

# A Spatial-Nonparametric Approach for Prediction of Claim Frequency in Motor Insurance

Gideon Kipngetich<sup>1</sup>, Ananda Kube<sup>2</sup>, Thomas Mageto<sup>3</sup>

<sup>1</sup>Pan African University, Nairobi, Kenya

<sup>2</sup>Department of Mathematics, Statistics and Actuarial Sciences, Kenyatta University, Nairobi, Kenya

<sup>3</sup>Department of Statistics and Actuarial Sciences, JKUAT, Nairobi, Kenya

Email: gideonlangat90@gmail.com, kube.ananda@ku.ac.ke, tttmageto@gmail.com

**How to cite this paper:** Kipngetich, G., Kube, A. and Mageto, T. (2021) A Spatial-Nonparametric Approach for Prediction of Claim Frequency in Motor Insurance. *Open Journal of Statistics*, 11, 493-505.  
<https://doi.org/10.4236/ojs.2021.114031>

**Received:** June 3, 2021

**Accepted:** August 10, 2021

**Published:** August 13, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Spatial modeling has largely been applied in epidemiology and disease modeling. Different methods such as Generalized linear models (GLMs) have been made available to prediction of the claim frequencies. However, due to heterogeneity nature of policies, the methods do not generate precise and accurate claim frequencies predictions; these parametric statistical methods extensively depend on limiting assumptions (linearity, normality, independence among predictor variables, and a pre-existing functional form relating the criterion variable and predictive variables). This study investigates how to derive a spatial nonparametric model estimator based on smoothing Spline for predicting claim frequencies. The simulation results showed that the proposed estimator is efficient for prediction of claim frequencies than the kernel based counterpart. The estimator derived was applied to a sample of 6500 observations obtained from Cooperative Insurance Company, Kenya for the period of 2018-2020 and the results showed that the proposed method performs better than the kernel based counterpart. It is worth noting that inclusion of the spatial effects significantly improves the estimator prediction of claim frequency.

## Keywords

Nonparametric, SAR, Smoothing Spline, Claims, CIC, Spatial

## 1. Introduction

Recent studies on spatial modeling have been rapidly applied in many fields: epidemiology, public health, and the insurance sector. Models such as Poisson,

Generalized linear models, Credibility models and Bayesian Models are the commonly used models for prediction of claim frequencies. However, from the available literature, these models appear to be relatively inflexible. Although the Generalized linear models provide accurate and fast analysis of insurance data, they fall short because they are defined based on the assumptions, and an incorrect model assumption can cause model misspecification leading to erroneous results. Nonparametric models are deemed to minimize the shortcoming of these standard parametric models since fewer assumptions are made for the model, therefore, suitable for modeling insurance data which are nonlinear as described by [1] where they concluded that the nonparametric models perform better than generalized linear models (GLMs); the only observable problem with modeling using Nonparametric models is the interpretation of some of the curves [2]. When modeling claims and risks, we need to determine their behavior and spatial dependence, and spatial heterogeneity of the data so that the insurer can determine which areas are associated with a higher riskier when determining premiums amount to be paid.

[3] proposed a Bayesian nonparametric approach for prediction of claims; here they found out that the model performs better compared to nonparametric GLMs in that it can capture the nonlinear random effects present in the data. [4] also proposed a flexible nonparametric loss model for prediction of the claims; they found out that having flexible multivariate model may allow actuaries to estimate the dependence between different risk classes and different lines of business and this topic needs to be explored further. [5] introduced the idea of using nonparametric data mining approach to modeling the claims and prediction of risk; here the approach classified the risk and predicted claim size based on the data. This study's research idea was to be built based on the idea proposed by [6] [7] where they introduce a nonparametric spatial regression model for prediction. This study's primary objective was to derive a spatial nonparametric estimator for the prediction of insurance claims. Therefore, this study's main contribution was to investigate the estimator's performance in the situation of additional covariates in the model and incorporate the aspect of spatial dependence in deriving a nonparametric estimator for the prediction of insurance claim frequencies.

The main difference between this research and [7] [8] [9] are as follows: 1) estimate a nonparametric spatial model where estimation of the unknown trend  $g(\cdot)$  is based on smoothing spline; 2) spatial heterogeneity and correlation were considered simultaneously rather than assuming that correlation satisfies specific form (such as in SAR).

The paper is organized as follows: Section 2 describes the development process of the model estimator based on smoothing spline; Section 3 presents the data description and main results; simulation study and analysis of CIC insurance claims data; Section 4 presents conclusion and suggestions for further research.

