*Electronic Research Archive*

https://www.aimspress.com/journal/era

ERA, 34(7): 4698–4724.
DOI: 10.3934/era.2026207
Received: 07 April 2026
Revised: 10 May 2026
Accepted: 18 May 2026
Published: 01 June 2026

*Research article*

# A heteroscedastic parametric method with Bayesian inference for interval-valued regression models

**Ruiqin Tian**[1], **Ke Liu**[1], **Guangyu Wang**[1] **and Dengke Xu**[2,*]

[1] School of Mathematics, Hangzhou Normal University, Hangzhou 311121, China
[2] School of Economics, Hangzhou Dianzi University, Hangzhou 310018, China

* **Correspondence:** Email: xudengke1983@163.com.

**Abstract:** As a typical form of symbolic data, interval-valued data provides an effective framework to analyze large-scale datasets. Most existing interval regression studies focus on classical methods, while research that incorporates heteroscedasticity within the Bayesian framework remains limited. This paper extends the existing parametric method for interval-valued data to a Bayesian heteroscedastic framework, and further develops the Bayesian Heteroscedastic Parametric Method (BHPM). By explicitly modeling heteroscedasticity in the regression structure, we conduct Bayesian inference using Gibbs sampling and the Metropolis–Hastings algorithm, thus enhancing the model's interpretability and generalization performance. Both simulation studies and real-data applications demonstrate that the extended BHPM achieves superior performance over traditional methods.

**Keywords:** interval-valued data; Bayesian regression; heteroscedasticity; parametrized method

## 1. Introduction

With the widespread adoption of the Internet of Things and large-scale data collection technologies, traditional single-value data can no longer fully reflect the uncertainty and variability of real observed values. Interval-valued data, as a typical representation of symbolic data, retain the fluctuation range of original observations in the form of upper and lower bounds, and are widely used in fields such as financial risk measurements and environmental quality assessments. For example, Liu et al. [1] developed a hybrid ensemble model for interval-valued carbon price forecasting via multi-scale decomposition and multi-model fusion. Wang et al. [2] proposed an interval-valued linear model, derived its least square estimator and properties, and validated it via simulations and an application to latitude's impact on urban temperatures.

Regarding linear regression modeling for interval-valued data, scholars have now developed various classical methods. Billard and Diday [3] introduced the center method (CM), which is a foundational

framework for early interval regressions that constructs regression equations using interval midpoints as single-valued features. Kong and Gao [4] proposed the Min-Max method (MM), which treats the upper and lower bounds of intervals as independent variables to construct a dual regression system. Lima-Neto et al. [5] proposed the center-range method (CRM), which further integrates interval midpoints and ranges to characterize the position and dispersion of intervals, thus enhancing the model's interpretability. Subsequently, Hao and Guo [6], Xu and Qin [7], and other scholars conducted constrained optimization and Bayesian extensions on this model. Souza et al. [8] proposed the parametrized method (PM), which finds the optimal reference point through fitting, thereby breaking the limitation of fixed reference points. Xu and Qin [9] further developed the Bayesian parametrized method (BPM) on this basis. Duary et al. [10] utilized the PM in an interval-valued production model, thus optimizing the average profit and verifying the model with numerical analyses. Additionally, there are many other studies on interval-valued models; see [11–13] for details.

Although the numerous methods mentioned above have achieved remarkable progress in interval-valued regression modeling, the ordinary least squares (OLS) estimation has multiple limitations when dealing with interval data. The Bayesian estimation framework provides an effective solution to these problems. By introducing a prior distribution and combining it with the likelihood function to obtain posterior inference, it can not only quantify the parameter uncertainty but also mitigate the impact of multicollinearity through a shrinkage prior. It should be noted that many researchers had explored this problem. For example, Xu and Qin [7] first introduced the Bayesian method into the CRM. Subsequently, Xu and Qin [14] further constructed a Bayesian framework based on Jeffreys' prior, thus solving the problem of prior specifications for interval-valued data. Ji et al. [15] extended the Bayesian method to panel interval-valued data with fixed effects. Zhang et al. [16] developed a Bayesian nonparametric interval regression model based on the asymmetric Laplace distribution and variational inference. Goffard and Laub [17] applied approximate a Bayesian computation to fit and compare insurance loss models with aggregated data. Additionally, several related studies have investigated Bayesian approaches for statistical modeling and inference; see [18–20] for details.

Heteroscedasticity is a widely encountered phenomenon in regression modeling. When neglected, it can severely compromise the validity of the model estimation and the accuracy of the predictive performance. Traditional methods, such as weighted least squares [21] and generalized least squares [22], are widely used to address heteroscedasticity in regressions. To better characterize varying the variance structures, the joint mean–variance modeling approach has attracted increasing attention. Xu and Zhang [23] considered the joint mean–variance modeling approach in the context of heteroscedasticity. Additionally, plenty of scholars have further optimized and extended the joint mean-variance framework in recent years [24, 25]. In the context of interval-valued data research, targeted heteroscedasticity correction methods have also been proposed. Souza et al. [8] adopted the Box-Cox transformation to alleviate heteroscedastic interference in interval range prediction. Zhong et al. [26] constructed a constrained interval regression model and achieved an effective heteroscedasticity adjustment through reasonable endpoint constraint settings. Nevertheless, current researches on estimating heteroscedastic interval-valued regression models using Bayesian methods remains relatively scarce.

Drawing on the above research background, this paper extends the existing parametrized method to a Bayesian heteroscedastic framework. Compared with classic interval-valued models [3–5], the Bayesian Heteroscedastic Parametric Method (BHPM) utilizes a parametric transformation approach to extract the optimal reference point of interval values; relative to the homoscedastic interval regression

framework [9] developed by Xu and Qin, this method further addresses heteroscedasticity via a shared log-linear heteroscedastic covariance matrix that characterizes the joint fluctuations of both interval bounds. Compared with [23], the extended method considers interval-valued data, which can better characterize the uncertainty in the real data. Moreover, grounded in the Bayesian framework [14], an appropriate prior is constructed for the model parameters, and the Bayesian estimates of unknown parameters are derived by combining Gibbs sampling and Metropolis-Hastings (MH) algorithms.

The remainder of this paper is organized as follows: in Section 2, we review the heteroscedastic model and three current regression models for interval-valued data; Section 3 provides a detailed introduction to the model structure of the BHPM, as well as its connections to other representative interval-valued data regression models; Section 4 presents the basic theory to estimate unknown parameters in the BHPM using Bayesian methods; Sections 5 and 6 report the estimation results based on synthetic and real datasets, respectively; and finally, Section 7 presents the concluding remarks.

## 2. Current models review

### 2.1. Heteroscedasticity in classical and joint mean-variance modeling

For the $i$-th individual observation ($i = 1, 2, \ldots, n$), the classical linear regression model is given by the following:

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta} + \varepsilon_i, \quad \varepsilon_i \overset{\text{ind.}}{\sim} N(0, \sigma^2), \tag{2.1}$$

where $\mathbf{X}_i = [X_{i1}, \ldots, X_{ip}]^T$ is a $p \times 1$ covariate column vector, and $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_p]^T$ is the corresponding $p \times 1$ unknown regression coefficient vector.

