Bridge Condition Rating Forecast—Survival-Based Models

Raka Goyal, Ph.D., P.E.

Introduction

The survival-based models on InfoBridge™ are probabilistic bridge deterioration models based on a methodology that combines survival analysis and Markov chain theory. The modeling framework was first developed and implemented on bridge deck, superstructure, and substructure components of the North Carolina State bridge inventory (Goyal 2015; Cavalline et al. 2015). In the current research, the model's implementation has been substantially expanded and improved for the more complex application to the entire nationwide bridge inventory.

Survival analysis is a statistical approach that analyzes the time until death or failure—in the case of bridge components, which is the time spent in a condition rating until it deteriorates to a lower condition rating. Survival analysis allows determination of the survival probability of the component in that condition rating at any point in time. The main advantage of survival analysis is that it can account for incompletely recorded durations commonly found in duration-based data, such as condition rating observations truncated at the beginning and end of the recording period. Instead of discarding the incomplete observations, a mechanism called "censoring" is used to include the observations in the analysis, which permits a more realistic estimate of duration compared to other statistical approaches.

In this study, the Cox proportional hazards model (PHM) (Cox 1972) has been used for the analysis of condition rating durations. Cox PHM is a semiparametric approach that does not make any assumptions about the shape of the distribution and can be used to analyze even unimodal hazard functions associated with some infrastructure components. For the survival-based models developed in this study, the transition probabilities of the Markov chain are calculated from the PHM survival functions at each condition rating, carrying the advantage of survival analysis probabilistically over the entire lifecycle. Additionally, multivariable effects at each condition rating are quantified in terms of PHM hazard ratios and are used to modify the Markov chain transition probabilities. In this way, the effects of time dependence and exogenous factors on deterioration, as analyzed through survival analysis, are incorporated in the Markov chain to develop a probabilistic lifecycle deterioration model.

Proportional Hazards Bridge Deterioration Model

The survival function ($S(t)$) associated with a bridge-component condition rating is the cumulative survival rate of bridge components in the condition rating. The hazard rate or hazard function is the instantaneous risk of transitioning to a lower rating at time (t) conditional to survival until that time. In the Cox PHM, the hazard rate (h(t,z)) is defined as the product of a time-dependent, nonparametric-baseline hazard function ($h_0(t)$) and a time-independent exponential function, representing the effects of design variables (z) through regression coefficients (β):

$$h(t,\vec{z}) = h_0(t)e^{\vec{z}\vec{\beta}} = h_0(t)e^{(z_1\beta_1+z_2\beta_2+\cdots+z_n\beta_n)} \tag{1}$$

In Equation 1, the time-independent exponential function represents the effects of covariates, or explanatory factors, on the hazard rate. The baseline hazard rate is associated with the baseline variables, which are assigned a value of zero.

Proportional hazards lifecycle deterioration modeling

The proportional hazards bridge deterioration model (Goyal 2015) can be represented in terms of structure-specific transition probability matrices (P_i) associated with each year (i) of the prediction period.

$$\boldsymbol{P_i} = \begin{bmatrix} P_{99}^{HR_9} & 1-P_{99}^{HR_9} & .. & .. & 0 & 0 & 0 & 0 \\ 0 & P_{88}^{HR_8} & .. & .. & 0 & 0 & 0 & 0 \\ \vdots & \vdots & .. & .. & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & .. & .. & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & .. & .. & P_{44}^{HR_4} & 1-P_{44}^{HR_4} & 0 & 0 \\ 0 & 0 & .. & .. & 0 & P_{33}^{HR_3} & 1-P_{33}^{HR_3} & 0 \\ 0 & 0 & .. & .. & 0 & 0 & 0.85 & 0.15 \\ 0 & 0 & .. & .. & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

The matrix elements (P_{ij}) represent the probability of the bridge component transitioning from condition rating i to condition rating j, assuming there is no improvement in condition rating and that the component deteriorates by no more than a single rating in 1 year. In each row, P_{kk}, on the diagonal, is the baseline stay-the-same transition probability for condition rating k. If S_k(t,z) is the baseline survival function of a bridge deck associated with condition rating k for a bridge described by the vector of covariates z, the baseline transition probability (Goyal 2015; Goyal et al. 2020) of staying at the same condition rating over the next annual reporting cycle at any time (t) is given by:

$$P_{kk} = \frac{S_k(t+\Delta,\vec{z})}{S_k(t,\vec{z})} = \frac{S_k(t+1,\vec{z})}{S_k(t,\vec{z})} \; for\ \Delta = 1\ year \tag{3}$$

