A Hybrid Neural Network Framework for Zero-Inflated Spatial Survival Analysis

Sepideh Asadi, Mohsen Mohammadzadeh

Department of Statistics, Tarbiat Modares University, Tehran, Iran.

Abstract

In this talk, we address limitations in traditional survival models when applied to data with zero-inflation and spatial dependence. While the zero-inflated discrete Weibull model previously handled these issues, its parametric nature can struggle with complex, nonlinear relationships. To overcome this, we propose a hybrid neural network framework that enhances the zero-inflated discrete Weibull model. The new framework uses a neural network to predict zero-inflation probabilities and graph neural networks to capture spatial correlations. An additional neural network is integrated to predict survival times with right censoring flexibly. This end-to-end training approach simultaneously learns complex patterns, nonlinear relationships, and spatial dependencies. Our framework improves predictive accuracy and interpretability in survival analysis with zero-inflated and spatially correlated data, as we will demonstrate through simulations and real applications.

Keywords: Survival data, Zero-Inflated, Neural network, Hybrid model.

Mathematics Subject Classification : 62H11**,** 62M30**,** 62F15**.**

1 Introduction

Survival analysis models the time until an event occurs. Traditional models, such as the Cox or Weibull, assume that the event cannot occur at time zero. This assumption is often unrealistic, as zero survival times or immediate failures are common, particularly in discrete data where events may occur before

[0]**Corespondin author:** Sepideh Asadi, `s.asadimit48@gmail.com`

the first measurement interval. When the number of these zero survival times exceeds expectations, the data is considered zero-inflated. Ignoring zero-inflation can lead to biased results, such as overestimating survival probabilities. To address this, specialized zero-inflated survival models have been developed, typically employing a two-part structure: one part models the probability of immediate failure (the ”zero” component), while the other models the survival distribution for non-immediate failures. Survival data collected from different locations often exhibit spatial dependence, where nearby areas show similar outcomes. Traditional survival models, which assume independence, fail to capture this spatial correlation, leading to inaccurate results. Bayesian models often use Conditional Autoregressive (CAR) priors to account for latent spatial heterogeneity. However, the challenge intensifies when both spatial dependence and zero-inflation are present. While some zero-inflated spatial models exist, they primarily focus on count data rather than right-censored time-to-event data common in survival analysis. For instance, Liu et al. (2020) proposed Zero-Inflated Spatial Frailty Models, which incorporate spatial effects in frailty-based survival models but are limited to frailty structures and do not explicitly address complex nonlinear relationships in the presence of zero-inflation. Similarly, Asadi and Mohammadzadeh (2025) introduced a Bayesian spatial censored zero-inflated survival model based on the Censored Zero-Inflated Discrete Weibull (CZIDW) distribution. However, such classical models often struggle with complex nonlinear relationships and are prone to model misspecification. Deep learning provides a promising alternative for addressing these limitations. Models like DeepSurv (Katzman et al., 2018), DeepHit (Lee et al., 2018), and Nnet-survival (Gensheimer and Culbreth, 2019) leverage neural networks to capture complex, nonlinear relationships in survival data, offering superior accuracy and robustness, especially with censored or incomplete data. These models are also effective at handling zero-inflated phenomena. To model spatial dependence, Graph Neural Networks (GNNs), such as Graph Convolutional Networks (GCNs) and Graph Attention Networks (GATs), have emerged as powerful tools. Recent work, such as GNN-Surv (Hu et al., 2021), integrates GNNs with survival analysis to improve predictive accuracy in spatially-indexed survival data. Similarly, Wang et al. (2022) proposed Spatial Deep Survival Analysis, which incorporates spatial information into deep learning frameworks for survival analysis. However, their approach does not explicitly address zero-inflation, limiting its applicability in datasets with immediate failures. This highlights the need for models that simultaneously address zero-inflation and spatial dependence in survival analysis, which we introduce in the following paragraph. To address these gaps, we propose a Hybrid Zero-Inflated Spatial Survival Neural Network (ZIS-SurvNet). This model features two interconnected components: a zero-inflation module using deep neural networks to identify immediate failures and a deep survival module, inspired by DeepSurv and DeepHit, to predict survival times for non-zero cases. Spatial dependence is incorporated using GNNs, which learn spatial embeddings from neighboring nodes and integrate them into both modules. Unlike Liu et al. (2020), which focuses on frailty-based spatial modeling, and Wang et al. (2022), which does not address zero-inflation, our model simultaneously handles zero-inflation, spatial correlation, and censoring in a unified deep learning framework. The model is trained end-to-end using a joint loss function, enabling simultaneous learning of both components. Our approach demonstrates superior predictive accuracy and spatial interpretability compared to traditional methods, as validated on a real-world dataset. The rest of this paper is structured as follows: Section 2 reviews CZIDW models. Section 3 applies our framework to a real-world dataset. Finally, Section 4 provides a summary and discusses future research directions.

