The differences between Post-1956 populations were most notable with the design loads associated with the failure population providing more consistency with the distribution of design load for the Pre-1956 waterway bridge populations. Most Post-1956 bridges consisted of HS-20 and HS-25 design loads. In contrast, most of failures in the same period were design for lower standards (codes 1–4) or marked as "other or unknown" (code 0); these ratings were similarly common for both the failure and total Pre-1956 populations.
- Scour critical ratings were well-correlated (0.698) for Pre-1956 populations and notably weaker (0.355) between Post-1956 populations. Two notable differences existed between Post-1956 populations: the large share of failures without assessments (code 6) and the share of current waterway inventory that rated as stable (codes 4, 5, 8, and 9).
- Superstructure and substructure rating also had similar correlation behavior as the scour critical rating and design load. While substructure rating had smaller differences between total, Pre-, and Post-1956 correlations (0.423, 0.598, and 0.494, respectively), the superstructure rating varied considerably between populations (0.486, 0.790, and 0.390, respectively).

Discussion of Pre- and Post-1956 Correlations

The distinction of bridges before and after highway standardization highlighted how modern designs have potentially reduced hydraulic vulnerability. The comparison between bridge failures and the waterway inventory built before 1956 suggested that failures had little characteristic differences excluding sufficiency rating. In contrast, the comparison between failed bridges and total waterway bridges built after 1956 found that failures were noticeably different except for structure length. In fact, the kernel distributions for failures built in that time imitated those of bridge populations built Pre-1956. This indicated that failures built more recently (Post-1956) were more similar to older (Pre-1956) failures than their contemporaries still in service.

The considerable difference in deck width between the failure population and waterway inventory may represent a hydraulic mechanistic phenomenon, considering that bridge deck width is well-correlated with foundation width. Solid wall-type substructures are most likely used considering that most failures were single-spanning, and as a result, wider bridges would generally accommodate more scour before reaching a point of instability. Hydraulic laboratory experiments found that bridge width (or pier length) has no appreciable effect on local scour depth unless the pier is skewed to the flow (AASHTO 2012). Scour depth is a hydraulic calculation that considers streambed conditions and quantifies the scour demand that a structure may experience. Structure width may instead contribute to a "scour capacity," or the ability to maintain stability at certain scour hole size or depth. This study also found that bridge skew had a correlation coefficient of 0.985 between bridge populations, and therefore was not a significant attribute in distinguishing hydraulic failures from the general bridge population.

Structure year built was very influential on many parameters including design type, design load, condition ratings, and geometry (considering standard designs). While these may represent causal effects, one age-dependent development that was not indicated by the NBI data was the increased prevalence of the use of deep foundations. The standardized construction of deep foundations was compelled by machine drilled shafts becoming more widespread during the early to mid-twentieth century (FHWA 2010). However, the costs of deep foundation construction may have prevented smaller structures from consideration in favor of more economic shallow footings. The extensive use of deep foundations for waterway structures was further adopted following the Interstate Highway Act of 1956 and certainly after the Schoharie Creek Bridge collapse.

Conclusions

A failure data set was used to construct a descriptive failure bridge population to evaluate hydraulic hazard vulnerability practices and to identify characteristics correlated with failure. Hydraulic failure records from three states were used to cross-identify 125 bridges in the NBI between 1992 and 2018. The cross-identification process required structure verification through matching common datapoints in the NBI and the failure data set. This intensive process could not be automated due to many inconsistencies, errors, and missing information. Incorporating a structure number with reported failures would have greatly expedited the process and helped form a more complete failure population. The last available NBI record was used to mark the structural condition and scour critical rating assigned prior to failure. Nearby USGS streamflow stations were used to characterize the hydraulic demands at failure. The bridge failure population was generated, analyzed, and compared to waterway inventory bridges from the NBI 2018 database in the same states. Significant findings of this work included:

