

Research Article

Complex Survival System Modeling for Risk Assessment of Infant Mortality Using a Parametric Approach

Hang Chen,¹ Maryam Sadiq ,² and Zishen Song ¹

¹Department of Electronics and Information, Xi'an Jiaotong University, Shaanxi 710049, China

²Department of Statistics, University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan

Correspondence should be addressed to Zishen Song; zishens@sina.com

Received 6 December 2021; Revised 6 March 2022; Accepted 31 March 2022; Published 19 April 2022

Academic Editor: Diego Pinto

Copyright © 2022 Hang Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Pakistan is still one of the five countries contributing to half of the child deaths worldwide and holds a low ratio of infant survival. A high rate of poverty, low level of education, limited health facilities, rural-urban inequalities, and political uncertainty are the main reasons for this condition. Survival models that evaluate the performance of models over simulated and real data set may serve as an effective technique to determine accurate complex systems. The present study proposed an efficient extension of the recent parametric technique for risk assessment of infant mortality to address complex survival systems in the presence of extreme observations. This extended method integrated four distributions with the basic algorithm using a real data set of infant survival without extreme observations. The proposed models are compared with the standard partial least squares-Cox regression (PLS-CoxR), and higher efficiency of these proposed algorithms is observed for handling complex survival time systems for risk assessment. The algorithm is also used to analyze simulated data set for further verification of results. The optimal model revealed that the mother's age, type of residence, wealth index, permission to go to a medical facility, distance to a health facility, and awareness about tuberculosis significantly affected the survival time of infants. The flexibility and continuity of extended parametric methods support the implementation of public health surveillance data effectively for data-oriented evaluation. The findings may support projecting targeted interventions, producing awareness, and implementing policies planned to reduce infant mortality.

1. Introduction

Strong statistical survival techniques are the demand of the era for authentic and reliable results for deeply examining complex survival and mortality patterns. Nonparametric survival techniques including the Kaplan-Meier product-limit method [1], the Gehan's generalized Wilcoxon test [2], and the log-rank test [3] were extensively used in older times. The Cox's regression model remained the most popular and widely used semiparametric survival technique if the proportional hazards assumption is fulfilled [4]. In recent times, flexible parametric models (FPM) are considered as a better alternative to nonparametric and semiparametric methods as they produce estimates with higher efficiency and lower standard errors [5]. In addition, these models consider full likelihood to draw more precise inferences

and easily interpretable results. So far, the FPM has been employed various probability distributions to estimate survival functions. The exponential probability distribution supports as the baseline to handle survival time. The Weibull, Gompertz, generalized gamma, and generalized F-distribution are commonly practiced too. The FPM is also able to efficiently investigate the relationship of covariates with survival response [5]. The partial least squares-Cox regression (PLS-CoxR) integrates PLS with the Cox model to address survival time response with collinear covariates [6] since the Cox regression is restricted with inflexible estimates of the cumulative hazard and survival functions as being incomplete. Hence, the PLS-CoxR model is restricted in the long-term estimation with unsmooth functions.

The flexible parametric models (FPMs) are recommended to compute hazard and cumulative hazard

functions for covariates to extrapolate the survival model. The FPM can estimate continuous survival and hazard functions instead of a step representation due to its flexibility [7].

Despite considerable improvement towards increasing infant survival, nearly six million child deaths are recorded every year, before attaining their fifth birthday [8]. By the end of 2015, a minor proportion of developing countries have met the fourth target of Millennium Development Goal (MDG) which is intended to increase the child survival rate by two-thirds [9]. The recently described Sustainable Development Goals (SDG) seek to forward the objectives originated by the MDG. The third SDG is to reduce the under-five mortality rate (U5MR) to 25 deaths per 1000 live births by 2030 [10]. Previous literature evidenced that five countries including China, Congo, Nigeria, India, and Pakistan possess nearly half of under-five mortality in the world [11]. Pakistan has the sixth largest population in the world with 188 million people [12]. In 2018, Pakistan's infant mortality rate (IMR) was 61 deaths per 1000 live births. Due to political instability, civil conflicts, poverty, lower educational level, unavailability of health facilities, and disparities regarding the area in Pakistan, 70% MDG targets were not achieved [13]. Understanding the factors affecting infant mortality is significantly informative to health professionals, practitioners, and health policymakers for the improvement of population health status through effective interventions.

Within this line, the partial least squares flexible parametric model (PLS-FPM) is developed to analyze the complex survival systems in the presence of extreme observations for risk and hazard assessment [14]. The present study extended the PLS-FPM to collinear predictors having moderate trend observations using four alternative probability distributions.