Since the sum of probabilities in each row should be equal to 1, the baseline transition probability of deteriorating to a lower rating is 1 – P_{kk}. The baseline transition probabilities are uniquely modified for each bridge or category of bridges using structure-specific hazard ratios (HR_k) calculated by multiplying the PHM hazard ratios for the factors associated with a bridge that are identified as significantly influencing deterioration at condition rating k. A stationary implementation of the proportional hazards deterioration model is used in this study, in which stationary or constant transition probabilities (P_{kk}) are obtained by averaging the yearly transition probabilities across the duration of the PHM-baseline survival functions for each condition rating k. The resulting stationary-transition probability matrix (P) is used in a homogeneous Markov chain to predict the future condition-state vector (Z_n) of a bridge component after n years if its present condition-state vector (Z_0) is known using:

$$Z_n = Z_0(P)^n \tag{4}$$

The condition-state vectors comprise of the probabilities of the bridge component being at all possible condition ratings at any given time. The future state vector (Z_n) can be multiplied by the column vector (R) of condition ratings to produce the expected condition rating (E) of the bridge component after n years.

$$E = Z_n R \tag{5}$$

This document briefly describes the implementation of the proportional hazards deterioration model to develop deck, superstructure, substructure, and culvert deterioration models for the nationwide bridge inventory, with more details to follow in a journal publication currently under preparation. Detailed background and theoretical development of the proportional hazards deterioration model can be found elsewhere (Goyal 2015; Goyal et al. 2017; Goyal et al. 2020).

Data Structuring

The nationwide survival-based models are based on the Federal Highway Administration (FHWA) National Bridge Inventory (NBI) data spanning from 1983 to 2019 for highway bridges nationwide (FHWA 2020b). The NBI files store historical inspection records of all bridges in the United States with a span greater than 20 feet (ft), with more than 100 data items per bridge, per year of recorded service (FHWA 1995). The first step in organizing the data for modeling was to identify, query, and extract the NBI fields relevant to deterioration modeling from the yearly inspection records of individual bridges, and assemble a continuous record of bridge condition-related data from 1983 to 2019. In addition to NBI data, yearly bridge-specific climate data attributes extracted from the National Aeronautics and Space Administration's (NASA) Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) database were similarly processed and included in the modeling dataset.

Bridges with concrete decks constitute an overwhelming percentage (greater than 80 percent) of the NBI and are a focus area of the Long-Term Bridge Performance (LTBP) Program. For survival-based deterioration modeling, bridges were classified based on deck material, main structure material, and design type. Further, since duration-based analysis is best served by long duration records, and in the interest of optimizing the use of computational resources, it was decided that only bridges with continuous records of 25 years or more be used for model development. The initial classification of major concrete deck bridge types was expanded to include all bridges in the NBI database, and in addition to deck models, deterioration models were developed for superstructure and substructure components and culverts in each category. The bridge categories for survival-based models with the number of bridges in the historical database selected for model development are provided in Table 1.

Table 1. Bridge categories for survival-based deterioration models.

Bridge Type	**Deck Type (NBI Item 107)**	**Main Structure Material (NBI Item 43A)**	**Main Structure Design (NBI Item 43B)**	**No. Bridges Included in Model Development (1983–2019)***
Steel Girder	Concrete (Cast-in-Place and Precast Panels)	Steel and Steel continuous	Stringer/Multibeam or Girder	37,906
Prestressed Concrete Girder		Prestressed concrete and Prestressed concrete continuous	Stringer/Multibeam or Girder	26,324
Prestressed Concrete Box Girder			Box Beam or Girders – Multiple and Single or Spread	16,834
Concrete Slab		Concrete and Concrete continuous	Slab	26,373
Concrete Girder			Stringer/Multibeam or Girder, Tee Beam, and Channel Beam	25,491
Concrete Deck – Other		All other concrete deck bridges		22,919
Timber Deck	Timber	Timber	Stringer/Multibeam or Girder	10,146
Steel Deck	Steel	Steel and Steel continuous	Stringer/Multibeam or Girder	5,018
Other	All other bridges			23,918
Culvert–Concrete		Concrete and Concrete continuous	Culvert	40,617
Culvert–Other	All other culverts			7,637

* Rebuilt and reconstructed bridges separated based on NBI items “Year Built” and “Year Reconstructed.”