## 2. Methods

### Model Estimation

The study proposed a nonparametric regression model to predict the number of claims  $Y_i, i = 1, \dots, n$  observed in region  $J$  in order to relax restrictive assumption on the distribution of number of claims and  $X_i$  covariates vector for the  $i^{\text{th}}$  claim. Since claims in each region,  $J$  has nonlinear relation with the covariates  $X_i$ 's.

The nonparametric form of the model is given by the general form [10] [11].

$$y_i = g(x_i) + Z_i^T b + \varepsilon_i$$

$g(\cdot)$  is unknown nonparametric function used to model fixed effects,  $Z_i^T b$  and  $\varepsilon_i$  cater for random effects.

Since the form of  $Z_i^T b = R_i$  for  $R_i$  is unknown. The main work of this study is to estimate the form of  $R_i$  then establish the functional form of  $Z_i^T b$  that captures the spatial effects.

Let  $n = (n_1, \dots, n_N)^T$  be two  $N$  dimensional vectors. We can make assumption about the spatial model as

$$Y_i = g(X_i) + R_i, \quad \text{var}(R) = \Sigma, \quad i \in \Lambda_n = \{1, \dots, n_1\} \times \dots \times \{1, \dots, n_N\} \quad (1)$$

where  $i = (i_1, \dots, i_N)$  in  $\Lambda_n$  will be referred to as site,  $R_i$  cater for the spatial effects (Random effects) and the cardinality of  $\Lambda_n$  is  $|\Lambda_n| = n$  [12].

Spatial data is modelled as finite realization of vector stochastic process indexed by  $i \in \Lambda_n$ ,  $R = (R_1, \dots, R_n)^T$  is assumed to follow a joint Gaussian distribution where  $E(R_i) = 0$ , is known  $\forall i \in \Lambda_n$ ,  $\Sigma = [\rho(R_i, R_j)]$  is the unknown correlation coefficient matrix (need to be estimated). The vector

$X_i = (X_{i1}, \dots, X_{id}) \in \mathfrak{R}^d$ ,  $Y_i \in \mathfrak{R}$  and  $g(\cdot)$  is the unknown trend function.

The aim is to estimate  $g(x)$  for some given  $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$ , the response variable  $Y_i$  is claim frequency and  $X_i$  is six dimensional vector consisting of the following explanatory variables: gender, claim amount, age of the policyholder, gender, vehicle age, model of the vehicle and age category of the policyholder.

Estimating  $g(x)$  at some point  $x \in \mathfrak{R}^d$ , for  $X_i$  in the neighbourhood of  $x$ ,  $g$  can be approximated using smoothing spline [13] [14].

To estimate the smoothing spline estimator  $\hat{g}(\cdot)$  of  $g(\cdot)$ , the study considers minimizing the equation

$$\sum_{i=1}^n (Y_i - g(x_i))^2 + \lambda \int (g''(x))^2 dx \quad (2)$$

over the function  $g$  This criterion trades-off least squares error of  $g$  over  $(x_i, y_i), i = 1, \dots, n$ , with a regularization term that grows large when the second derivative of  $g$  is wiggly. The coefficients are chosen to minimize Equation (3) which is a simplified form of Equation (2)

$$\frac{1}{n} \sum_{i=1}^n \{Y_i - g(X_i; \beta)\}^2 + \lambda \beta^T \Omega \beta \quad (3)$$

which can be represented as

$$\|Y_i - G\beta\|^2 + \lambda\beta^T\Omega\beta$$

where  $G \in \mathbb{R}^{n \times n}$  is basis matrix defined as

$$G_{ij} = \psi_j(x_i), \quad i, j = 1, \dots, n$$

where  $\psi_1, \dots, \psi_n$  are the truncated power basis functions with knots at  $x_1, \dots, x_n$  which is evaluated at the data values

$$\psi_j(x) = \begin{cases} x_i^j & (0 \leq n \leq p) & i, j = 1, \dots, n \\ (x_i - N_{j+1-p})_+^p & (p+1) \leq j \leq N \end{cases} \quad (4)$$

$(x - N_{j+1-p})_+^p = \max(0, x_i - N_j)^p, j \in \phi$  where  $\phi$  is compact interval.  $p$  is the degree of the spline and  $j_1 < \dots < j_{N-p}$  are fixed points or knots in  $\phi$ .