The results exposed the flexible dynamics of the extended method to obtain smooth survival and hazards estimates in the presence of multicollinearity. This model can be implemented in the field of genetics, biology, engineering, medicine, social sciences, or behavioral sciences for system reliability and risk assessment. The formal statements of the problem are the following:

- (i) Selection of optimum model by execution of four distribution integrated with the PLS-FPM oversimulated and real data set having collinear predictors and moderate observation
- (ii) Identification of significant risk factors of infant mortality in Pakistani

2. Methodology

The PLS-CoxR model is considered as the benchmark method in the present study, and the PLS-FP model with four different distributions is the proposed technique.

2.1. The Cox Regression Model. The Cox model has the form

$$\lambda(t) = \lambda_o(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p) = \lambda_o(t) \exp[\beta' X], \quad (1)$$

where $\lambda_o(t)$ represents the baseline hazard function, β is the vector of regression estimates, and X denotes a $(n * p)$ matrix of predictors.

2.2. The Partial Least Squares-Cox Regression Model. The PLS-CoxR model is employed as the reference method in the present study. Suppose the survival time is represented by t and $x_j = x_{1j}, x_{2j}, \dots, x_{nj}$ be the vector of p correlated covariates with n samples. The model estimates k components for p correlated predictors and assumes the hazard estimate as

$$\lambda(t) = \lambda_o(t) \exp(\beta_1 S_1 + \beta_2 S_2 + \dots + \beta_p S_k) = \lambda_o(t) \exp[\beta' S], \quad (2)$$

where S represents a $(n * k)$ matrix of components.

2.3. Flexible Parametric Survival Model (FPSM). Let T represent a nonnegative continuous survival response and let X is the vector of predictors x_1, \dots, x_p over a sample of size n . The survival function is the probability of being alive at time t and is represented by $S(t) = \Pr(T > t)$ for a vector of covariates at time t with the cumulative distribution function $F(t) = \Pr(T \leq t)$. Then the cumulative hazard or risk function is

$$\Lambda(t) = \int_0^t \lambda(x) dx. \quad (3)$$

Any distribution ranges over $t \in [0, \infty]$, and it may serve as survival distribution. The survival distributions included in this study as FPSM are as follows:

2.3.1. The Gompertz Distribution. A survival response T following a Gompertz distribution with parameters $(b > 0, \eta > 0)$ exhibits the survival function

$$S(t) = \exp\left(-\frac{b}{\eta}(e^{\eta t} - 1)\right), \quad (4)$$

and the cumulative hazard function as

$$\Lambda(t) = \frac{b}{\eta}(e^{\eta t} - 1). \quad (5)$$

The Gompertz distribution is also an extreme value distribution with increasing hazard function.

2.3.2. The Generalized Gamma Distribution. The generalized gamma distribution with parameters (β, σ, κ) has survival function as

$$S(t) = 1 - \Gamma \kappa^{-2} \left(e^{-\beta t} t^{\kappa/\sigma}; \kappa^{-2} \right). \quad (6)$$

The hazard function of the generalized gamma function is increasing, decreasing, bathtub, and arc-shaped [15].

2.3.3. *The Generalized F-Distribution.* The density function of generalized F-distribution with $2\nu_1$ and $2\nu_2$ is

$$f(t) = (\nu_1 e^t / \nu_2)^{\nu_1} (1 + \nu_1 e^t / \nu_2)^{-(\nu_1 + \nu_2)} \beta(\nu_1, \nu_2)^{-1}, \quad (7)$$

where $\beta(\nu_1, \nu_2)$ is the beta function and then the survival function is

$$S(t) = \int_0^{\nu_2(\nu_2 + \nu_1 e^t)^{-1}} \chi^{\nu_2 - 1} (1 - \chi)^{\nu_1 - 1} \beta(\nu_2, \nu_1)^{-1} dx, \quad (8)$$

where χ denotes the chi-square distribution. This distribution is useful for testing different parametric forms as it includes other distributions as limiting or special cases.

2.3.4. *The Exponential Distribution.* The survival time T has an exponential distribution with rate parameter λ having density function

$$f(t) = \lambda \exp(-\lambda t), \quad (9)$$

then the survival function is

$$S(t) = \exp(-\lambda t), \quad (10)$$

and the cumulative hazard function is

$$\Lambda(t) = \lambda t. \quad (11)$$

Several other probability distributions can be employed in FPM. The interpretation for regression coefficients of FPM is the same as for semiparametric models. The FPM provides a more stabilized cumulative hazard function than the semiparametric model. For instance, the Weibull models produce the hazard function as a continuous straight trend. The PLSR model integrated with FPM addressing generalized gamma (GG), generalized F (GF), exponential, and Gompertz distribution is included in the present study for improved model performance for multicollinear covariates.