Design variables for proportional hazards survival analysis

Data was further preprocessed to extract all observations of the response variable, which is the observed continuous duration at each NBI condition rating analyzed. For each observed duration, corresponding censoring information was compiled in a separate vector of the same size as the response variable, but stored as a binary variable of either 0 or 1, depending on whether the observations were classified as completely observed or censored. In this study, all continuous observations that were truncated at the beginning (1983) or end year (2019) of the NBI database were classified as right censored, presuming that the actual duration of the condition rating was longer than observed due to the limited timespan of the data recording period. Further, all observations where an increase in condition rating was observed rather than deterioration were considered as complete or uncensored under the assumption that observed improvements in condition rating reflected maintenance action performed because of an unrecorded transition to the lower rating.

For proportional hazards analysis, the descriptive information on each structure, such as its functional classification, traffic characteristics, design parameters, and other details contained within the NBI historical records that could potentially influence the deterioration rates of specific bridge components, were considered as potential explanatory factors. Some factors were derived from two or more NBI fields. For example, “age” refers to the age of the bridge at the beginning of the observed condition rating duration and is calculated based on the corresponding data year and the year when the bridge was last built or reconstructed. Additionally, average daily traffic per lane (ADTL) and average daily truck traffic per lane (ADTTL) were calculated by dividing the average daily traffic (ADT) and average daily truck traffic (ADTT), respectively, by the number of through-traffic lanes on a bridge also recorded in the NBI (ADTT mentioned here refers to the number of trucks, which was first calculated from the percentage ADTT recorded in the NBI). ADTL and ADTTL are new variables that were introduced in the nationwide study to account for the width of the roadway in studying the impact of traffic on bridge deterioration. In addition to NBI-based variables, potentially influential climate factors, such as the frequency of snow days and freeze–thaw cycles from NASA’s climate database, which were associated with each bridge based on the geospatial information in the NBI, were also included in the analysis.

Each explanatory factor is organized into categories designated by one or more design variables to which bridges are classified based on either binary or reference cell coding. Variables such as ADTL, ADTTL, age, and maximum span length, which are continuously recorded, are divided into categories of approximately equal frequency of occurrence for each component based on weighted averages computed across the available bridge records. The categorical ranges for the variables differ for disparate bridge categories (Table 1), depending on the different statistical distributions of variables associated with each bridge category. The design variables included in the development of survival-based models for concrete deck bridges nationwide with corresponding baseline categories are listed in Table 2. Category ranges developed for steel stringer bridges are provided for illustration. The subsequent steps involving multivariable proportional hazards regression were performed individually on each of the distinct condition rating-specific sets of dependent and independent PHM regression inputs, extracted for each category of bridges.

Table 2. Design variables included in proportional-hazards survival analysis of concrete-deck bridges nationwide. © 2020 Raka Goyal.

Factor	Baseline Category*	Design Variable*
Deck Type	Cast-in-Place Deck	Precast Panel Deck
Span Type	Simple Span	Continuous Span
Functional Class	Noninterstate	Interstate
Average Daily Traffic per Lane (ADTL)	ADTL (≤ 124)	ADTL2 (124–1037)
		ADTL3 (1037–4113)
		ADTL4 (> 4113)
Age (years)	Age (≤ 19)	Age2 (19–30)
		Age3 (30–42)
		Age4 (> 42)
Skew	No Skew	Skew
Reconstruction	Original/Rebuilt	Reconstructed
Average Daily Truck Traffic per Lane (ADTTL)	ADTTL (≤ 4)	ADTTL2 (4–63)
		ADTTL3 (63–335)
		ADTTL4 (> 335)
Wearing Surface	No Wearing Surface	Monolithic Concrete
		Integral Concrete
		Latex Concrete
		Low Slump Concrete
		Epoxy Overlay
		Bituminous
		Timber
		Gravel
		Other
Deck Membrane	No Membrane	Deck Membrane
Deck Protection	No Protection	Deck Protection
Maximum Span (m)	Max Span (≤ 13)	MaxSpan2 (13–20)
		MaxSpan3 (20–28)
		MaxSpan4 (> 28)
Number of Spans	Single Span	Multiple Spans
Highway System	Non-NHS	NHS
Snow Days	Zero Snow Days	SnowDays2 (1–50)
		SnowDays3 (50–84)
		SnowDays4 (> 84)
Freeze–Thaw Cycles	Zero Freeze–Thaw Cycles	Freeze–Thaw Cycles > 0