2 Methodology

This section presents ZIS-SurvNet, a deep learning framework for discrete-time survival data with zero inflation, spatial dependence, and right censoring. It integrates the Zero-Inflated Discrete Weibull (ZIDW) model with Graph Neural Networks (GNNs) and deep survival analysis to address nonlinear relationships and spatial autocorrelation. The dataset comprises $n = \sum_{r=1}^{R} n_r$ individuals across R spatial units, with n_r individuals in unit r. Units form an undirected graph $G = (V, E)$, where vertices $v_r \in V$ represent units and edges E denote spatial proximity. Each individual i in region r has discrete survival time $T_{ir} \in \{0, 1, \dots, T_{\max}\}$, observed as $Y_{ir} = \min(T_{ir}, C_{ir})$ under right censoring (C_{ir} independent of T_{ir}). Censoring indicator: $\delta_{ir} = 1(T_{ir} \leq C_{ir})$. Covariates: $X_{ir} \in \mathbb{R}^p$. Zero-inflation indicator: $J_{ir} = 1(Y_{ir} = 0)$. Data cases (assuming no zero-censoring; handled in likelihood if present):

$$\begin{cases} J_{ir} = 1, \delta_{ir} = 1 & Y_{ir} = 0, \text{uncensored}, \\ J_{ir} = 0, \delta_{ir} = 1 & Y_{ir} > 0, \text{uncensored}, \\ J_{ir} = 0, \delta_{ir} = 0 & Y_{ir} > 0, \text{censored}. \end{cases}$$

ZIDW models excess zeros via mixture:

$$P(T_{ir} = t) = \begin{cases} \pi_{ir} + (1 - \pi_{ir}) f_1(0; q_{ir}, \beta), & t = 0, \\ (1 - \pi_{ir}) f_1(t; q_{ir}, \beta), & t \geq 1, \end{cases}$$

where $\pi_{ir} = P(J_{ir} = 1)$ and $f_1(t; q_{ir}, \beta) = q_{ir}^{t^\beta} - q_{ir}^{(t+1)^\beta}$ ($q_{ir} \in (0, 1)$, $\beta > 0$). The goals of the model are threefold: to capture spatial dependencies through the use of graphs; to model the probability of zero inflation by estimating $P(J_{ir} = 1)$; and to estimate $P(T_{ir} = t \mid J_{ir} = 0, X_{ir}, r)$ under censoring. Areal units form a graph $G = (V, E)$ with a weighted adjacency matrix $A \in \mathbb{R}^{R \times R}$ based on shortest-path distances. We use a Graph Convolutional Network (GCN) (Kipf and Welling, 2017) to learn spatial embeddings $S_r \in \mathbb{R}^d$ from the graph structure:

$$H^{(\ell+1)} = \text{ReLU}\left(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^{(\ell)} W^{(\ell)}\right)$$

where $\tilde{A} = A + I$, $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$, $H^{(0)}$ initial features, $W^{(\ell)}$ weights, and two layers used d tunable. Estimates $\pi_{ir} = P(J_{ir} = 1 \mid X_{ir}, S_r)$ via MLP on $[X_{ir}, S_r] \in \mathbb{R}^{p+d}$: input layer, hidden layers (ReLU, dropout 0.3 (Srivastava et al., 2014)), output sigmoid σ: $\pi_{ir} = \sigma(\mathrm{MLP}_{\mathrm{ZI}}([X_{ir}, S_r]; \theta_{\mathrm{ZI}}))$.

For non-zeros: $P(T_{ir} = t \mid J_{ir} = 0, X_{ir}, S_r) = f_1(t; q_{ir}, \beta)$, $t \geq 1$. Model q_{ir} via log-log link: $q_{ir} = \exp(-\exp(f_{\mathrm{DW}}(X_{ir}, S_r)))$, f_{DW} MLP on $[X_{ir}, S_r]$ (ReLU, dropout 0.3). β global learnable parameter. End-to-end training minimizes negative log-likelihood from ZIDW, handling uncensored ($t = 0$, $t > 0$) and censored cases:

$$\mathcal{R} = -\frac{1}{n}\sum_{r=1}^{R}\sum_{i=1}^{n_r}\Big\{\delta_{ir}\Big[J_{ir}\log\left(\pi_{ir} + (1-\pi_{ir})(1-q_{ir})\right) + (1-J_{ir})\log\Big((1-\pi_{ir})(q_{ir}^{Y_{ir}^{\beta}} - q_{ir}^{(Y_{ir}+1)^{\beta}})\Big)\Big]$$
$$+ (1-\delta_{ir})\log\Big((1-\pi_{ir})q_{ir}^{(Y_{ir}+1)^{\beta}}\Big)\Big\}. \quad (1)$$