2.4. *The Partial Least Squares Flexible Parametric (FP) Model.* The proposed model assumes the occurrence of an event e at time t in the presence of censoring, and let X be the matrix of p correlated predictors x_1, \dots, x_p for a sample of size n . The method computes the FP model for S components (as $S \leq p$) computed from PLSR for survival response and X as a matrix of predictors. The PLS-FP model assumes that some A is equal to the number of components to be predicted (where $A \leq p$), then for $a = 1, 2, \dots, A$, the algorithm runs:

(1) Loading weights are computed by

$$w_a = X'_{a-1} t_{a-1}. \quad (12)$$

Loading weights are normalized to have length equal to 1 by

$$w_a \leftarrow \frac{X'_{a-1} t_{a-1}}{\|X'_{a-1} t_{a-1}\|}. \quad (13)$$

(2) Score vector s_a is computed by

$$s_a = X_{a-1} w_a. \quad (14)$$

The risk function for FPSM is computed as

$$\Lambda(t) = \int_0^t \lambda(s) ds. \quad (15)$$

(3) If $a < A$ return to 1

The PLS-FP model is a two-stage procedure. At the first stage, the PLS-FP regression model computes components of PLS regression with time as response outcome and correlated covariates as predictors. Then, it executes the FP model with survival time as response and components of PLSR as explanatory factors at the later stage. This method produces efficient estimates with increased accuracy for collinear predictors. Hence, it is recommended to use in the case of collinear data as it is a conjugate of PLS and FP models. The PLSR model is also coupled with a filter-based factor selection method, namely, "loading weights" to identify the significant factors [16, 17].

2.5. *Simulated Survival Data Generated from Gompertz Distributions.* The R-package namely "simsurv" is used for the generation of simulated survival data [18] with moderate observation and collinear predictors. The data follows Gompertz distribution with 0.1 and 0.1 scale and shape parameters, respectively. The correlation among predictors is established as (0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0) for 100 samples with 30 predictors.

2.6. *Infant Survival Times Data.* This study used secondary data, obtained from the Demographic and Health Surveys (DHS), gathered during 2012-2013 from Pakistan. Hence, no ethical concerns are required to conduct this study [19]. The present analysis used data set of infants aged 1-12 months in Pakistan. Due to missing and incomplete information, infants dead within one month of birth are excluded from the analysis. A total of 697 infants belonging to Pakistan and 83 predictor variables are included.

3. Results

The PLS-FPM parameterized with generalized gamma, generalized F, exponential, and Gompertz distribution are modeled on simulated data generated from Gompertz distribution to observe the variation in efficiency for multicollinear data. The left panel of Figure 1 showed the efficiency of models established by AIC and indicated that coupled with PLSR, the FPM models showed the higher efficiency over simulated

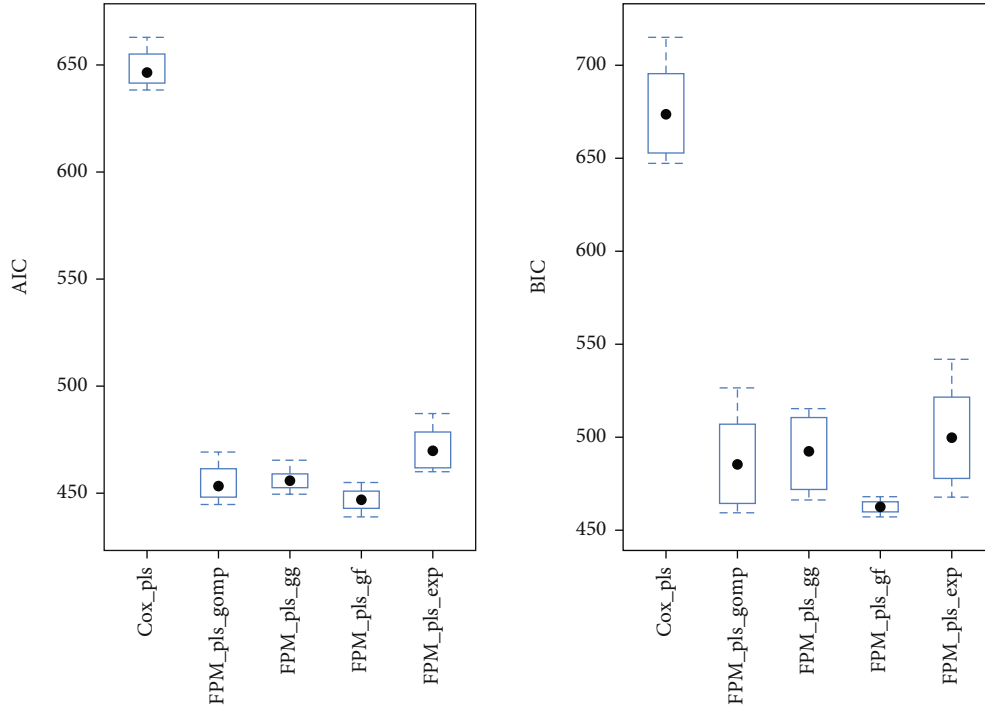


FIGURE 1: The comparison of the PLS-Cox model with the PLS-FPM parameterized over generalized gamma (GG), generalized F (GF), exponential, and Gompertz distribution for simulating survival response generated from Gompertz distribution based on AIC and BIC.