$\Omega \in \mathbb{R}^{n \times n}$  is the penalty matrix defined as

$$\Omega_{ij} = \int g_i''(x)\psi_j''(x)dx, \quad i, j = 1, \dots, n$$

Given the optimal coefficients  $\hat{\beta}$  minimizing (3) through penalized least squares, the smoothing spline estimator at  $x$  is therefore defined as

$$\hat{g}(x_i) = \sum_{j=1}^n \hat{\beta}_j \psi_j(x) \quad (5)$$

The term affects shrinking the components of estimation  $\hat{\beta}$  towards zero. The parameter  $\lambda \geq 0$  is the smoothing parameter.

Each computed coefficient  $\hat{\beta}_j$  corresponds to a particular basis function  $\psi_j$ . The term  $\beta^T\Omega\beta$  in (3) imparts more shrinkage on the coefficients  $\hat{\beta}_j$  that correspond to wigglier functions  $\psi_j(x)$ . Hence, as we increase  $\lambda$ , we are shrinking away from the wiggler basis functions.

Similar to least squares regression, the coefficients  $\hat{\beta}$  minimizing (3) is

$$\hat{\beta} = (G^T G + \lambda\Omega)^{-1} G^T Y = (X^T X + n\lambda D)^{-1} X^T Y$$

where  $X$  is a design matrix with entries  $x_i$  for  $i = 1, \dots, n$ ,  $Y$  is a vector of the response variables,  $D$  is a diagonal matrix with  $p + 1$  zeros on the diagonal followed by  $N$  ones and  $n\lambda D$  is a penalty term.

Smoothing splines can be seen as a linear smoother, where

$k(x) = (\psi_1(x_1), \dots, \psi_n(x_n))$ . Therefore, Equation (5) can be represented as

$$\hat{g}(x) = k(x)^T \hat{\beta} = k(x)^T (X^T X + n\lambda D)^{-1} X^T Y \quad (6)$$

which is linear combination of the points  $y_i, i = 1, \dots, n$ ,  $\lambda$  is estimated using Generalized Cross Validation (GCV) method given by

$$GCV(\lambda) = \frac{1}{n} \sum_{i=1}^n \left( \frac{Y(z_i) - \hat{Y}_\lambda^{-i}(z_i)}{1 - (p + tr(S_\lambda))/n} \right)^2 \quad (7)$$

where  $Y(z_i)$  is the observation in point  $z_i$ ,  $Y_\lambda^{-i}(z_i)$  is the predicted value from a fitted smoothing spline model from the data less the  $i^{\text{th}}$  data and  $S_\lambda$  is the degree of the smoother.

As proposed by [6] [7],  $R^2$  is used to assess the performance of predictor function, given by

$$R^2 = 1 - \frac{\sum_{i=1}^n [g(x_i) - \hat{g}(x_i)]^2}{\sum_{i=1}^n [g(x_i) - \bar{g}]^2} \quad (8)$$

where  $\bar{g}$  is the sample mean of  $g(x_i), i = 1, \dots, n$ .

After estimating the function  $g(\cdot)$ , then from (1)  $R_i$  is estimated as  $\hat{R}_i = Y_i - \hat{g}(X_i)$ . Since  $\Sigma$  in model Equation (1) is unknown, we assume that  $R_i, i = 1, 2, \dots, n$  is 2<sup>nd</sup>-order stationary and isotopic process (does not depend on direction).

Before prediction can be performed on spatial data sets, the variogram is usually estimated at various lags and a nonparametric model is fitted to those estimates.

Then let  $C(h)$  and  $2\gamma(h)$  be covariogram and variogram of the process where  $h$  represents the distance between 2 points at which the process is obtained [12] [15]. The two quantities are related by

$$C(h) = C(0) - \gamma(h) \quad (9)$$

where  $C(0) = \sigma^2 = \text{var}(Y(z))$ ,  $Y(z)$  is the value of the process at spatial location  $z$  within region  $C$ .

$$\lim_{h \rightarrow \infty} C(h) = 0$$

implies

$$\lim_{h \rightarrow \infty} \gamma(h) = \text{Var}(Y(z)) = C(0)$$

for validity of variogram the condition that

$$\lim_{h \rightarrow \infty} \frac{2\gamma(h)}{h^2} = 0$$

must be met [16].

$\Sigma = [\rho(R_i, R_j)] = [C(\|z_i - z_j\|) / \sigma^2]$ , while  $z_i$  and  $z_j$  are the spatial locations associated with the error values  $R_i$  and  $R_j$  thus to estimate  $\Sigma$  it is sufficient to estimate  $\gamma(h)$  [16] [17].

$$2\hat{\gamma}(h) = \sum_{S(h)} [z_i - z_j]^2 / N(h) \quad (10)$$

$S(h) = \{(z_i, z_j) : |z_i - z_j| = h\}$ ,  $h \in \mathfrak{R}^d$ ,  $N(h)$  is a number of distinct pairs in  $S(h)$  since  $r(z_i)$  the error at location  $z_i$  is unobserved, the quantity is to be estimated as well.