Its core homoscedasticity assumption of constant conditional response variance $\mathrm{Var}(Y_i \mid \mathbf{X}_i) = \sigma^2$ is often violated in empirical work by heteroscedasticity, namely systematic variation in $\mathrm{Var}(Y_i \mid \mathbf{X}_i) = \sigma_i^2$ with covariates, thus yielding inefficient OLS estimates, biased standard errors, and invalid inference.

Joint mean-variance models address this by simultaneously modeling the response's conditional mean and variance, thus offering greater flexibility and predictive performances under heteroscedasticity. The standard parametric form [23] is as follows:

$$\begin{cases} Y_i \sim N(\mu_i, \sigma_i^2), \\ \mu_i = \mathbf{X}_i^T \boldsymbol{\beta}, \\ \log(\sigma_i^2) = \mathbf{Z}_i^T \boldsymbol{\gamma}, \end{cases} \tag{2.2}$$

where $\mu_i$ is the conditional mean of $Y_i$, $\sigma_i^2$ is the conditional variance (log link ensures non-negativity), $\mathbf{Z}_i$ is the covariate vector of the variance model (may fully overlap with $\mathbf{X}_i$), and $\boldsymbol{\gamma}$ is the corresponding unknown coefficient vector.

Designed for point-valued data, these models cannot be directly applied to interval-valued observations without discarding critical boundary information or generating incoherent predictions. Thus, we extend this joint framework to interval-valued data.

### 2.2. Regression modeling for interval-valued data

Interval-valued data, where each observation is represented as a closed interval $[y_i^l, y_i^u]$ with $y_i^l \le y_i^u$, have garnered significant attention in symbolic data analyses. To model the relationship between

interval-valued responses and predictors, several parametric approaches have been proposed.

In this paper, we focus on three classical parametric interval-valued regression models: the MM, CRM, and PM. Let $\mathbf{Y}^I = (Y_1^I, \ldots, Y_n^I)^T$ denote the interval-valued response variable, where $Y_i^I = [y_i^l, y_i^u]$. Additionally, the predictor matrix is interval-valued, denoted as $\mathbf{X}^I = (\mathbf{X}_1^I, \ldots, \mathbf{X}_p^I)$, where $\mathbf{X}_j^I = (X_{1j}^I, \ldots, X_{nj}^I)^T$ and $X_{ij}^I = [x_{ij}^l, x_{ij}^u]$. An interval can be equivalently represented by its center and range as follows:

$$y_i^c = \frac{y_i^l + y_i^u}{2}, \quad y_i^r = \frac{y_i^u - y_i^l}{2}, \quad x_{ij}^c = \frac{x_{ij}^l + x_{ij}^u}{2}, \quad x_{ij}^r = \frac{x_{ij}^u - x_{ij}^l}{2}, \tag{2.3}$$

then we denote the center and range vectors for the response and predictors as $\mathbf{Y}^c, \mathbf{X}^c$ and $\mathbf{Y}^r, \mathbf{X}^r$, respectively.

#### 2.2.1. Min–Max method

The MM [4] directly constructs separate linear regression models for the lower and upper bounds of the intervals as follows:

$$\mathbf{Y}^l = \mathbf{X}^l\boldsymbol{\beta}^l + \boldsymbol{\varepsilon}^l, \quad \mathbf{Y}^u = \mathbf{X}^u\boldsymbol{\beta}^u + \boldsymbol{\varepsilon}^u, \tag{2.4}$$

where $\mathbf{Y}^l = (y_1^l, \ldots, y_n^l)^T$, $\mathbf{Y}^u = (y_1^u, \ldots, y_n^u)^T$, and $\mathbf{X}^l$ and $\mathbf{X}^u$ are the design matrices formed by the lower and upper bounds of the predictors, respectively. The coefficient vectors $\boldsymbol{\beta}^l$ and $\boldsymbol{\beta}^u$ are independently estimated. Although other estimation approaches are available, we focus on the OLS method for illustration, thus yielding the following estimates:

$$\hat{\boldsymbol{\beta}}^l = ((\mathbf{X}^l)^T\mathbf{X}^l)^{-1}(\mathbf{X}^l)^T\mathbf{Y}^l, \quad \hat{\boldsymbol{\beta}}^u = ((\mathbf{X}^u)^T\mathbf{X}^u)^{-1}(\mathbf{X}^u)^T\mathbf{Y}^u. \tag{2.5}$$

The prediction is directly performed on the bounds as follows:

$$\hat{Y}^l = \tilde{\mathbf{X}}^l\hat{\boldsymbol{\beta}}^l, \quad \hat{Y}^u = \tilde{\mathbf{X}}^u\hat{\boldsymbol{\beta}}^u. \tag{2.6}$$

#### 2.2.2. Center and Range method

The CRM [5] decomposes the interval-valued regression into two independent classical linear models for the center and range as follows:

$$\mathbf{Y}^c = \mathbf{X}^c\boldsymbol{\beta}^c + \boldsymbol{\varepsilon}^c, \quad \mathbf{Y}^r = \mathbf{X}^r\boldsymbol{\beta}^r + \boldsymbol{\varepsilon}^r, \tag{2.7}$$

where $\boldsymbol{\beta}^c = (\beta_0^c, \beta_1^c, \ldots, \beta_p^c)^T$ and $\boldsymbol{\beta}^r = (\beta_0^r, \beta_1^r, \ldots, \beta_p^r)^T$ are the vectors of regression coefficients for the center and range models, respectively. The error terms are assumed to follow $\boldsymbol{\varepsilon}^c \sim N(\mathbf{0}, \sigma_c^2\mathbf{I}_n)$, $\boldsymbol{\varepsilon}^r \sim N(\mathbf{0}, \sigma_r^2\mathbf{I}_n)$, and are mutually independent. The OLS estimates of the parameters are given by the following:

$$\hat{\boldsymbol{\beta}}^c = ((\mathbf{X}^c)^T\mathbf{X}^c)^{-1}(\mathbf{X}^c)^T\mathbf{Y}^c, \quad \hat{\boldsymbol{\beta}}^r = ((\mathbf{X}^r)^T\mathbf{X}^r)^{-1}(\mathbf{X}^r)^T\mathbf{Y}^r. \tag{2.8}$$

For a new interval-valued predictor observation $\tilde{\mathbf{X}}^I = ([\tilde{x}_1^l, \tilde{x}_1^u], \ldots, [\tilde{x}_p^l, \tilde{x}_p^u])$, the corresponding interval response prediction is obtained by the following:

$$\hat{Y}^c = \tilde{\mathbf{X}}^c\hat{\boldsymbol{\beta}}^c, \quad \hat{Y}^r = \tilde{\mathbf{X}}^r\hat{\boldsymbol{\beta}}^r, \tag{2.9}$$

$$\hat{Y}^l = \hat{Y}^c - \hat{Y}^r, \quad \hat{Y}^u = \hat{Y}^c + \hat{Y}^r, \tag{2.10}$$

where $\tilde{\mathbf{X}}^c$ and $\tilde{\mathbf{X}}^r$ are the vectors of center and range for the new predictor intervals, respectively.