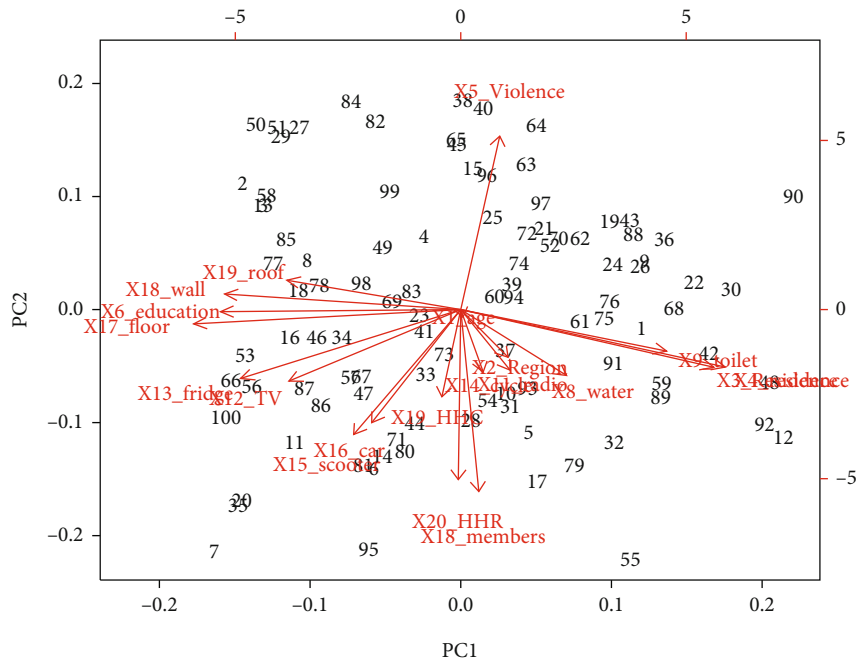


FIGURE 2: A biplot visualizing the correlations between the covariates on the first two principal components for infant survival data set.

data having known correlation structure. Similar results based on BIC, as shown in Figure 1(b), are observed. The simulation analysis demonstrated that the proposed models are efficient and reliable in terms of performance for the corresponding distributions. The analysis over simulation recommended the practical application of proposed models to examine sur-

vival response along with correlated covariates in a more flexible manner.

Before analyzing the real data set, multicollinearity among covariates is verified to justify the application of PLS. For this purpose, correlations structure for infant survival data is examined. The biplot for infant survival data