Since we have to estimate the variogram  $\hat{\gamma}(h)$  in Equation (10) in nonparametric approach [18] [19], then  $\gamma(h)$  can be estimated as

$$\gamma(h) = \int_0^\infty (1 - \omega_d(ht)) dM(t) \quad (11)$$

$M(t)$  is nonnegative bounded nondecreasing function for nodes(or location of the jumps)  $t \geq 0$  and  $\omega_d$  is a basis for functions in  $\mathbb{R}^d$  ( $d$  is the dimension

of the spatial domain  $D$ ) given by

$$\omega_d(ht) = (2/ht)^{(d-2)/2} \Gamma(d/2) J_{(d-2)/2}(ht)$$

$\Gamma(d/2)$  is the gamma function, and  $J_{(\cdot)}$  is the Bessel function of the first kind. Some familiar examples of  $\omega_d$  are  $\omega_1(ht) = \cos(ht)$ ,  $\omega_2(ht) = J_0(ht)$ , and here  $\omega_3(ht) = \frac{\sin(ht)}{ht}$  is chosen which yields a non-parametric estimate which is conditionally negative definite for spatial data from 1 - 3 dimensions.

The characteristics of the estimator (11) are estimated using Integrated square error [20], given by

$$\text{ISE}(\gamma) = \int_{h_1}^{h_k} \{\hat{\gamma}(h) - \gamma(h)\}^2 dh \quad (12)$$

where  $h_1$  and  $h_k$  are the smallest and largest distances for which variogram estimates are available [17].

Model (1) can therefore be represented as

$$Y(z_i) = g(X_i(z_i)) + R(z_i), \quad i = 1, \dots, n \quad (13)$$

where  $Y(z_i): i = 1, \dots, n$  is the observations (claims) in region  $z_i$  associated with independent variables  $X_i(z_i)$  in region  $z_i$ ,  $R(z_i)$  is the unobserved error in region  $z_i$  and  $g(\cdot)$  is the estimated function in (6).

To evaluate performance of the proposed method we used  $R^2$  to assess prediction accuracy of the method

$$R^2 = 1 - \frac{\sum_{i=1}^n [Y(z_i) - \hat{Y}(z_i)]^2}{\sum_{i=1}^n [Y(z_i) - \bar{Y}]^2} \quad (14)$$

### 3. Main Results

#### 3.1. Data Description

The study used motor third party liability data for 2018-2020 from the insurance company Cooperative Insurance Company (CIC). The data include 6500 policies, out of which many policies have total claim sizes other than zero, and an appropriate number of policies without any claims were taken. The following policy data were used: the region where the policy was taken, age, gender, type of vehicle, number of claims per policy, years of policy ownership, claim amount, insured cases number for a user, and average claim size. In the process of preparation, data was cleaned, and imputation of data will be done; age is categorized into old (over the age of 50), Young (up to the age of 25), and Middle (aged 25 - 50) age. Policies with extremely low and extremely high average claim sizes are removed; categorical variables with multiple categories were replaced with dummy (indicator) variables.

**Table 1** shows that there is a very large number of observations with no claims in claims dataset where the maximum number of claims made in a region was 4 in an observation.

**Table 1.** Summary of claim frequencies in the data.

No. of Claims	Freq of Observations	% of Observations
0	4015	61.8
1	1967	30.3
2	431	6.6
3	71	1.1
4	16	0.2

### 3.2. Simulation Study

This section describes the simulation and their analysis results of the proposed method, we simulate spatial data with a length of  $n = 100$  observations. This is to ensure that the simulated data mimic the real claims dataset so that the results can be inferred to evaluate the performance of our method in data analysis. 65 spatial sampling locations were selected randomly and denoted by  $z_1, \dots, z_n$ . The responses  $Y(z_i)$  for  $i = 1, \dots, n$  are the observations and were simulated from the spatial nonparametric model (13) with  $p = 2$

$$Y(z) = g(x_i(z)) + R(z)$$

$R(z)$  is the term for spatial effects  $z_i$  and  $z_j$  in 2-dimensional space [16] with mean 0 and covariance given by (11). The covariates,  $x_i(z_i)$  for  $i = 1, \dots, n$ , were generated as iid  $N(0,1)$  and are independent of each other, before the simulations *i.e.*, the variables were treated as fixed terms when  $Y(z_i)$  were generated repeatedly. Within each simulation, the spatial random effects  $R(z_i)$  were generated from a Gaussian process with mean zero and the covariance function (11), for  $i = 1, \dots, n$ .