*Category values in parentheses are for concrete deck steel stringer bridges.

Proportional Hazards Deterioration Model Development

The variables that were found to be statistically significant at each condition rating using PHM regression were further processed through a best subset selection algorithm to optimize the size of the model without compromising its reliability. The survival functions developed using the best subset model incorporate the effect of the most significant explanatory variables on the deterioration rate over individual condition ratings. The condition rating-dependent best subsets and associated hazard ratios obtained for the significant explanatory factors identified in the proportional hazards deck model for steel stringer bridges nationwide are summarized in Table 3.

Table 3. Best subset factors and hazard ratios in proportional hazards deck deterioration model for concrete deck steel stringer bridges nationwide. © 2020 Raka Goyal.

Best Subset Factor	**Hazard Ratios at Condition Rating**						
	9	8	7	6	5	4	3
Precast Panel Deck	1	1	1	1	1.3441	1	1
Continuous Span	1	1.1706	1.1159	1	1.1335	1.0834	1
Interstate	1.1563	1.1236	1	1.0490	1.0961	1	1
ADTL2	1.2038	1.3289	1.2514	1.0899	1.1891	1.2324	1
ADTL3	1.4176	1.5090	1.3309	1.0836	1.2801	1.3543	1
ADTL4	1.5699	1.5758	1.3086	1	1.2509	1.4047	1
Age2	1.3224	1.8314	1.3247	1	1.0897	1	1
Age3	1.2510	2.2992	1.5920	0.9229	1	1	1.3195
Age4	1	2.1078	1.5662	0.8961	0.8493	0.7681	1
Skew	1	1.0867	1	1.0332	1.0559	1.0666	1
Reconstructed	1.2342	1.3057	1.1734	0.9396	1	0.8950	1
ADTTL2	1	1	0.9049	1.0949	1	1	1.2290
ADTTL3	1.2680	1.1079	0.9411	1.1228	1	1	1.3699
ADTTL4	1	1.1271	0.9752	1.2282	0.9745	1	1.4987
Monolithic Concrete	1	0.7393	1	1	1	1	1
Integral Concrete	1	0.8819	1.2956	1.3557	1.2698	1	1
Latex Concrete	1	1	1.3748	1.2275	1.3355	1.2089	1
Low Slump Concrete	1	1	1	1	1	1	1.5727
Epoxy Overlay	1	1	1.4442	1	1.8996	1	1
Bituminous	1.2685	0.8080	0.9168	1.0340	1	0.9149	1
Timber	1	1	1	2.0292	1	1	1
Gravel	1	0.6090	1	1	1	1	1
Other	1.5512	1	1.3531	1.3578	1.4503	1	1
Deck Membrane	1.3861	1.2141	1.1512	1	1	1	1
Deck Protection	1.2274	1.1343	0.9547	1	1	1	1
MaxSpan2	1	1	1	1.1227	1.1196	1.1522	1
MaxSpan3	1	1.1308	1.0559	1.1497	1.1937	1.2091	1
MaxSpan4	1	1.1428	1.0739	1.1821	1.1802	1.4226	1.2537
Multiple Spans	1	1	0.9389	1	1	1	1
NHS	1	1	1	1	1.0752	1	1
SnowDays2	1	1	1.1574	0.7227	0.8134	1	0.6655
SnowDays3	1	1	1.5243	1	1	1	0.6756
SnowDays4	1.2057	1	1.8469	1.1313	1	1	1
Freeze–Thaw Cycles	1	1.4493	0.6982	1	1.2735	1	1