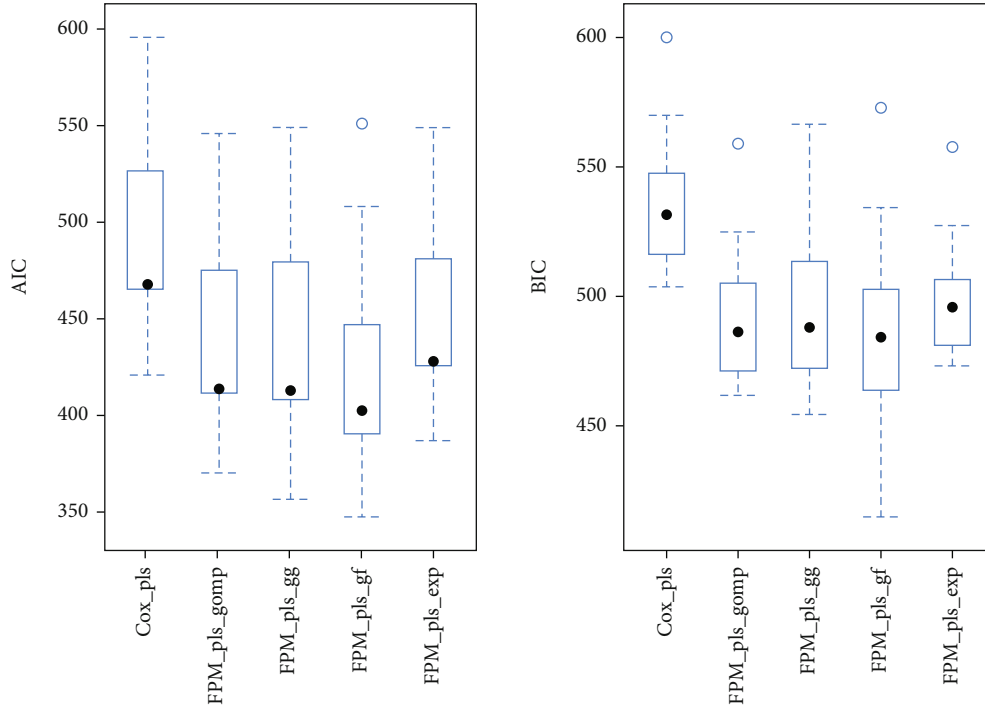


FIGURE 3: The comparison of reference model with the PLS-FPM parameterized over generalized gamma (GG), generalized F (GF), exponential, and Gompertz distribution on the basis of AIC and BIC for infant survival are presented.

presented in Figure 2 clearly portrayed the correlation between covariates showing close points of occurrence.

Real data set of infant survival with 12 months of censoring is considered in this analysis. Discarding outliers, 83 covariates measured over 577 observations (infants) were included in the final sample to compare survival models. The data set is randomly split into testing (30%) and training data (70%) for reliable results. After verification of multicollinearity among covariates, the PLS-FPM parameterized over Gompertz, generalized gamma, generalized F, and exponential distribution are analyzed. The PLS-Cox model for survival time is considered as the reference method. Figure 3 showed the efficiency of models measured by AIC and BIC which demonstrated the higher performance of modified models compared to the PLS-Cox over infant survival data. The proposed models based on the parametric approach performed better due to their additional flexibility. Flexible parametric models integrated with PLSR parameterized with generalized gamma (GG), generalized F (GF), exponential, and Gompertz distribution showed increased accuracy compared to the Cox model integrated with PLS.

The Gompertz distribution is modeled into the innovation-imitation paradigm, and its hazard function works as a convex function. These properties develop their flexibility to use as flexible parametric distribution in survival models. Hence, it increased the performance of the model incorporated with PLS compared to the semiparametric model, due to its flexible nature. Based on AIC and BIC, it is concluded that the PLS-FPM parameterized over generalized F (GF) is the best-fitted model and hence further executed for influential factor selection. PLS-FP model based on generalized F-distribution with location parameter μ is found to be the most

TABLE 1: A description of corresponding parameter of each distribution used in the PLS-FPM for infant survival data set is presented.

Model	Parameter			
	Location	Scale	Shape	Rate
<i>FPM_pls_gomp</i>	—	—	0.231	-7.20
<i>FPM_pls_gg</i>	4.32	-0.32	0.56	—
<i>FPM_pls_gf</i>	4.37	-1.44	-0.11, 2.96	—
<i>FPM_pls_g</i>	—	—	0.77	-3.56

efficient model over infant survival times data. In this model, covariates on the corresponding parameter represent the accelerated failure time (AFT) model which speeds up or slows down the passage of time. A detailed illustration of PLS-FP model parameterization is presented in Table 1 to describe the corresponding location, scale, shape, and rate parameter of the associated distribution.

Figure 4 showed the cumulative hazards regression estimates for the reference method and the PLS-FPM integrated with generalized gamma (GG), generalized F (GF), exponential, and Gompertz distribution for infant mortality data. The proposed PLS-FPM delivered smooth regression coefficients of the hazard functions extrapolated to a time of 12 months showing consistent estimates. The reference model showed unsmooth hazard trends with odd fluctuations for certain time intervals shown in Figure 4.

For modeling the survival time data, the PLS-FPM parameterized over generalized F (GF) is applied, and a well-known factor selection method of PLS, namely, loading weights, is used to estimate the regression coefficients of