The  $ISE(\gamma)$  defined as  $ISE(\gamma) = \int_{h_1}^{h_k} \{\hat{\gamma}(h) - \gamma(h)\}^2 dh$  was approximated numerically from simulated data for the proposed estimator (11) and NW kernel. **Table 2** present the mean values of ISE. From the results in Table, the proposed estimator (11) offers a better performance compared to NW kernel estimator.

Assessing how well the proposed method performs, we compare the proposed method under which  $R = \int_0^\infty (1 - \omega_d(ht)) dM(t)$  with the method under which the spatial component ( $R$ ) is based on kernel estimation, we calculated the MSE and the  $R^2$  of the estimators from 100 simulations and present the results in **Table 3**. From the table it was found that MSE for the proposed method is smaller ranging from 0.0221 to 0.0102 in all the sample sizes taken while the MSE of the kernel based estimator ranges between 0.308 to 0.0176, in addition, the  $R^2$  for the proposed method were larger ranging between 0.7003 to 0.99963 in all the sample sizes compared those of kernel based estimator which ranges from 0.6751 to 0.9694. Thus the results demonstrate the superior performance of the proposed method compared to the kernel based estimator.

The results in **Table 3** were visualized in **Figure 1** and **Figure 2**.

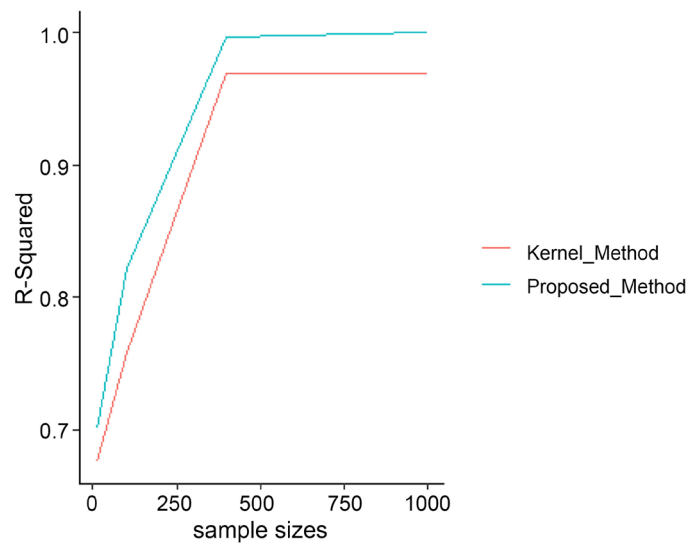
Based on performance of the proposed method, the method was applied to the simulated data to check its performance in prediction of future values. **Table 4** describes the distribution of the predicted values out of 100 simulation.

**Table 2.** Mean values of the standardized ISE from the estimators (Sample size of  $n = 1000$ ).

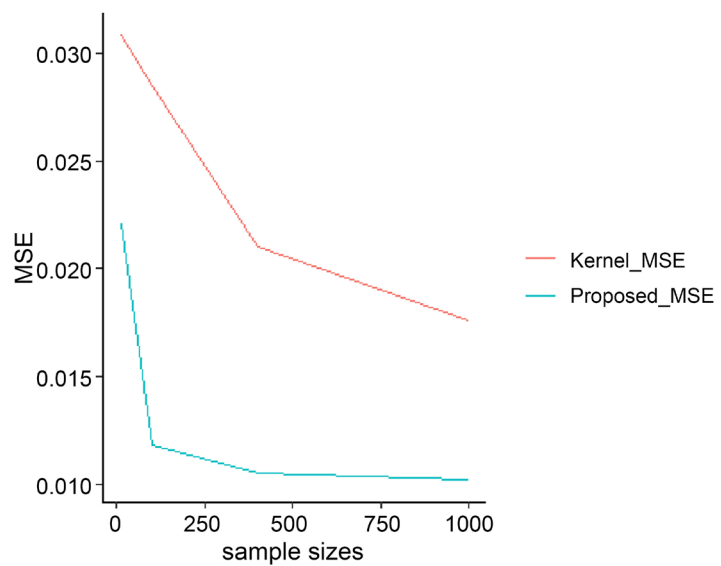
	$h = 1$	$h = 2$	$h = 3$	$h = 4$
NW kernel	0.53883	0.53407	0.5177	0.40779
Proposed estimator	0.30731	0.28762	0.26737	0.23037