While the CRM is straightforward and interpretable, the predicted range may be negative, thus leading to logically invalid predicted intervals (i.e., $\hat{Y}^l > \hat{Y}^u$). The most common ad-hoc remedy is to truncate any negative predicted range to zero.

### 2.2.3. The Parametrized method

The conventional MM and CRM methods rely on fixed summary statistics (e.g., center and range) that may not adapt to the underlying structure of interval-valued data. A more flexible alternative is the PM [8], which represents each interval-valued predictor via a data-driven reference point obtained through a convex combination of its lower and upper bounds [7]. For the $j$-th predictor interval $[x_{ij}^l, x_{ij}^u]$, the reference point for the $i$-th observation is defined as follows:

$$p_{ij}(\lambda_j) = (1 - \lambda_j)x_{ij}^l + \lambda_j x_{ij}^u, \qquad \lambda_j \in [0, 1], \tag{2.11}$$

where $\lambda_j$ is a data-driven parameter to be estimated, which determines the relative weight assigned to the lower and upper bounds of the $j$-th predictor. Then, the PM constructs separate linear models for the lower and upper bounds of the interval response using the same set of parameterized predictor points for both bounds. Substituting the reference point definition directly into the model specification yields the following:

$$\begin{aligned} y_i^l &= \sum_{j=1}^{p} \beta_j^l p_{ij}(\lambda_j) + \varepsilon_i^l = \sum_{j=1}^{p} \beta_j^l \left[(1 - \lambda_j)x_{ij}^l + \lambda_j x_{ij}^u\right] + \varepsilon_i^l, \\ y_i^u &= \sum_{j=1}^{p} \beta_j^u p_{ij}(\lambda_j) + \varepsilon_i^u = \sum_{j=1}^{p} \beta_j^u \left[(1 - \lambda_j)x_{ij}^l + \lambda_j x_{ij}^u\right] + \varepsilon_i^u, \end{aligned} \tag{2.12}$$

where $\beta_j^l$ and $\beta_j^u$ are the regression coefficients for the $j$-th covariate in the lower and upper bound models, respectively, and the errors $\varepsilon_i^l$ and $\varepsilon_i^u$ are assumed to have mean zero and finite variances.

For notational and computational simplicity, we introduce a reparameterization by defining following:

$$\alpha_j^l = \beta_j^l(1 - \lambda_j), \quad \omega_j^l = \beta_j^l \lambda_j, \quad \alpha_j^u = \beta_j^u(1 - \lambda_j), \quad \omega_j^u = \beta_j^u \lambda_j. \tag{2.13}$$

This transforms the PM into an equivalent and more interpretable form:

$$y_i^l = \sum_{j=1}^{p} \left(\alpha_j^l x_{ij}^l + \omega_j^l x_{ij}^u\right) + \varepsilon_i^l, \quad y_i^u = \sum_{j=1}^{p} \left(\alpha_j^u x_{ij}^l + \omega_j^u x_{ij}^u\right) + \varepsilon_i^u. \tag{2.14}$$

This reparameterization reveals that the PM effectively models the response bounds as linear combinations of all predictor bounds, with coefficients implicitly constrained by the shared $\lambda_j$ parameters. These parameters are typically estimated by optimizing a goodness-of-fit criterion, thus allowing the method to adaptively extract the most informative linear combination of predictor bounds for each variable.

**Remark 1** (Particular cases of PM)**.** The PM unifies the mean structures of classical interval regression models. Specifically, the PM degenerates to the CM when the reference-point weights are fixed at $\lambda_j^l = \lambda_j^u = 0.5$ for all predictors $j$. Similarly, setting $\lambda_j^l = 0$ and $\lambda_j^u = 1$ reduces the PM to the MM.

Alternatively, the PM reduces to the CRM by imposing specific linear constraints on the parameters $\alpha_j^b$ and $\omega_j^b$ ($b \in \{l, u\}$). This is achieved by setting the lower-bound coefficients as $\alpha_j^l = \frac{\beta_j^c + \beta_j^r}{2}$ and $\omega_j^l = \frac{\beta_j^c - \beta_j^r}{2}$, and the upper-bound coefficients as $\alpha_j^u = \frac{\beta_j^c - \beta_j^r}{2}$ and $\omega_j^u = \frac{\beta_j^c + \beta_j^r}{2}$.

Building upon these foundations, this paper focuses on their heteroscedastic extensions: Heteroscedastic Min-Max Method (HMM), Heteroscedastic Center and Range Method (HCRM), Heteroscedastic Parametrized Method (HPM), and introduces the extended BHPM.

## 3. The Bayesian heteroscedastic parametrized method

Interval-valued data frequently exhibit heteroscedasticity, whereby the dispersion of the lower and upper boundaries systematically varies with the predictor levels—for instance, higher price levels in financial data are often accompanied by larger daily ranges. The parametrized representation offers a natural extension to variance modeling, as it allows the lower and upper bounds of predictors to exert distinct influences on conditional dispersion. Moreover, a unified Bayesian framework supports coherent uncertainty quantifications while ensuring the mathematical consistency of the predicted intervals. Motivated by these considerations, the extended BHPM embeds a joint mean-variance structure within the PM, thereby enabling the simultaneous capture of interval uncertainty and variance heterogeneity. In particular, the BHPM fully exploits interval information by automatically recovering data-driven reference-point weights.

Consider $n$ observations with an interval-valued response variable $\mathbf{Y}^I = \{[y_i^l, y_i^u]\}_{i=1}^n$ and $p$ interval-valued predictor variables $\{\mathbf{X}_j^I\}_{j=1}^p$, where each predictor observation is given by $[x_{ij}^l, x_{ij}^u]$ for $i = 1, \ldots, n$. Assume the errors are independent and normally distributed as follows:

$$y_i^l \sim N(\mu_i^l, (\sigma_i^l)^2), \qquad y_i^u \sim N(\mu_i^u, (\sigma_i^u)^2). \tag{3.1}$$

Following the interval parametrization approach in Section 2.2.3, for each predictor $j = 1, \ldots, p$, we introduce a value $\lambda_j^l \in [0, 1]$ that identifies a reference point inside the interval $[x_{ij}^l, x_{ij}^u]$. This reference point is expressed as follows:

$$p_{ij}(\lambda_j^l) = (1 - \lambda_j^l)x_{ij}^l + \lambda_j^l x_{ij}^u. \tag{3.2}$$

We model the conditional mean of the lower bound using the following reference point:

$$\mu_i^l = \sum_{j=1}^{p} \tilde{\beta}_j^l \, p_{ij}(\lambda_j^l), \tag{3.3}$$

where $\tilde{\beta}_j^l$ denotes the regression coefficient associated with the reference point for predictor $j$.