ZIS-SurvNet is trained end-to-end using the Adam optimizer (Kingma and Ba, 2015) (learning rate 10^{-3}). The process involves normalizing covariates X_{ir}, building graph G and adjacency A (e.g., $h = 1.0$), initializing GCN with Glorot (Glorot and Bengio, 2010) and MLPs with He (He et al., 2015) ($\beta = 1.0$). Mini-batches ($B = 64$) compute S_r, π_{ir}, q_{ir}, and loss $\mathcal{R}$. Backpropagation updates parameters with L2 regularization (10^{-4}) and dropout (0.3). Early stopping monitors validation loss (20% data, 20-epoch patience). Training runs up to 200 epochs, saving the best model. Implemented in PyTorch (Paszke et al., 2019) with PyTorch Geometric (Fey and Lenssen, 2019) for GCN. GCN has two layers ($d_0 = 16$, $d_1 = 32$); MLPs have three hidden layers (128, 64, 32), ReLU, dropout 0.3. Bandwidth $h \in \{0.5, 1.0, 2.0\}$ is tuned; $\beta = \log(1 + e^{\beta_{\mathrm{raw}}})$. The code is modular and available on request. The training procedure for ZIS-SurvNet is implemented as an end-to-end optimization framework that jointly learns zero-inflation probabilities, spatial embeddings, and survival time distributions under censoring. Algorithm 1 outlines the complete training pipeline. Given the input dataset, spatial graph, and model hyperparameters, the algorithm first normalizes covariates and constructs the spatial adjacency matrix. It then initializes the Graph Convolutional Network (GCN) and Multi-Layer Perceptron (MLP) weights. For each epoch, mini-batches are processed to compute spatial embeddings, predict zero-inflation and survival parameters, and calculate the negative log-likelihood loss. The model parameters are updated using the Adam optimizer, and early stopping is applied based on validation loss to prevent overfitting. The final output consists of the trained parameters used for inference and evaluation.

The dataset is split into 60% training, 20% validation, and 20% test sets. Performance is evaluated on the test set using Negative Log-Likelihood (NLL), Concordance index (C-index) (Harrell et al., 1982) for ranking non-zero censored times, and MAE for expected vs. observed non-zero uncensored times. We note that traditional information criteria like AIC and BIC, while useful for classical parametric models, are not directly applicable to complex deep learning models due to the difficulty in defining the number

Algorithm 1 Training ZIS-SurvNet: Hybrid Neural Network for Zero-Inflated Spatial Survival Analysis

Require: Dataset $\{(X_{ir}, Y_{ir}, \delta_{ir}, J_{ir})\}_{i=1,r=1}^{n,R}$, graph $G = (V, E)$ with adjacency matrix A, hyperparameters (learning rate η, batch size B, max epochs $E_{\max}$, dropout rate, regularization, bandwidth h)

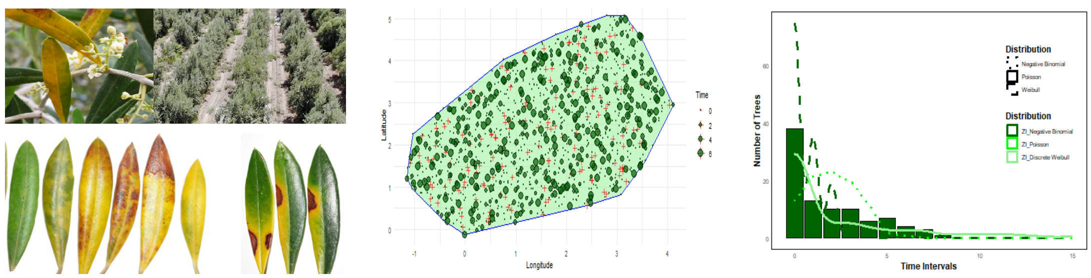
Ensure: Trained model parameters $\theta_{\mathrm{ZI}}, \theta_{\mathrm{DW}}, \beta$