A hazard ratio value of 1 signifies a lack of influence on the deterioration rate, and indicates that the factor was not included in the best subset for that rating. For example, in Table 3, precast panel deck is included only in the best subset associated with condition rating 5. A hazard ratio value less than 1 indicates that the factor is associated with a reduced rate of deterioration, and a value greater than 1 indicates that the factor is associated with an increased rate of deterioration relative to the baseline category. As seen in Table 3, revealing the varying effects of the same factors at different condition ratings across the bridge lifecycle is a unique aspect of the proportional hazards deterioration model.

The final step in the proportional hazards deterioration model development is the calculation of baseline transition probabilities at each condition rating and assembling the Markov chain transition probability matrix, shown in equation 2, for future condition forecasting. Sufficient historical condition rating data was available in the nationwide NBI database to develop survival function-based transition probabilities for all the condition ratings from 9 to 3. The stay-the-same transition probability for condition rating 2 was prescribed as 0.85 to prevent the deterioration models from converging abruptly to condition rating 2. Condition rating 1 is the lowest, or absorbing state, with a stay-the-same transition probability of 1. Figure 1 shows deck deterioration models which exhibit the combined effects of variables including ADTL, age, and maximum span length on the deck deterioration rates of concrete deck steel stringer bridges, with other variables at baseline values. Figure 1 illustrates the depth of information revealed by the proportional hazards deterioration models based on the NBI.

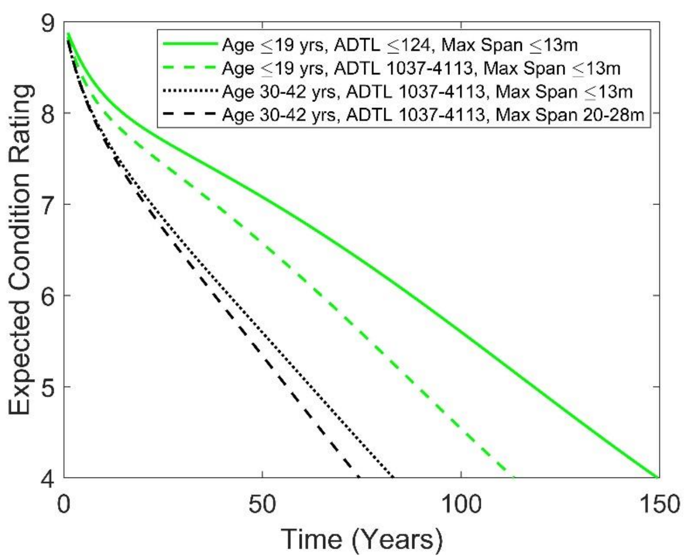
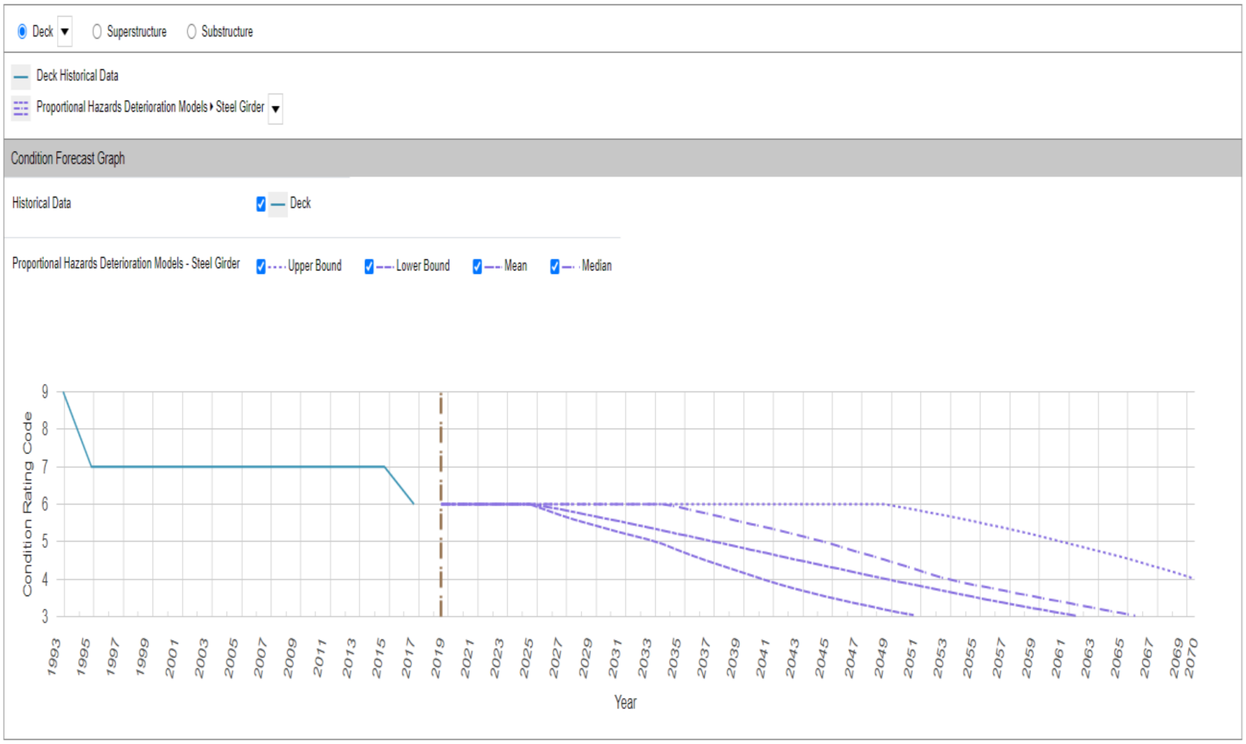