Substituting Eq (3.2) into Eq (3.3) immediately yields the equivalent representation:

$$\mu_i^l = \sum_{j=1}^{p} \tilde{\beta}_j^l [(1 - \lambda_j^l)x_{ij}^l + \lambda_j^l x_{ij}^u]. \tag{3.4}$$

By grouping the coefficients and the parameters $\lambda_j^l$, we define the new coefficients as follows:

$$\alpha_j^l = \tilde{\beta}_j^l(1 - \lambda_j^l), \qquad \omega_j^l = \tilde{\beta}_j^l \lambda_j^l. \tag{3.5}$$

Inserting Eq (3.5) into Eq (3.4) produces the following linear form in the original predictor bounds:

$$\mu_i^l = \sum_{j=1}^{p} (\alpha_j^l x_{ij}^l + \omega_j^l x_{ij}^u). \tag{3.6}$$

Solving Eq (3.5) for the parametrization weight (provided $\alpha_j^l + \omega_j^l \neq 0$) gives the following:

$$\lambda_j^l = \frac{\omega_j^l}{\alpha_j^l + \omega_j^l}. \tag{3.7}$$

An analogous derivation for the upper bound yields

$$\mu_i^u = \sum_{j=1}^{p} (\alpha_j^u x_{ij}^l + \omega_j^u x_{ij}^u), \tag{3.8}$$

and the corresponding weight

$$\lambda_j^u = \frac{\omega_j^u}{\alpha_j^u + \omega_j^u}. \tag{3.9}$$

In matrix notation, both conditional-mean models share the common design matrix $\mathbf{X}^{lu} \in \mathbb{R}^{n \times 2p}$ constructed from the lower and upper bounds of all predictors:

$$\mathbf{X}^{lu} = (\mathbf{X}_1^{lu}, \ldots, \mathbf{X}_n^{lu}) = \begin{pmatrix} x_{11}^l & x_{11}^u & \cdots & x_{1p}^l & x_{1p}^u \\ x_{21}^l & x_{21}^u & \cdots & x_{2p}^l & x_{2p}^u \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{n1}^l & x_{n1}^u & \cdots & x_{np}^l & x_{np}^u \end{pmatrix}. \tag{3.10}$$

Consequently, the observation models become the following:

$$\mathbf{y}^l = \mathbf{X}^{lu}\boldsymbol{\eta}^l + \boldsymbol{\varepsilon}^l, \qquad \mathbf{y}^u = \mathbf{X}^{lu}\boldsymbol{\eta}^u + \boldsymbol{\varepsilon}^u, \tag{3.11}$$

where $\boldsymbol{y}^l = (y_1^l, \ldots, y_n^l)^T$, $\boldsymbol{y}^u = (y_1^u, \ldots, y_n^u)^T$, $\boldsymbol{\eta}^l = (\alpha_1^l, \omega_1^l, \ldots, \alpha_p^l, \omega_p^l)^T$, $\boldsymbol{\eta}^u = (\alpha_1^u, \omega_1^u, \ldots, \alpha_p^u, \omega_p^u)^T$, and $\boldsymbol{\varepsilon}^l$, $\boldsymbol{\varepsilon}^u$ are the respective error vectors. Therefore, the parametrization weights $\lambda_j^l$ and $\lambda_j^u$ are directly recovered from the fitted coefficients, thereby automatically selecting the data-driven reference points that best explain each response bound.

Furthermore, if we have variance heterogeneity, it is convenient to assume an explicit variance modeling related to some explanatory variables. In this paper, we consider the scenario where the upper and lower bound data are independent; thus, the variances of the two bounds are separately modeled. Based on Eq (3.1), the variance model is as follows:

$$\log((\sigma_i^l)^2) = \sum_{k=1}^{q} \xi_k^l z_{ik}^l + \psi_k^l z_{ik}^u, \tag{3.12}$$

$$\log((\sigma_i^u)^2) = \sum_{k=1}^{q} \xi_k^u z_{ik}^l + \psi_k^u z_{ik}^u, \tag{3.13}$$

These point-valued vectors are defined as $\mathbf{Z}_i^{lu} = (z_{i1}^l, z_{i1}^u, \ldots, z_{iq}^l, z_{iq}^u)^T$, and $\mathbf{Z}^{lu} = (\mathbf{Z}_1^{lu}, \ldots, \mathbf{Z}_n^{lu})^T \in \mathbb{R}^{n\times 2q}$ denotes the matrix of point-valued covariates. Here, some components of $\mathbf{Z}_i^{lu}$ may partially overlap with those of $\mathbf{X}^{lu}$. Consequently,

$$\boldsymbol{\Sigma}^l = \operatorname{diag}\left((\sigma_1^l)^2, \ldots, (\sigma_n^l)^2\right) = \operatorname{diag}\left(\exp(\mathbf{Z}^{lu}\boldsymbol{\gamma}^l)\right), \tag{3.14}$$

$$\boldsymbol{\Sigma}^u = \operatorname{diag}\left((\sigma_1^u)^2, \ldots, (\sigma_n^u)^2\right) = \operatorname{diag}\left(\exp(\mathbf{Z}^{lu}\boldsymbol{\gamma}^u)\right), \tag{3.15}$$

where $\boldsymbol{\Sigma}^l$ and $\boldsymbol{\Sigma}^u$ are the $n \times n$ conditional variance-covariance matrices of $\boldsymbol{\varepsilon}^l$ and $\boldsymbol{\varepsilon}^u$, respectively, $\boldsymbol{\gamma}^l = (\xi_1^l, \psi_1^l, \ldots, \xi_q^l, \psi_q^l)^T$, and $\boldsymbol{\gamma}^u = (\xi_1^u, \psi_1^u, \ldots, \xi_q^u, \psi_q^u)^T$.

This joint mean-variance structure with parameterized predictors offers high flexibility, with constraints recovering the CM, MM, and CRM models.

Therefore, in this paper, we consider the following BHPM model:

$$\begin{cases} y_i^b \sim N(\mu_i^b, (\sigma_i^b)^2), \\ \mu_i^b = \mathbf{X}_i^{lu}\boldsymbol{\eta}^b, \\ \log\left((\sigma_i^b)^2\right) = (\mathbf{Z}_i^{lu})^T\boldsymbol{\gamma}^b, \\ b \in \{l, u\}. \end{cases} \tag{3.16}$$

The BHPM extends the classical parametrized representation of interval-valued data into a unified joint mean-variance framework. Unlike the CM (which discards dispersion information by collapsing intervals to midpoints) and the CRM (which models centers and ranges separately without addressing heteroscedasticity), the BHPM enables the lower and upper bounds of each predictor to exert distinct, data-driven effects on both the conditional means and log-variances of the response bounds. This structure fully exploits interval information, automatically recovers optimal reference-point weights, and supports rigorous Bayesian posterior inference.

## 4. Bayesian analysis

### *4.1. Posterior distributions*

We adopt a Bayesian inferential framework to achieve parameter estimation and rigorous uncertainty quantification for the proposed BHPM. Accordingly, we first specify conditionally prior distributions for all unknown parameters and derive the explicit likelihood function under the normal heteroscedastic error assumption. From the model Eq (3.16),

$$\mathbf{y}^b = \mathbf{X}^{lu}\boldsymbol{\eta}^b + \boldsymbol{\varepsilon}^b, \quad b \in \{l, u\}, \tag{4.1}$$

where $\boldsymbol{\varepsilon}^b \sim N(\mathbf{0}, \boldsymbol{\Sigma}^b)$ and $\boldsymbol{\Sigma}^b = \operatorname{diag}(\exp(\mathbf{Z}^{lu}\boldsymbol{\gamma}^b))$, we obtain the following likelihood function:

$$L(\boldsymbol{\eta}^b, \boldsymbol{\gamma}^b \mid \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) \propto |\boldsymbol{\Sigma}^b|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}^b - \mathbf{X}^{lu}\boldsymbol{\eta}^b)^T(\boldsymbol{\Sigma}^b)^{-1}(\mathbf{y}^b - \mathbf{X}^{lu}\boldsymbol{\eta}^b)\right\}. \tag{4.2}$$

where $\propto$ denotes proportionality, thus signifying that the two sides only differ by a constant factor independent of the model parameters.