**Table 3.** MSE and  $R^2$  for the model over 100 simulation under different sample sizes.

	$n = 10$		$n = 100$		$n = 400$		$n = 1000$	
	MSE	$R^2$	MSE	$R^2$	MSE	$R^2$	MSE	$R^2$
$N(kernel)$	0.0308	0.6751	0.0285	0.7585	0.0210	0.9691	0.0176	0.9694
Proposed Method	0.0221	0.7003	0.0118	0.8217	0.0105	0.9962	0.0102	0.99963

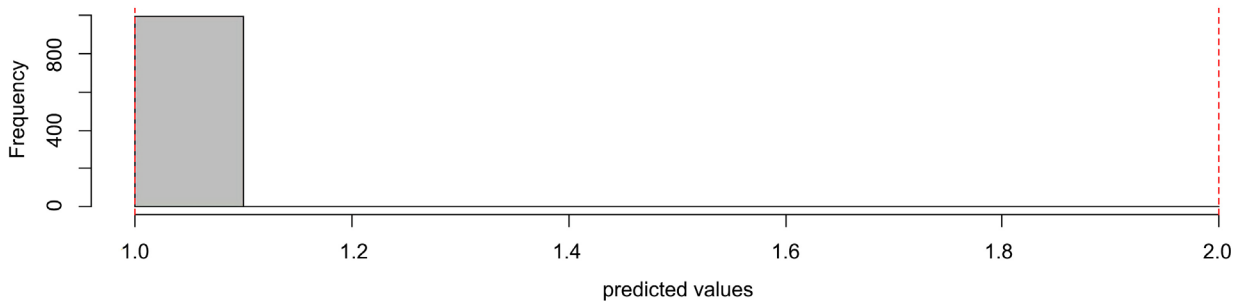


**Figure 1.** R-squared plot.



**Figure 2.** Mean squared plot of both K (kernel) and proposed estimator.





**Figure 3.** Histogram for the predicted values.

**Table 4.** Summary of predicted claim frequencies from simulation.

No. of Claims	Freq of Observations	% of Observations
1	998	99.8
2	2	0.2

The predicted values generated by the proposed method as presented in **Table 4** were graphically presented and the prediction intervals superimposed on the distributional histogram of the predicted values as shown in **Figure 3**, the prediction interval (in red dotted lines) showed that a larger number of predicted values lies between 1 - 2 this means that there are higher chances of getting future values as 1 and 2.

### 3.3. Analysis for Claims Data

The study considered claims data from CIC insurance observed in different parts of 7 counties of Kenya to exhibit the performance of the proposed method. The main interest of this study was predicting claims frequencies, the study considers a set of 6500 observations. Let  $Y_i$  denote the claim frequency, and  $\mathbf{X}_i = (X_1, \dots, X_6)^T$  be a vector which consists of the following explanatory variables: gender, claim amount, age of the policyholder, gender, vehicle age, model of the vehicle and age category of the policyholder. Using the estimated model (13) we predict claim frequencies. The observations were from random process over a countable sample of spatial locations. The claim data at a particular location typically represent the entire region (**Figure 4**).

Using the proposed method future claim frequencies were predicted and the results were presented in **Table 5**.

**Figure 5** shows graphical representation of the predicted claims as described in **Table 5** with the prediction interval (red dotted vertical lines), the future values will lie between 1 and 4.

From the prediction results,  $R^2$  values using Equation (14) were evaluated to access the performance of two methods, the results presented in **Table 6**.

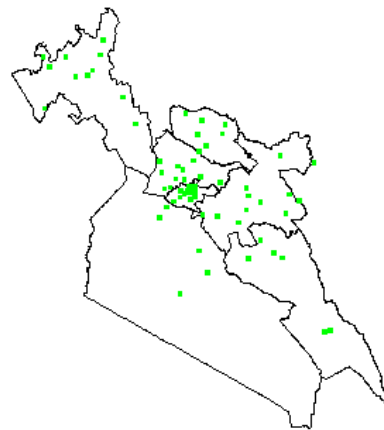
From the results in **Table 6** the  $R^2$  for  $N(kernel)$  was 0.543, and that from the proposed method is 0.566, this showed that the proposed method for prediction has a higher prediction accuracy than the kernel based estimator. Therefore the