Figure 1. Proportional hazards deck deterioration models for concrete deck steel stringer bridges showing the effect of ADTL, age, and maximum span length over the bridge lifecycle.

Proportional hazards deterioration model plots for individual bridges in each bridge category (Table 1) are available on InfoBridge (FHWA 2020a). In addition to the mean predicted condition rating, prediction curves associated with cumulative probabilities of 75 percent (lower bound), 50 percent (median), and 25 percent (upper bound) are also displayed, as shown in Figure 2.



Source: FHWA.
Figure 2. Survival based deck condition forecasting curves on LTBP InfoBridge (FHWA 2020a).

References

Cavalline, T.L., Whelan, M.J., Tempest, B.Q., Goyal, R., and Ramsey, J.D. (2015). *Determination of Bridge Deterioration Models and Bridge User Costs for the NCDOT Bridge Management System.* Final Report No. FHWA/NC/2014-07, North Carolina Department of Transportation, Raleigh, NC.

Cox, D.R. (1972). "Regression Models and Life-Tables." *Journal of the Royal Statistical Society, Series B (Methodological), 34*(2), pp.187–220. JSTOR, New York, NY.

FHWA. (1995). *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges*. Report FHWA-PD-96-001, Office of Engineering, Bridge Division, Federal Highway Administration, Washington, D.C.

FHWA. (2020a). "LTBP Infobridge." (website) Washington, D.C. Available online: https://infobridge.fhwa.dot.gov/data, last accessed November 20, 2020.

FHWA. (2020b). "Download NBI ASCII Files." (website). Washington, D.C. Available online: https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm, last accessed March 1, 2020.

Goyal, R. (2015). *Development of Survival-Based Framework for Bridge Deterioration Modeling with Large-Scale Application to the North Carolina Bridge Management System*. Ph.D. Dissertation, University of North Carolina at Charlotte. Charlotte, NC.

Goyal, R., Whelan, M., and Cavalline, T.L. (2017). "Characterizing the Effect of External Factors on Deterioration Rates of Bridge Components Using Multivariate Proportional Hazards Regression." *Structure and Infrastructure Engineering, 13*(7), pp.894–905. Taylor & Francis Online, Abingdon-on-Thames, United Kingdom.

Goyal, R., Whelan, M., and Cavalline, T.L. (2020). "Multivariable Proportional Hazards-Based Probabilistic Model for Bridge Deterioration Forecasting." *Journal of Infrastructure Systems, 26*(2). American Society of Civil Engineers, Reston, VA.