Assuming prior independence among parameters, $p(\boldsymbol{\eta}^l, \boldsymbol{\eta}^u, \boldsymbol{\gamma}^l, \boldsymbol{\gamma}^u) = p(\boldsymbol{\eta}^l)p(\boldsymbol{\eta}^u)p(\boldsymbol{\gamma}^l)p(\boldsymbol{\gamma}^u)$, the priors are specified as $\boldsymbol{\eta}^l \sim N(\boldsymbol{\mu}_{\eta^l}, \boldsymbol{\Sigma}_{\eta^l})$, $\boldsymbol{\eta}^u \sim N(\boldsymbol{\mu}_{\eta^u}, \boldsymbol{\Sigma}_{\eta^u})$, $\boldsymbol{\gamma}^l \sim N(\boldsymbol{\mu}_{\gamma^l}, \boldsymbol{\Sigma}_{\gamma^l})$, and $\boldsymbol{\gamma}^u \sim N(\boldsymbol{\mu}_{\gamma^u}, \boldsymbol{\Sigma}_{\gamma^u})$.

**Theorem 1** (Conditional posterior distribution of $\boldsymbol{\eta}^b$)**.** For model (3.16), the conditional posterior $p(\boldsymbol{\eta}^b \mid \boldsymbol{\gamma}^b, \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu})$ follows a normal distribution $N(\boldsymbol{\mu}^*_{\eta^b}, \boldsymbol{\Sigma}^*_{\eta^b})$, where

$$\boldsymbol{\Sigma}^*_{\eta^b} = \left(\boldsymbol{\Sigma}^{-1}_{\eta^b} + (\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^b)^{-1}\mathbf{X}^{lu}\right)^{-1}, \quad \boldsymbol{\mu}^*_{\eta^b} = \boldsymbol{\Sigma}^*_{\eta^b}\left(\boldsymbol{\Sigma}^{-1}_{\eta^b}\boldsymbol{\mu}_{\eta^b} + (\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^b)^{-1}\mathbf{y}^b\right). \tag{4.3}$$

*Proof.* By Bayes' theorem, the conditional posterior distribution of $\boldsymbol{\eta}^b$ satisfies the core formula as follows:

$$p(\boldsymbol{\eta}^b \mid \boldsymbol{\gamma}^b, \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) \propto p(\mathbf{y}^b \mid \boldsymbol{\eta}^b, \boldsymbol{\gamma}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) \cdot p(\boldsymbol{\eta}^b), \tag{4.4}$$

where $p(\mathbf{y}^b \mid \cdot)$ is the likelihood function defined in Eq (4.2), and $p(\boldsymbol{\eta}^b)$ is the independent normal prior of $\boldsymbol{\eta}^b$.

Given $\boldsymbol{\gamma}^b$, the covariance matrix $\boldsymbol{\Sigma}^b = \mathrm{diag}(\exp(\mathbf{Z}^{lu}\boldsymbol{\gamma}^b))$ is known and fixed. Let $(\boldsymbol{\Sigma}^b)^{-1/2}$ denote the inverse Cholesky factor of $\boldsymbol{\Sigma}^b$. Multiplying both sides of Eq (4.1) by $(\boldsymbol{\Sigma}^b)^{-1/2}$ yields the following transformed homoscedastic model:

$$(\boldsymbol{\Sigma}^b)^{-1/2}\mathbf{y}^b \sim N\left((\boldsymbol{\Sigma}^b)^{-1/2}\mathbf{X}^{lu}\boldsymbol{\eta}^b, \mathbf{I}\right).$$

The corresponding likelihood kernel (simplified from Eq (4.2)) is as follows:

$$p(\mathbf{y}^b \mid \boldsymbol{\eta}^b, \boldsymbol{\gamma}^b, \cdot) \propto \exp\left\{-\frac{1}{2}\left((\boldsymbol{\Sigma}^b)^{-1/2}\mathbf{y}^b - (\boldsymbol{\Sigma}^b)^{-1/2}\mathbf{X}^{lu}\boldsymbol{\eta}^b\right)^T\left((\boldsymbol{\Sigma}^b)^{-1/2}\mathbf{y}^b - (\boldsymbol{\Sigma}^b)^{-1/2}\mathbf{X}^{lu}\boldsymbol{\eta}^b\right)\right\}.$$

By expanding the quadratic form inside the exponent, we get the following simplified likelihood kernel:

$$p(\mathbf{y}^b \mid \boldsymbol{\eta}^b, \boldsymbol{\gamma}^b, \cdot) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\eta}^b)^T(\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^b)^{-1}\mathbf{X}^{lu}\boldsymbol{\eta}^b + (\boldsymbol{\eta}^b)^T(\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^b)^{-1}\mathbf{y}^b\right\}.$$

The kernel of the normal prior for $\boldsymbol{\eta}^b$ is as follows:

$$p(\boldsymbol{\eta}^b) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\eta}^b - \boldsymbol{\mu}_{\eta^b})^T\boldsymbol{\Sigma}^{-1}_{\eta^b}(\boldsymbol{\eta}^b - \boldsymbol{\mu}_{\eta^b})\right\}.$$

By substituting the likelihood kernel and prior kernel into the Bayes' formula Eq (4.4), the unnormalized posterior kernel is the product of the two:

$$p(\boldsymbol{\eta}^b \mid \boldsymbol{\gamma}^b, \cdot) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\eta}^b)^T\left[(\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^b)^{-1}\mathbf{X}^{lu} + \boldsymbol{\Sigma}^{-1}_{\eta^b}\right]\boldsymbol{\eta}^b + (\boldsymbol{\eta}^b)^T\left[(\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^b)^{-1}\mathbf{y}^b + \boldsymbol{\Sigma}^{-1}_{\eta^b}\boldsymbol{\mu}_{\eta^b}\right]\right\}.$$

Note that the exponent of a normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ has the following canonical form:

$$-\frac{1}{2}\boldsymbol{\theta}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta} + \boldsymbol{\theta}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \text{constant},$$

where $\boldsymbol{\Sigma}^{-1}$ is the precision matrix, and $\boldsymbol{\mu}$ is the mean vector.