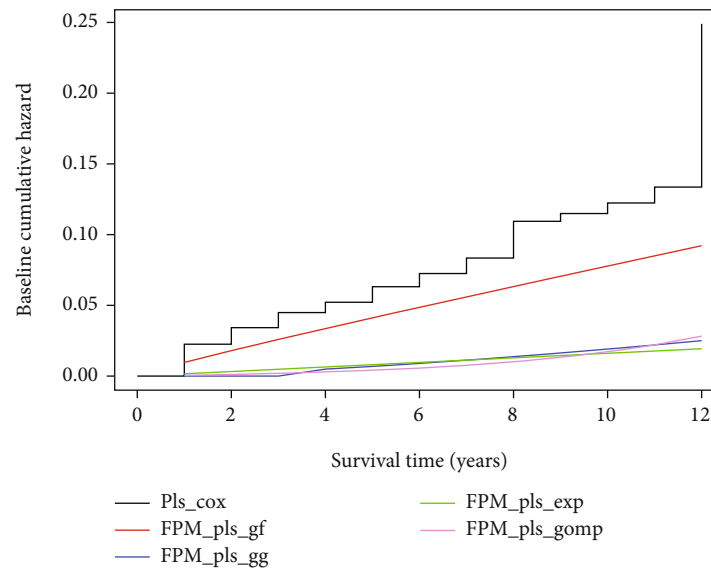


FIGURE 4: The cumulative hazard estimates of the PLS-Cox and the PLS-FPM parameterized over generalized gamma (GG), generalized F (GF), exponential, and Gompertz distribution for infant survival data.

TABLE 2: The coefficient estimates of influential factors for infant survival obtained by the PLS-FPM coupled with GF-distribution.

Factor	Coefficient
Mother's age	-0.2599634
Province	0.1152137
Type of place of residence	-0.2259600
Selected for domestic violence module	-0.1314814
Mother's educational level	0.1150733
Type of toilet facility	0.1306607
Household has: television	0.1123271
Main roof material	-0.1824950
Relationship to household head	-0.1462302
Sex of household head	0.1263040
Toilet facilities shared with other households	0.1128781
Wealth index	-0.2221866
Total children ever born	0.1440127
Sons died	0.1444297
Daughters died	0.1380171
Used contraceptive methods	0.1255910
Have mosquito bed net for sleeping	-0.1472814
Getting medical help for self: getting permission to go	0.2765390
Getting medical help for self: getting money needed for treatment	0.1091858
Getting medical help for self: distance to health facility	0.2711795
Getting medical help for self: having to take transport	0.1405939
Getting medical help for self: not wanting to go alone	0.1932727
Heard of tuberculosis or TB	-0.2165038
Person who usually decides on visits to family or relatives	-0.1273039
Preceding birth interval (in months)	-0.1647803
Duration of breastfeeding	-0.1697750
Blood relation with husband	-0.1285042
Total pregnancy outcomes	-0.1926369

significant factors. The estimates of important predictors associated with infant mortality are presented in Table 2.

After analysis, 28 influential factors out of 80 which significantly affect infant survival in Pakistan are observed. A negative relationship of mother's age, region, selection for domestic violence, main roof material, relationship to household head, wealth index, availability of mosquito bed net, awareness about tuberculosis (TB), decision power to visit family, preceding birth interval, duration of breastfeeding, blood relation with husband, and total pregnancy outcomes are found for infant survival. Furthermore, positive association of province, mother's education, toilet facility, availability of television, sex of household head, shared toilet, number of total children, number of dead son and daughters, use of contraception, availability of permission, money, transport, and attendant for medical facility and distance to a medical facility was observed.

4. Discussion

Estimating the hazard and survival functions that flexibly explain complex systems remained a hard and computationally challenging task. Hence, the candidate models are usually limited in studies to allow for evaluations and comparisons. However, nonparametric and semiparametric survival methods can speculate model structures as unsmooth estimates are evaluated. The present study extended the PLS-FPM [14] to correlated predictors having moderate trend observations using four alternative probability distributions. The PLS-FPM extends previous survival approaches that either perform semiparametric analyses or use nonparametric methods, while analysis of all previous methods was limited due to their inflexible nature. To administrate all four shaped hazard functions, distribution fitting is implemented over defined simulated survival data set.

Most previous literature used the Cox regression model for infant survival analysis [20]. Very few recent studies used FPM to examine infant survival analysis [21]. The PLS-FPM is compared with the reference method for both simulated and a real data set for collinear covariates. A previous study proposed the PLS-FPM integrated with Gamma, Weibull, log-logistic, and log-normal distributions for data with extreme observations to examine four real data sets of breast cancer survival time and identify the significantly associated gene signatures for each data set. The study found that the PLS-FPM has higher performance than the traditional PLS-Cox model [14]. Consistent with the previous study, the present study found the higher efficiency of the PLS-FPM compared to the PLS-Cox regression method for data sets with moderate observations. The PLS-FPM coupled with Gompertz distribution is found to be the optimum model to estimate hazard functions using AIC for simulated survival data following Gompertz distribution. The efficiency of the algorithms flexibly increases the model accuracy to a greater extent even considering correlated predictors. This accuracy suggested that hazard, as well as survival functions, can be accurately computed by smooth trends for the survival response. A recent study proposed the partial least squares spline modeling approach by inte-

grating PLS with restricted cubic spline model and compared it with the PLS-Cox model [22]. The study estimated the risk factors of infant mortality in Pakistan by using the PLS-spline model based on the odds scale with one knot. This study also examine the important factors of infant mortality by executing the optimal model, namely, the PLS-FPM parameterized over generalized F (GF), and identified the influential factors which are also determined by various previous studies. Consistent with the recent literature, the present study evidenced that mother's age, region, selection for domestic violence, relationship to household head, wealth index, awareness about tuberculosis (TB), decision power to visit family, preceding birth interval, and blood relation with husband [22] are significantly associated with infant mortality. Some other previous literature also supported the association of main roof material [23, 24], availability of mosquito bed net [25], duration of breastfeeding [26], and total pregnancy outcomes [27] with infant survival similar to the present study.