- Most (78%) peak flow measurements at failure exceeded a nearby USGS station peak flood scenario. Failures in which the average daily flow exceeded a peak flood scenario were likely associated with a total collapse. The 100-year flood was the most frequently exceeded scenario by peak flow for both total and partial collapses. For peak flow, the larger flood scenarios were more represented than the smaller scenarios, and the most extreme hydraulic demands reported by the respective USGS stations occurred for most of the observed hydraulic failures.
- Scour assessments have improved in identifying structures most vulnerable to hydraulic hazards. Although bridges deemed scour critical or stable within scour limits (ratings 0–5) made up only 45% of the entire failure population, most failures that occurred before 2005 had incomplete or overestimated scour assessments. In subsequent years, failures had significantly lower scour critical ratings, indicating that the vulnerability assessments have improved in identifying structures most vulnerable to failure.
- The completion of scour critical assessments for the NBI inventory had a noticeable influence on the failure population. One quarter (26%) of the failure population did not have a scour assessment conducted, which constituted one-third (33%) of all failures that occurred before 2005. Through the completion of scour assessments between 1992 and 2005, the frequency of scour critical bridges in the NBI surged from 0% to a peak of 14%. State DOTs have since addressed scour-critical bridges (bringing their total percentage down to 10%) and completed all scour assessments with no bridges coded 6 within the NBI 2018. Failures lacking scour assessments serve to highlight the importance of performing hydraulic evaluations on a regular basis.
- Generally, the most common bridge types within the inventory failed most frequently. Steel stringers and concrete T-beams were overrepresented in the failure data by a factor of two and together made up 65% of the failure population. Failure frequencies can be attributed to design age; older designs such as masonry arches and steel trusses were similarly overrepresented. In contrast, newer designs such as prestressed concrete and continuous designs were significantly underrepresented.
- Bridge designs have evolved to reduce hydraulic vulnerability, because failures generally affected bridge types in the early years of implementation. Most failures were structures built before 1956, the beginning of the Interstate Era. Nearly half of the failure population was constructed between 1920 and 1940, a two-decade period when bridge construction was expanding in the United States.
- Structure year built had a strong correlation with structure condition ratings for the NBI 2018 waterway bridges (ranging 0.54–0.61) than the failure population (ranging 0.28–0.39). Scour critical rating was also well-correlated to condition ratings for waterway inventory, but it was not significantly correlated to any NBI field for the failure population.
- The stronger correlation between structure length and max span length in the failure population (0.79) than current population (0.35) suggested that most of the failures were single-span bridges.
- Deck width was more positively correlated with sufficiency rating for the failure population (0.56) than the waterway inventory (0.15). Considering that sufficiency rating rewards structures for increased structural adequacy and traffic serviceability, the stronger correlation for failures was attributed to a penalty attributed to deck narrowness. Correlations also found that newer bridges had slight increases in deck width within the waterway bridge population (0.14) but not within the failure population (0.01).
- The bridge failure population was significantly older, narrower, and had lower structural condition ratings than the 2018 waterway bridge inventory. While older and lower condition ratings were expected, the increased prevalence of failure among bridges with smaller widths represents a new finding and it appears well-correlated with increased hydraulic vulnerability. Considering that 75% of failures occurred in single-span bridges with narrower foundations, the observed correlation is consistent with the fact that such foundations can accommodate lower levels of scours (as measured transverse to the direction of traffic) prior to reaching instability. Over the previous century of bridge construction, bridge width has gradually increased while failed structures have had consistently narrower dimensions, with a weak correlation coefficient of 0.1. The average deck width for the failure population was 6.7 m (22 ft), compared to 9.1 m (30 ft) and 9.7 m (32 ft) for Pre and Post-1956 waterway bridges, respectively.
- Waterway bridges built before the Interstate Era (Pre-1956) have a strong resemblance to reported failures built in that era. In contrast, waterway structures built before 1956 are more dissimilar to failed structures built after 1956 and more alike pre-Interstate Era bridges.

The significant characteristics that varied between the failure population and current NBI waterway inventory can be leveraged to identify structures most vulnerable to hydraulic failure in the future. Future work in this field can determine the contribution of each significant parameter, such as deck width, to bridge hydraulic vulnerability. A framework for identifying vulnerable structures may be developed using the findings of failure data to combat expected changes in hydraulic demand due to climate change.

Data Availability Statement

Some or all data, models, or code used during the study were provided by a third party. Direct requests for these materials may be made to the provider as indicated in the Acknowledgments.

Acknowledgments

This research was conducted by a student sponsored by the Rutgers Coastal Climate Risk and Resilience (C2R2) Program, a fellowship supported by the National Science Foundation (NSF) Grant No. 1633557. Special thanks to Sreenivas Alampalli from NYSDOT for communication and the use of the bridge failure data set.

References

AASHTO. 2012. *Hydraulic engineering circular No. 18: Evaluating scour at bridges*. 5th ed. Washington, DC: AASHTO.

Cook, W., P. J. Barr, and M. W. Halling. 2015. "Bridge failure rate." *J. Perform. Constr. Facil.* 29 (3): 04014080. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000571.

FHWA (Federal Highway Administration). 2001. *ACTION: Revision of coding guide, item 113—Scour critical bridges*. Washington, DC: FHWA.

FHWA (Federal Highway Administration). 2006. *ACTION: Compliance with the national bridge inspection standards—Background/plan of action for scour critical bridges/scour focus states/item 113 coding issues/unknown foundations/tidal bridges*. Washington, DC: FHWA.

FHWA (Federal Highway Administration). 2010. *Drilled shafts: Construction procedures and LFRD design methods*. Geotechnical Engineering Circular No. 10. Washington, DC: FHWA.

Harik, I. E., A. M. Shaaban, H. Gesund, G. Y. Valli, and S. T. Wang. 1990. "United States bridge failures, 1951–1988." *J. Perform. Constr. Facil.* 4 (4): 272–277. https://doi.org/10.1061/(ASCE)0887-3828(1990)4:4(272).

NCHRP (National Cooperative Highway Research Program). 2005. "*A context for common historic bridges types*." Project 25–25, Task 15. Washington, DC: NCHRP. http://onlinepubs.trb.org/onlinepubs/archive/NotesDocs/25-25(15)_FR.pdf.

Neumann, J. E., et al. 2015. "Climate change risks to US infrastructure: Impacts on roads, bridges, coastal development, and urban drainage." *Clim. Change* 131 (1): 97. https://doi.org/10.1007/s10584-013-1037-4.

Wardhana, K., and F. C. Hardipriono. 2003. "Analysis of recent bridge failures in the United States." *J. Perform. Constr. Facil.* 17 (3): 144–150. https://doi.org/10.1061/(ASCE)0887-3828(2003)17:3(144).