Define the posterior precision matrix

$$\boldsymbol{\Sigma}^{*-1}_{\eta^b} = (\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^b)^{-1}\mathbf{X}^{lu} + \boldsymbol{\Sigma}^{-1}_{\eta^b},$$

and the corresponding posterior mean

$$\boldsymbol{\mu}_{\eta^b}^* = \boldsymbol{\Sigma}_{\eta^b}^* \left( \boldsymbol{\Sigma}_{\eta^b}^{-1} \boldsymbol{\mu}_{\eta^b} + (\mathbf{X}^{lu})^T (\boldsymbol{\Sigma}^b)^{-1} \mathbf{y}^b \right),$$

with $\boldsymbol{\Sigma}_{\eta^b}^* = \left( \boldsymbol{\Sigma}_{\eta^b}^{*-1} \right)^{-1}$.

Completing the square in the exponent of the posterior kernel confirms that the conditional posterior of $\boldsymbol{\eta}^b$ follows a normal distribution as follows:

$$\boldsymbol{\eta}^b \mid \boldsymbol{\gamma}^b, \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu} \sim N\left( \boldsymbol{\mu}_{\eta^b}^*, \boldsymbol{\Sigma}_{\eta^b}^* \right),$$

with the mean and covariance matrix as defined above. □

**Theorem 2** (Conditional posterior distribution of $\boldsymbol{\gamma}^b$)**.** For model (3.16), the conditional posterior $p(\boldsymbol{\gamma}^b \mid \boldsymbol{\eta}^b, \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu})$ is proportional to the following:

$$|\boldsymbol{\Sigma}^b|^{-\frac{1}{2}} \exp\left( -\frac{1}{2} (\mathbf{e}^b)^T (\boldsymbol{\Sigma}^b)^{-1} \mathbf{e}^b \right) \exp\left( -\frac{1}{2} (\boldsymbol{\gamma}^b - \boldsymbol{\mu}_{\gamma^b})^T \boldsymbol{\Sigma}_{\gamma^b}^{-1} (\boldsymbol{\gamma}^b - \boldsymbol{\mu}_{\gamma^b}) \right),$$

where $\mathbf{e}^b = \mathbf{y}^b - \mathbf{X}^{lu} \boldsymbol{\eta}^b$.

*Proof.* By Bayes' theorem, the conditional posterior distribution of $\boldsymbol{\gamma}^b$ satisfies the following core formula:

$$p(\boldsymbol{\gamma}^b \mid \boldsymbol{\eta}^b, \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) \propto p(\mathbf{y}^b \mid \boldsymbol{\eta}^b, \boldsymbol{\gamma}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) \cdot p(\boldsymbol{\gamma}^b), \tag{4.5}$$

where $p(\mathbf{y}^b \mid \cdot)$ is the likelihood function defined in Eq (4.2), and $p(\boldsymbol{\gamma}^b)$ is the independent normal prior of $\boldsymbol{\gamma}^b$.

Given $\boldsymbol{\eta}^b$, the residuals $\mathbf{e}^b = \mathbf{y}^b - \mathbf{X}^{lu} \boldsymbol{\eta}^b$ are fixed and known (directly derived from Eq (4.1)). By substituting the residuals into the likelihood function Eq (4.2), the likelihood kernel simplifies to the following:

$$p(\mathbf{y}^b \mid \boldsymbol{\eta}^b, \boldsymbol{\gamma}^b, \cdot) \propto |\boldsymbol{\Sigma}^b|^{-\frac{1}{2}} \exp\left( -\frac{1}{2} (\mathbf{e}^b)^T (\boldsymbol{\Sigma}^b)^{-1} \mathbf{e}^b \right),$$

where $\boldsymbol{\Sigma}^b = \mathrm{diag}(\exp(\mathbf{Z}^{lu} \boldsymbol{\gamma}^b))$ is fully determined by $\boldsymbol{\gamma}^b$.

The kernel of the normal prior for $\boldsymbol{\gamma}^b$ is as follows:

$$p(\boldsymbol{\gamma}^b) \propto \exp\left( -\frac{1}{2} (\boldsymbol{\gamma}^b - \boldsymbol{\mu}_{\gamma^b})^T \boldsymbol{\Sigma}_{\gamma^b}^{-1} (\boldsymbol{\gamma}^b - \boldsymbol{\mu}_{\gamma^b}) \right).$$

By substituting the likelihood kernel and prior kernel into the Bayes' formula Eq (4.5), we obtain the unnormalized conditional posterior kernel as follows:

$$p(\boldsymbol{\gamma}^b \mid \boldsymbol{\eta}^b, \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) \propto |\boldsymbol{\Sigma}^b|^{-\frac{1}{2}} \exp\left( -\frac{1}{2} (\mathbf{e}^b)^T (\boldsymbol{\Sigma}^b)^{-1} \mathbf{e}^b \right) \exp\left( -\frac{1}{2} (\boldsymbol{\gamma}^b - \boldsymbol{\mu}_{\gamma^b})^T \boldsymbol{\Sigma}_{\gamma^b}^{-1} (\boldsymbol{\gamma}^b - \boldsymbol{\mu}_{\gamma^b}) \right),$$

which exactly matches the form stated in the proposition. Note that this posterior is non-conjugate due to the non-linear dependence of $\boldsymbol{\Sigma}^b$ on $\boldsymbol{\gamma}^b$ via the logarithmic link, and thus requires numerical sampling methods (e.g., Metropolis-Hastings algorithm) for posterior inference. □

### 4.2. Gibbs-MH sampling procedure

Given the observed data $\mathbf{y}^b$, design matrix $\mathbf{X}^{lu}$, and covariate matrix $\mathbf{Z}^{lu}$, the posterior distribution for the model parameters $(\boldsymbol{\eta}^b, \boldsymbol{\gamma}^b)$ is as follows:

$$p(\boldsymbol{\eta}^b, \boldsymbol{\gamma}^b \mid \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) \propto L(\boldsymbol{\eta}^b, \boldsymbol{\gamma}^b \mid \mathbf{y}^b, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) \times p(\boldsymbol{\eta}^b)p(\boldsymbol{\gamma}^b), \quad b \in \{l, u\},$$

where we can sample from the joint posterior distribution by Gibbs sampling along the following process.

**Iterative Sampling Procedure**

**Step 1. Initialization:** Set initial values $\boldsymbol{\eta}^{b(0)}$, $\boldsymbol{\gamma}^{b(0)}$.

**Step 2.** For each iteration $t = 1, 2, \ldots, T$, perform the following steps:

(a) **Update $\boldsymbol{\eta}^{b(t)}$ (conditioned on $\boldsymbol{\gamma}^{b(t-1)}$):**

$$\boldsymbol{\eta}^{b(t)} \mid \boldsymbol{\gamma}^{b(t-1)}, \mathbf{y}^b \sim N\left(\boldsymbol{\mu}^*_{\eta^b}, \boldsymbol{\Sigma}^*_{\eta^b}\right),$$

where

$$\boldsymbol{\Sigma}^*_{\eta^b} = \left(\boldsymbol{\Sigma}^{-1}_{\eta^b} + (\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^{b(t-1)})^{-1}\mathbf{X}^{lu}\right)^{-1}, \tag{4.6}$$

$$\boldsymbol{\mu}^*_{\eta^b} = \boldsymbol{\Sigma}^*_{\eta^b}\left(\boldsymbol{\Sigma}^{-1}_{\eta^b}\boldsymbol{\mu}_{\eta^b} + (\mathbf{X}^{lu})^T(\boldsymbol{\Sigma}^{b(t-1)})^{-1}\mathbf{y}^b\right), \tag{4.7}$$

$$\boldsymbol{\Sigma}^{b(t-1)} = \mathrm{diag}(\exp(\mathbf{Z}^{lu}\boldsymbol{\gamma}^{b(t-1)})). \tag{4.8}$$

The sampling of $\boldsymbol{\eta}^b$ is done using a normal distribution in each iteration, where the current $\boldsymbol{\gamma}^{b(t-1)}$ is used to update $\boldsymbol{\eta}^{b(t)}$.