Normalize covariates X_{ir}
Construct weighted adjacency matrix A using spatial distances and bandwidth h
Initialize GCN and MLP weights (Glorot, He initialization)
Initialize global Weibull parameter β
for epoch = 1 to $E_{\max}$ **do**
  Shuffle dataset and partition into mini-batches of size B
  for each batch **do**
    Compute spatial embeddings S_r for regions r using GCN on adjacency A
    Concatenate covariates and embeddings: $Z_{ir} = [X_{ir}, S_r]$
    Predict zero-inflation probabilities: $\pi_{ir} = \sigma(\mathrm{MLP}_{\mathrm{ZI}}(Z_{ir}; \theta_{\mathrm{ZI}}))$
    Predict discrete Weibull parameter: $q_{ir} = \exp(-\exp(\mathrm{MLP}_{\mathrm{DW}}(Z_{ir}; \theta_{\mathrm{DW}})))$
    Compute negative log-likelihood loss $\mathcal{R}$ using Equation (1)
    Backpropagate and update parameters $(\theta_{\mathrm{ZI}}, \theta_{\mathrm{DW}}, \beta)$ using Adam optimizer with learning rate η
  end for
  Evaluate validation loss and apply early stopping if no improvement for patience epochs
end for
return Trained parameters $(\theta_{\mathrm{ZI}}, \theta_{\mathrm{DW}}, \beta)$

of effective parameters. Instead, we use a combination of performance metrics on a held-out test set and regularization techniques like $L^2$and dropout, alongside early stopping, to ensure the model finds an optimal balance between fit and complexity, thereby preventing overfitting. Sensitivity analyses test GCN layer and dimension variations, binary vs. weighted adjacency with varying h, and robustness to 10-20% covariate missingness using mean, kNN, or embedding-based imputation. Metrics are reported with 95% CI via bootstrap (Efron and Tibshirani, 1994), confirming model robustness.

3 Application to Olive Orchard Data

Cercosporiosis, a fungal disease causing leaf discoloration and reduced olive yield, was studied in a 5,000-square-meter orchard in western Iran with 2,200 saplings monitored over two months. Time-to-infection, discretized into nine 1-week intervals ($T_i \in \{0, \dots, 8\}$, $T_i = 0$ for infections within the first week), shows 20% structural zeros (early infections) and 20.6% right censoring (454/2,200 censored). A histogram of infection times indicates zero inflation and overdispersion (dispersion index 2.81), justifying the ZIDW model. Covariates include tree age (mean 4.50 years, SD 0.83), olive species (Type I: 23%, II: 46%, III: 29%), and tree height (mean 3.90 meters, SD 0.84). Spatial autocorrelation, driven by spore dispersal via precipitation and wind, increases infection risk near infected trees. The orchard is partitioned into 89 hexagons, forming an undirected graph $G = (V, E)$ with adjacency matrix $A \in \mathbb{R}^{89 \times 89}$ based on shared boundaries ($A_{ij} = 1$ if adjacent, else 0). Hexagon features aggregate tree covariates (mean age, species proportions, mean height, infection rate). Missing covariates (<5%) are imputed using the median for numerical and the mode for categorical variables.



(a) (b) (c)

Figure 1: a: Cercosporiosis symptoms in an olive orchard, b: Location and time of infection of trees in the olive orchard, observations ($\bullet$) and censored data (+), and c: Histogram and Zero-inflated distributions of "time-to-infestation" among the olive trees.

ZIS-SurvNet, implemented as in Section 2, uses a GCN (2 layers, 64 units) to generate hexagon embeddings $S_i \in \mathbb{R}^{64}$, an MLP (3 hidden layers: 128, 64, 32, ReLU, dropout 0.3) to predict zero-inflation probability $\pi_i = P(J_i = 1 \mid X_i, S_i)$, and another MLP (4 layers: 256, 128, 64, 32, ReLU, dropout 0.3) to estimate ZIDW parameters q_i, β for non-zero survival times. The model is trained on 70% (1,540 trees), validated on 15% (330 trees), and tested on 15% (330 trees) using Adam (learning rate 0.001, batch size 64, early stopping with 20-epoch patience) to minimize the negative log-likelihood (Equation 1). Performance is evaluated using negative log-likelihood (NLL, overall fit), C-index, integrated Brier score (IBS, calibration), AUC-ROC (zero-inflation discrimination), and MAE (expected vs. observed non-zero times). Table 1 shows ZIS-SurvNet outperforms CZIDW and DeepSurv, achieving NLL 1.92 (vs. 2.15, 2.20), C-index 0.84 (vs. 0.79, 0.77), IBS 0.12 (vs. 0.15, 0.16), AUC-ROC 0.87 (vs. 0.81, –), and MAE 0.08 (vs. 0.11, –) (p < 0.01). Spatial embeddings confirm high-risk clusters in northern and eastern regions, driven by Type II species and older trees, demonstrating ZIS-SurvNet's effectiveness in capturing spatial and covariate-driven infection dynamics.