**Table 5.** Summary of predicted claim frequencies.

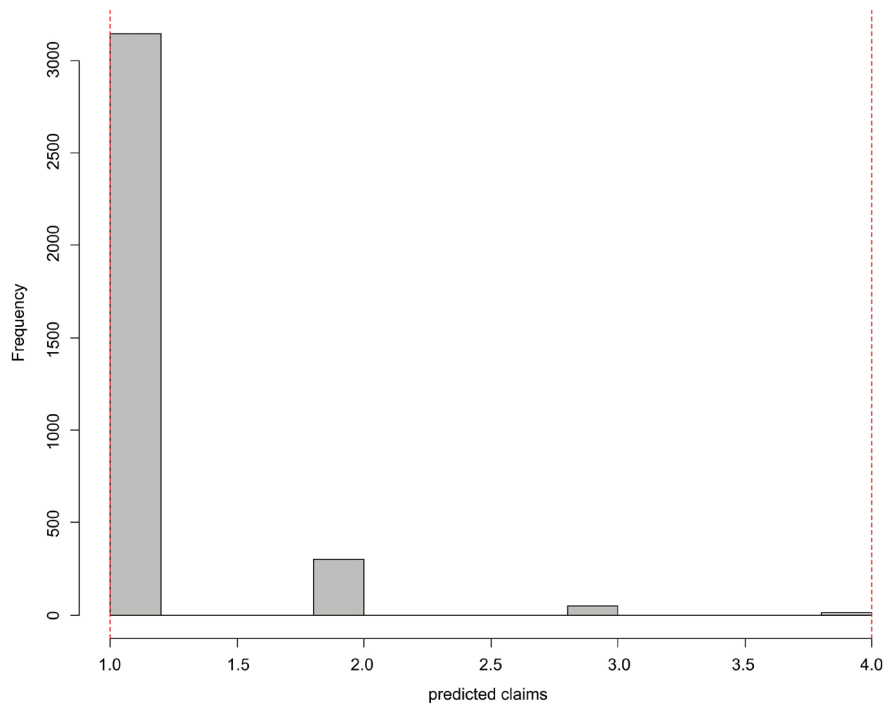
No. of Claims	Freq of Observations	% of Observations
1	3145	89.52
2	302	8.60
3	50	1.42
4	16	0.50

**Table 6.**  $R^2$  for the estimators.

Method	$N(\text{kernel})$	Proposed Method
$R^2$	0.543	0.566



**Figure 4.** Hotspots locations in Nakuru, Nairobi, Kajiado, Muranga, Kiambu, Machakos, Makueni counties.



**Figure 5.** Predicted number of claims.

study concluded that the proposed method is more efficient than  $N$  (*kernel*) model, this implies that the predicted value was more likely to be more identically equal to the observed claims.

#### **4. Conclusions and Suggestions**

The idea of deriving an appropriate estimator in predicting frequency claims in the insurance industry has gained more interest in finance and statistical research. Many researchers heavily rely on parametric estimators; however, the insurance datasets have some aspect of non-linearity. Hence, researchers in statistics and econometrics are currently developing nonparametric models incorporating spatial effects to improve on the prediction based on the existing parametric models such as aggregate claim models and GLMs which are rather more restrictive on their transformed mean of the response; the nonparametric methods provide a more flexible method for prediction. The study proposed a spatial nonparametric (based on splines) estimator for predicting claim frequencies in motor insurance.

The simulation study showed that the proposed method performs better than the kernel based estimator; here the Mean Squared Error values of the proposed method were smaller than those of the kernel estimator which also implies a higher value of R-squared, particularly in presence of spatial dependence. Case study findings also showed that the proposed method performs better than the kernel based estimator on predicting the future claim frequencies. Therefore, the proposed method compared to kernel based estimator provides a more efficient prediction method for motor insurance claim data and ultimately leads to more accurate predictions.

#### **Suggestions**

Some additional exogenous variables such as environmental among other institutional factors may have effect on claim frequencies therefore, more robust spatial estimator need to be constructed using the proposed idea to investigate how these factors may affect claim frequencies. Further research can also be done on the theoretical properties of this proposed model estimator. In addition, this study made the assumption that the errors were correlated, for this reason future studies could consider a case of uncorrelated error structure.

#### **Acknowledgements**

Sincere thanks to my supervisors Dr. Kube Anada and Dr. Thomas Mageto for their professional contribution and performance, and special thanks to my parents for their moral support and rare attitude of high quality.

#### **Conflicts of Interest**

The authors declare no conflicts of interest.