(b) **Update $\boldsymbol{\gamma}^{b(t)}$ (conditioned on $\boldsymbol{\eta}^{b(t)}$):**
Given the current value $\boldsymbol{\eta}^{b(t)}$, the posterior distribution for $\boldsymbol{\gamma}^{b(t)}$ is as follows:

$$p(\boldsymbol{\gamma}^b \mid \boldsymbol{\eta}^{b(t)}, \mathbf{y}^b) \propto |\boldsymbol{\Sigma}^b|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}^b - \mathbf{X}^{lu}\boldsymbol{\eta}^{b(t)})^T(\boldsymbol{\Sigma}^b)^{-1}(\mathbf{y}^b - \mathbf{X}^{lu}\boldsymbol{\eta}^{b(t)})\right\} \times p(\boldsymbol{\gamma}^b).$$

Since this posterior distribution is complex, we use the Metropolis-Hastings algorithm for sampling. The specific steps are as follows:

- **Proposal distribution:** A normal distribution $N(\boldsymbol{\gamma}^{b(t-1)}, \sigma^2_\gamma \mathbf{V}^{-1}_{\gamma^b})$ is used as the proposal distribution([27,28]), where $\mathbf{V}_{\gamma^b}$ is a pre-specified or adaptively tuned covariance matrix, $\sigma^2_\gamma$ is a scalar adjustment parameter, and $\sigma^2_\gamma$ is a scalar adjustment parameter chosen such that the average acceptance rate is about between 0.25 and 0.45 ( [29]).
- **Acceptance probability:** For the current candidate $\boldsymbol{\gamma}^{b*}$, the acceptance probability $\alpha$ is calculated as follows:

$$\alpha = \min\left\{1, \ \frac{p(\boldsymbol{\gamma}^{b*} \mid \boldsymbol{\eta}^{b(t)}, \mathbf{y}^b)}{p(\boldsymbol{\gamma}^{b(t-1)} \mid \boldsymbol{\eta}^{b(t)}, \mathbf{y}^b)}\right\}.$$

This acceptance probability compares the posterior density ratio between the current candidate $\boldsymbol{\gamma}^{b*}$ and the previous value $\boldsymbol{\gamma}^{b(t-1)}$. If the ratio is greater than 1, then the candidate $\boldsymbol{\gamma}^{b*}$ is accepted. Otherwise, the acceptance is based on this ratio. If accepted, then set $\boldsymbol{\gamma}^{b(t)} = \boldsymbol{\gamma}^{b*}$. If rejected, then keep $\boldsymbol{\gamma}^{b(t)} = \boldsymbol{\gamma}^{b(t-1)}$.

**Step 3. Repeat Step 2:** Repeat the process from Step 2 for $T$ iterations, continuously updating $\boldsymbol{\eta}^{b(t)}$ and $\boldsymbol{\gamma}^{b(t)}$. This iterative procedure allows the parameters to gradually converge to their posterior distribution.

### 4.3. Bayesian inference

MCMC samples $\{\boldsymbol{\theta}^{(t)} = (\boldsymbol{\eta}^{l(t)}, \boldsymbol{\gamma}^{l(t)}, \boldsymbol{\eta}^{u(t)}, \boldsymbol{\gamma}^{u(t)}) : t = B+1, \ldots, T\}$ (after discarding burn-in period of length $B$ and optional thinning) are generated from the joint posterior via the above hybrid algorithm. The Bayesian point estimates are the following posterior means:

$$\hat{\boldsymbol{\eta}}^l = \frac{1}{T-B}\sum_{t=B+1}^{T} \boldsymbol{\eta}^{l(t)}, \quad \hat{\boldsymbol{\eta}}^u = \frac{1}{T-B}\sum_{t=B+1}^{T} \boldsymbol{\eta}^{u(t)}, \tag{4.9}$$

$$\hat{\boldsymbol{\gamma}}^l = \frac{1}{T-B}\sum_{t=B+1}^{T} \boldsymbol{\gamma}^{l(t)}, \quad \hat{\boldsymbol{\gamma}}^u = \frac{1}{T-B}\sum_{t=B+1}^{T} \boldsymbol{\gamma}^{u(t)}. \tag{4.10}$$

As is shown by Geyer [30], these posterior means are consistent estimators of the corresponding posterior mean vector as $T$ goes to infinity.

The posterior covariance matrix is estimated by the sample covariance of the retained draws as follows:

$$\widehat{\mathrm{Cov}}(\boldsymbol{\theta} \mid \mathbf{y}^l, \mathbf{y}^u, \mathbf{X}^{lu}, \mathbf{Z}^{lu}) = \frac{1}{T-B-1}\sum_{t=B+1}^{T} \left(\boldsymbol{\theta}^{(t)} - \hat{\boldsymbol{\theta}}\right)\left(\boldsymbol{\theta}^{(t)} - \hat{\boldsymbol{\theta}}\right)^T.$$

The posterior standard deviations are the square roots of the diagonal elements.

With the obtained point estimates of $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$, the predicted lower and upper bounds of the response variable can be directly calculated via the BHPM model specified in Eq (3.16). For any sample unit $i$ to be predicted, the point predictions of the lower and upper bounds (i.e., the estimated conditional means of the response bounds) are given by the following:

$$\hat{y}_i^l = \mathbf{X}_i^{lu}\hat{\boldsymbol{\eta}}^l, \tag{4.11}$$

$$\hat{y}_i^u = \mathbf{X}_i^{lu}\hat{\boldsymbol{\eta}}^u, \tag{4.12}$$

where $\mathbf{X}_i^{lu}$ is the row vector of mean-model covariates that correspond to sample $i$, which is consistent with the covariate structure used in the model fitting.

## 5. Simulation study

This section presents a simulation study to validate the effectiveness of the Bayesian method in estimating the parameters of the BHPM, and compares the BHPM with the HPM, HMM, and HCRM. All simulations are implemented using the R statistical software environment. Section 5.1 designs three different prior scenarios for the parameter estimation. Section 5.2 presents the results of the model comparison.

### 5.1. Parameter estimation of the BHPM

In this section, the model structure is identified $\mathbf{y}^b = \mathbf{X}^{lu}\boldsymbol{\eta}^b + \boldsymbol{\varepsilon}^b$ according to Eq (4.1). When $p = 3$, the parameter settings for data generation are as follows:

1) First, we generate $x_{ij}^c \sim U[5, 10]$, $x_{ij}^r \sim U[4, 7] (i = 1, 2, \ldots, n)$.

2) Then, we calculate $x_{ij}^l = x_{ij}^c - x_{ij}^r$ and $x_{ij}^u = x_{ij}^c + x_{ij}^r$, respectively.