Various previous studies also observed the positive association of province [28], mother's educational level [29], type of toilet facility [30], availability of television [31], sex of household head [32], shared toilet [33], number of total children [34], number of died son and daughters [35], use of contraception [36], and availability of permission, money, transport, distance and attendant for medical facility [37, 38] with infant survival which is consistent with the current study. Last but not least, the PLS-FPM not only can extrapolate survival response besides the availability of follow-up information but also sponsors variant hazard shapes. The PLS-FPM is suggested as a helpful parametric addition for the estimation and prediction of survival response. This model is recommended to use in reliability theory for risk assessment.

Data Availability

Data is freely available at <https://dhsprogram.com/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] E. T. Lee and O. T. Go, "Survival analysis in public health research," *Annual review of public health.*, vol. 18, no. 1, pp. 105–134, 1997.
- [2] E. A. Gehan, "A generalized Wilcoxon test for comparing arbitrarily singly-censored samples," *Biometrika*, vol. 52, no. 1-2, pp. 203–224, 1965.
- [3] R. Peto and J. Peto, "Asymptotically efficient rank invariant test procedures," *Journal of the Royal Statistical Society: Series A (General)*, vol. 135, no. 2, pp. 185–198, 1972.
- [4] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 34, no. 2, pp. 187–202, 1972.
- [5] J. P. Klein and M. L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*, Springer Science & Business Media, 2006.