## References

- [1] Doupe, P., Faghmous, J. and Basu, S. (2019) Machine Learning for Health Services Researchers. *Value in Health*, **22**, 808-815. <https://doi.org/10.1016/j.jval.2019.02.012>
- [2] Islamiyati, A., Chamidah, N., *et al.* (2020) Penalized Spline Estimator with Multi Smoothing Parameters in Bi-Response Multi-Predictor Nonparametric Regression Model for Longitudinal Data. *Songklanakarin Journal of Science & Technology*, **42**, 897-909.
- [3] Fellingham, G.W., Kottas, A. and Hartman, B.M. (2015) Bayesian Nonparametric Predictive Modeling of Group Health Claims. *Insurance: Mathematics and Economics*, **60**, 1-10. <https://doi.org/10.1016/j.insmatheco.2014.10.011>
- [4] Hong, L. and Martin, R. (2017) A Flexible Bayesian Nonparametric Model for Predicting Future Insurance Claims. *North American Actuarial Journal*, **21**, 228-241. <https://doi.org/10.1080/10920277.2016.1247720>
- [5] Kaščelan, V., Kaščelan, L. and Burić, M.N. (2016) A Nonparametric Data Mining Approach for Risk Prediction in Car Insurance: A Case Study from the Montenegrin Market. *Economic Research—Ekonomika istraživanja*, **29**, 545-558. <https://doi.org/10.1080/1331677X.2016.1175729>
- [6] Wang, H.X., Wang, J.D. and Huang, B. (2012) Prediction for Spatiotemporal Models with Autoregression in Errors. *Journal of Non-Parametric Statistics*, **24**, 217-244. <https://doi.org/10.1080/10485252.2011.616893>
- [7] Wang, H.X., Lin, J.G. and Wang, J.D. (2016) Nonparametric Spatial Regression with Spatial Autoregressive Error Structure. *Statistics*, **50**, 60-75. <https://doi.org/10.1080/02331888.2015.1094068>
- [8] Li, Y., Qin, Y. and Li, Y. (2021) Empirical Likelihood for Nonparametric Regression models with Spatial Autoregressive Errors. *Journal of the Korean Statistical Society*, **50**, 447-478. <https://doi.org/10.1007/s42952-020-00088-z>
- [9] Xu, G.Y. and Bai, Y. (2020) Estimation of Nonparametric Additive Models with High Order Spatial Autoregressive Errors. *Canadian Journal of Statistics*, **49**, 311-343.
- [10] Rice, J.A. and Wu, C.O. (2001) Nonparametric Mixed Effects Models for Unequally Sampled Noisy Curves. *Biometrics*, **57**, 253-259. <https://doi.org/10.1111/j.0006-341X.2001.00253.x>
- [11] Karcher, P. and Wang, Y.D. (2001) Generalized Nonparametric Mixed Effects Models. *Journal of Computational and Graphical Statistics*, **10**, 641-655. <https://doi.org/10.1198/106186001317243377>
- [12] Wang, H.X., Wu, Y.H. and Chan, E. (2017) Efficient Estimation of Nonparametric Spatial Models with General Correlation Structures. *Australian & New Zealand Journal of Statistics*, **59**, 215-233. <https://doi.org/10.1111/anzs.12192>
- [13] Wang, Y.D. (2019) Smoothing Splines: Methods and Applications. Chapman and Hall/CRC, Boca Raton, FL.
- [14] Tait, A. and Woods, R. (2007) Spatial Interpolation of Daily Potential Evapotranspiration for New Zealand Using a Spline Model. *Journal of Hydrometeorology*, **8**, 430-438. <https://doi.org/10.1175/JHM572.1>
- [15] Laslett, G.M. (1994) Kriging and Splines: An Empirical Comparison of Their Predictive Performance in Some Applications. *Journal of the American Statistical Association*, **89**, 391-400. <https://doi.org/10.1080/01621459.1994.10476759>
- [16] Cressie, N. (2015) Statistics for Spatial Data. John Wiley & Sons, Hoboken.
- [17] Huang, C.F., Hsing, T. and Cressie, N. (2011) Nonparametric Estimation of the Variogram and Its Spectrum. *Biometrika*, **98**, 775-789.

<https://doi.org/10.1093/biomet/asr056>

- [18] Fernández-Casal, R., Castillo-Páez, S. and Francisco-Fernández, M. (2018) Nonparametric Geostatistical Risk Mapping. *Stochastic Environmental Research and Risk Assessment*, **32**, 675-684. <https://doi.org/10.1007/s00477-017-1407-y>
- [19] Qadir, G.A. and Sun, Y. (2020) Semiparametric Estimation of Cross-Covariance Functions for Multivariate Random Fields. *Biometrics*, **77**, 547-560. <https://doi.org/10.1111/biom.13323>
- [20] Yu, K.M., Mateu, J. and Porcu, E. (2007) A Kernel-Based Method for Nonparametric Estimation of Variograms. *Statistica Neerlandica*, **61**, 173-197. <https://doi.org/10.1111/j.1467-9574.2007.00326.x>