Table 1: Performance of ZIS-SurvNet vs. baselines (standard errors in parentheses).

Model	**NLL**	**C-index**	**IBS**	**AUC-ROC**	**MAE**
CZIDW	2.15 (0.03)	0.79 (0.03)	0.15 (0.02)	0.81 (0.03)	0.11 (0.02)
DeepSurv	2.20 (0.03)	0.77 (0.03)	0.16 (0.02)	–	–
ZIS-SurvNet	**1.92 (0.02)**	**0.84 (0.02)**	**0.12 (0.01)**	**0.87 (0.02)**	**0.08 (0.01)**

4 Discussion and Results

Our study successfully applied ZIS-SurvNet to olive orchard data, demonstrating its superior ability to model survival times. Unlike conventional methods like DeepSurv and CZIDW, ZIS-SurvNet effectively handles zero inflation, spatial dependence, and censoring by integrating deep neural networks with graph convolutional networks. The model's strong performance, evidenced by high C-index, AUC-ROC, and low IBS and MAE scores, confirms its potential for epidemiological applications. We identified high-risk clusters associated with factors like tree age and species, providing valuable insights for targeted interventions like pruning or localized treatments. While ZIS-SurvNet's computational complexity and reliance on proper spatial aggregation are noted limitations, its robust performance confirms its applicability for spatial survival analysis with zero inflation. Future work will explore incorporating temporal dynamics and alternative spatial graph structures.

Acknowledgement

The authors thank Dr Kiomars Motargem for his help to access to the olive orchard data. Receiving support from the Center of Excellence in Analysis of Spatio-Temporal Correlated Data at Tarbiat Modares University is acknowledged.

References

Asadi, S. and Mohammadzadeh, M. (2025), Bayesian Analysis of Spatial Zero-Inflated and Right-Censored Survival Data, *Journal of Agricultural, Biological and Environmental Statistics*, 1-38.

Efron, B. and Tibshirani, R. J. (1994), *An Introduction to the Bootstrap*, CRC Press.

Fey, M. and Lenssen, J. E. (2019), Fast Graph Representation Learning with PyTorch Geometric, *arXiv Preprint arXiv:1903.02428*.

Gensheimer, M. F. and Narasimhan, B. (2019), A Scalable Discrete-time Survival Model for Neural Networks, *PeerJ Computer Science*, 5:e146.

Glorot, X. and Bengio, Y. (2010), Understanding the Difficulty of Training Deep Feedforward Neural Networks, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, **9**, 249–256.

Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L. and Rosati, R. A. (1982), Evaluating the Yield of Medical Tests, *JAMA*, **247**(18), 2543-2546.

He, K., Zhang, X., Ren, S. and Sun, J. (2015), Delving Deep Into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, *Proceedings of the IEEE International Conference on Computer Vision*, 1026–1034.

Hu, G., et al. (2023), GNN-Surv: Graph Neural Network for Survival Analysis with Spatially Dependent Data, *Biometrics*.

Katzman, J. L., Shaham, U., Cloninger, A., Bates, J., Jiang, T. and Kluger, Y. (2018), DeepSurv: Personalized Treatment Recommender System Using a Cox Proportional Hazards Deep Neural Network, *BMC Medical Research Methodology*, **18**(24).

Kingma, D. P. and Ba, J. (2015), Adam: A Method for Stochastic Optimization, *ICLR*.

Kipf, T. N. and Welling, M. (2017), Semi-Supervised Classification with Graph Convolutional Networks, *International Conference on Learning Representations*.

Lee, C., Zame, W. R., Yoon, J. and van der Schaar, M. (2018), DeepHit: A Deep Learning Approach to Survival Analysis with Competing Risks, *AAAI Conference on Artificial Intelligence*, **32**(1).

Liu, C., Wang, T. and Zhang, H. (2020), A Zero-Inflated Spatial Frailty Model for Disease Mapping with Excess Zeros, *Journal of the American Statistical Association*.

Paszke, A. et al. (2019), PyTorch: An Imperative Style, High-Performance Deep Learning Library, *NeurIPS*.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014), Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *Journal of Machine Learning Research*, **15**(1), 1929–1958.

Wang, M., Hu, G. and Liu, H. (2022), Spatial Deep Survival Analysis with Application to Cancer Epidemiology, *Statistics in Medicine*, **41**(12), 2099-2115.