- [6] P. Bastien, V. E. Vinzi, and M. Tenenhaus, "PLS generalised linear regression," *Computational Statistics & data analysis.*, vol. 48, no. 1, pp. 17–46, 2005.
- [7] P. Royston and P. C. Lambert, *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model*, vol. 347, Stata press College Station, TX, 2011.
- [8] J. Bryce, R. E. Black, and C. G. Victora, "Millennium development goals 4 and 5: progress and challenges," *BMC medicine.*, vol. 11, no. 1, pp. 1–4, 2013.
- [9] R. E. Black, C. Levin, N. Walker et al., "Reproductive, maternal, newborn, and child health: key messages from _Disease Control Priorities 3rd Edition_," *The Lancet.*, vol. 388, no. 10061, pp. 2811–2824, 2016.
- [10] W. A. Rosa, *New Era in Global Health: Nursing and the United Nations 2030 Agenda for Sustainable Development*, Springer Publishing Company, 2017.
- [11] A. Helova, K. R. Hearld, and H. Budhwani, "Associates of neonatal, infant and child mortality in the Islamic Republic of Pakistan: a multilevel analysis using the 2012–2013 demographic and health surveys," *Maternal and child health journal.*, vol. 21, no. 2, pp. 367–375, 2017.
- [12] W. H. Organization, *World Health Statistics 2016: Monitoring Health for the SDGs Sustainable Development Goals*, World Health Organization, 2016.
- [13] National Institute of Population Studies (Pakistan), Macro International, Institute for Resource Development, and Demographic, & Health Surveys, *Pakistan Demographic and Health Survey*, National Institute of Population Studies, 2012.
- [14] M. Sadiq and T. Mehmood, "A flexible and robust approach to analyze survival systems in the presence of extreme observations," *Mathematical Problems in Engineering.*, vol. 2021, pp. 1–11, 2021.
- [15] C. Cox and M. Matheson, "A comparison of the generalized gamma and exponentiated Weibull distributions," *Statistics in medicine.*, vol. 33, no. 21, pp. 3772–3780, 2014.
- [16] M. Sadiq, T. Mehmood, and M. Aslam, "Identifying the factors associated with cesarean section modeled with categorical correlation coefficients in partial least squares," *PLoS One*, vol. 14, no. 7, article e0219427, 2019.
- [17] T. Mehmood, M. Sadiq, and M. Aslam, "Filter-based factor selection methods in partial least squares regression," *IEEE Access.*, vol. 7, pp. 153499–153508, 2019.
- [18] S. L. Brilleman, R. Wolfe, M. Moreno-Betancur, and M. J. Crowther, "Simulating survival data using the simsurvR package," *Journal of Statistical Software.*, vol. 97, no. 3, pp. 1–27, 2021.
- [19] P. Demographic, "Health Survey 2012–13," in *Islamabad and Calverton, MA: National Institute of Population Studies and ICF International; 2013, 2015*, <https://dhsprogram.com/data>.
- [20] A. K. Iddrisu, A. Alhassan, and N. Amidu, "Survival analysis of birth defect infants and children with pneumonia mortality in Ghana," *Public Health*, vol. 2019, pp. 1–7, 2019.
- [21] R. K. Saroj, K. H. H. V. S. S. N. Murthy, M. Kumar, R. Singh, and A. Kumar, "Survival parametric models to estimate the factors of under-five child mortality," *Journal of Health Research and Reviews*, vol. 6, no. 2, 2019.
- [22] M. Sadiq, D. K. F. Alnagar, A. T. Abdulrahman, and R. Alharbi, "The partial least squares spline model for public health surveillance data," *Computational and Mathematical Methods in Medicine.*, vol. 2022, pp. 1–7, 2022.
- [23] S. A. Adebowale, O. M. Morakinyo, and G. R. Ana, "Housing materials as predictors of under-five mortality in Nigeria: evidence from 2013 demographic and health survey," *BMC pediatrics.*, vol. 17, no. 1, p. 30, 2017.
- [24] E. Bendavid, "Changes in child mortality over time across the wealth gradient in less-developed countries," *Pediatrics*, vol. 134, no. 6, pp. e1551–e1559, 2014.
- [25] D. Meekers and J. O. Yukich, "The association between household bed net ownership and all-cause child mortality in Madagascar," *Malaria journal.*, vol. 15, no. 1, p. 475, 2016.
- [26] D. Phukan, M. Ranjan, and L. Dwivedi, "Impact of timing of breastfeeding initiation on neonatal mortality in India," *International breastfeeding journal*, vol. 13, no. 1, 2018.
- [27] S. W. Masho and P. W. Archer, "Does maternal birth outcome differentially influence the occurrence of infant death among African Americans and European Americans?," *Maternal and child health journal.*, vol. 15, no. 8, pp. 1249–1256, 2011.
- [28] S. A. Collins, P. Surmala, G. Osborne et al., "Causes and risk factors for infant mortality in Nunavut, Canada 1999–2011," *BMC pediatrics*, vol. 12, no. 1, 2012.
- [29] G. T. Kiross, C. Chojenta, D. Barker, T. Y. Tiruye, and D. Loxton, "The effect of maternal education on infant mortality in Ethiopia: a systematic review and meta-analysis," *PLoS One*, vol. 14, no. 7, p. e0220076, 2019.
- [30] O. Ezeh, K. Agho, M. Dibley, J. Hall, and A. Page, "The impact of water and sanitation on childhood mortality in Nigeria: evidence from demographic and health surveys, 2003–2013," *International journal of environmental research and public health.*, vol. 11, no. 9, pp. 9256–9272, 2014.
- [31] T. Dendup, Y. Zhao, and D. Dema, "Factors associated with under-five mortality in Bhutan: an analysis of the Bhutan National Health Survey 2012," *BMC Public Health*, vol. 18, no. 1, p. 1375, 2018.
- [32] R. Adhikari and C. Podhisita, "Household headship and child death: evidence from Nepal," *BMC international health and human rights*, vol. 10, no. 1, 2010.
- [33] J. Golding, R. Greenwood, A. McCaw-Binns, and P. Thomas, "Associations between social and environmental factors and perinatal mortality in Jamaica," *Paediatric and perinatal epidemiology.*, vol. 8, no. s1, pp. 17–39, 1994.
- [34] A. van Soest and U. R. Saha, "Relationships between infant mortality, birth spacing and fertility in Matlab, Bangladesh," *PLoS One*, vol. 13, no. 4, article e0195940, 2018.
- [35] B. B. Gubhaju, "The effect of previous child death on infant and child mortality in rural Nepal," *Studies in Family Planning.*, vol. 16, no. 4, pp. 231–236, 1985.
- [36] U. R. Saha and A. van Soest, "Contraceptive use, birth spacing, and child survival in Matlab, Bangladesh," *Bangladesh. Studies in family planning*, vol. 44, no. 1, pp. 45–66, 2013.
- [37] S. Adebowale and E. Udjo, "Maternal health care services access index and infant survival in Nigeria," *Ethiopian journal of health sciences.*, vol. 26, no. 2, pp. 131–146, 2016.
- [38] D. Kadobera, B. Sartorius, H. Masanja, A. Mathew, and P. Waiswa, "The effect of distance to formal health facility on childhood mortality in rural Tanzania, 2005–2007," *Global health action.*, vol. 5, no. 1, p. 19099, 2012.