3) Set the regression parameters $\boldsymbol{\eta}^l = (6, 3, 3, 2, 4, 6)^T$ and $\boldsymbol{\eta}^u = (2, 3, 4, 6, 6, 9)^T$.

4) The structure of the variance model is $\log((\sigma_i^b)^2) = (\mathbf{Z}_i^{lu})^T \boldsymbol{\gamma}^b$. Set the variance parameters $\boldsymbol{\gamma}^l = (0.2, -0.1, 0.1, -0.2)^T$ and $\boldsymbol{\gamma}^u = (0.3, -0.2, 0.2, -0.3)^T$ to attain heteroscedasticity.

5) Specifically, we construct $\mathbf{Z}_i^{lu}$ by extracting the interval bounds of the first two covariates from $\mathbf{X}_i^{lu}$. Formally, we define the $4 \times 1$ vector as follows: $\mathbf{Z}_i^{lu} = (x_{i1}^l, x_{i1}^u, x_{i2}^l, x_{i2}^u)^T$.

To investigate the sensitivity of a Bayesian estimation to the prior distributions, three scenarios of prior distributions for the unknown parameters in the upper-bound and lower-bound models are set in Table 1. Specifically, Type I represents the case with no prior information, Type II represents the case with good prior information, and Type III represents the case with imprecise prior information. For example, Figure 1 displays a scatter plot of interval-valued symbolic data of variable $X_1$, with only 50 data points presented.

**Table 1.** Prior settings of parameters.

| Model | Type | Parameters | | | |
|---|---|---|---|---|---|
| Upper bound | I | $\boldsymbol{\mu}_{\eta^u} = (0, 0, 0, 0, 0, 0)^T$ | $\boldsymbol{\Sigma}_{\eta^u} = \mathbf{I}_6$ | $\boldsymbol{\gamma}^u = (0, 0, 0, 0)^T$ | $\boldsymbol{\Sigma}_{\gamma^u} = \mathbf{I}_4$ |
| | II | $\boldsymbol{\mu}_{\eta^u} = (2, 3, 4, 6, 6, 9)^T$ | $\boldsymbol{\Sigma}_{\eta^u} = \mathbf{I}_6$ | $\boldsymbol{\gamma}^u = (0.3, -0.2, 0.2, -0.3)^T$ | $\boldsymbol{\Sigma}_{\gamma^u} = \mathbf{I}_4$ |
| | III | $\boldsymbol{\mu}_{\eta^u} = 3 \times (2, 3, 4, 6, 6, 9)^T$ | $\boldsymbol{\Sigma}_{\eta^u} = \mathbf{I}_6$ | $\boldsymbol{\gamma}^u = 3 \times (0.3, -0.2, 0.2, -0.3)^T$ | $\boldsymbol{\Sigma}_{\gamma^u} = \mathbf{I}_4$ |
| Lower bound | I | $\boldsymbol{\mu}_{\eta^l} = (0, 0, 0, 0, 0, 0)^T$ | $\boldsymbol{\Sigma}_{\eta^l} = \mathbf{I}_6$ | $\boldsymbol{\gamma}^l = (0, 0, 0, 0)^T$ | $\boldsymbol{\Sigma}_{\gamma^l} = \mathbf{I}_4$ |
| | II | $\boldsymbol{\mu}_{\eta^l} = (6, 3, 3, 2, 4, 6)^T$ | $\boldsymbol{\Sigma}_{\eta^l} = \mathbf{I}_6$ | $\boldsymbol{\gamma}^l = (0.2, -0.1, 0.1, -0.2)^T$ | $\boldsymbol{\Sigma}_{\gamma^l} = \mathbf{I}_4$ |
| | III | $\boldsymbol{\mu}_{\eta^l} = 3 \times (6, 3, 3, 2, 4, 6)^T$ | $\boldsymbol{\Sigma}_{\eta^l} = \mathbf{I}_6$ | $\boldsymbol{\gamma}^l = 3 \times (0.2, -0.1, 0.1, -0.2)^T$ | $\boldsymbol{\Sigma}_{\gamma^l} = \mathbf{I}_4$ |

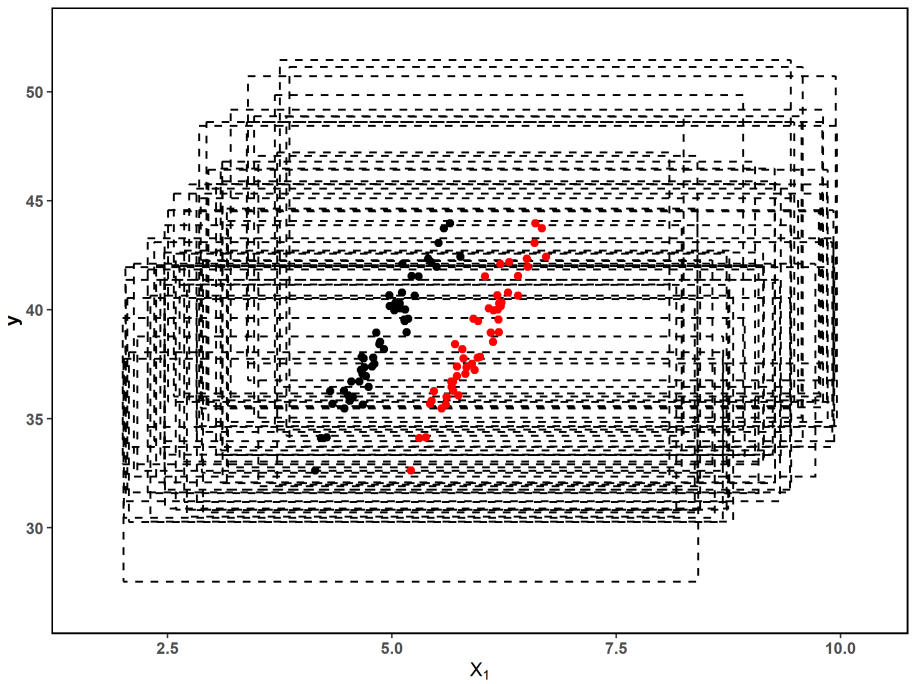


**Figure 1.** Scatter plot of interval-valued symbolic data with $p = 1$ (Black points: interval endpoints driven by $\lambda$; Red points: interval centers).

Under three different prior scenarios, we employ a hybrid algorithm that combines joint Gibbs sampling and the MH algorithm (in Section 4.2) to estimate the unknown parameters. In the simulation, we conduct experiments with sample sizes $n = 100$ and $n = 200$. For each prior scenario, the computations are repeated 100 times. Figure 2 presents the convergence plots of partial parameters in the upper bound model. Accordingly, we run the algorithm for 5000 iterations in each computation, with a burn-in period of 3000 iterations. Tables 2 and 3 present the Bayesian estimation results for the unknown parameters of the upper-bound and lower-bound models under different prior distributions, respectively. Here, Bias denotes the difference between the estimated value and the true value averaged over 100 replications, the standard deviation (SD) represents the average estimate of the posterior standard error, and the root mean square error (RMS) is the root mean square error of the Bayesian estimates across 100 replications.

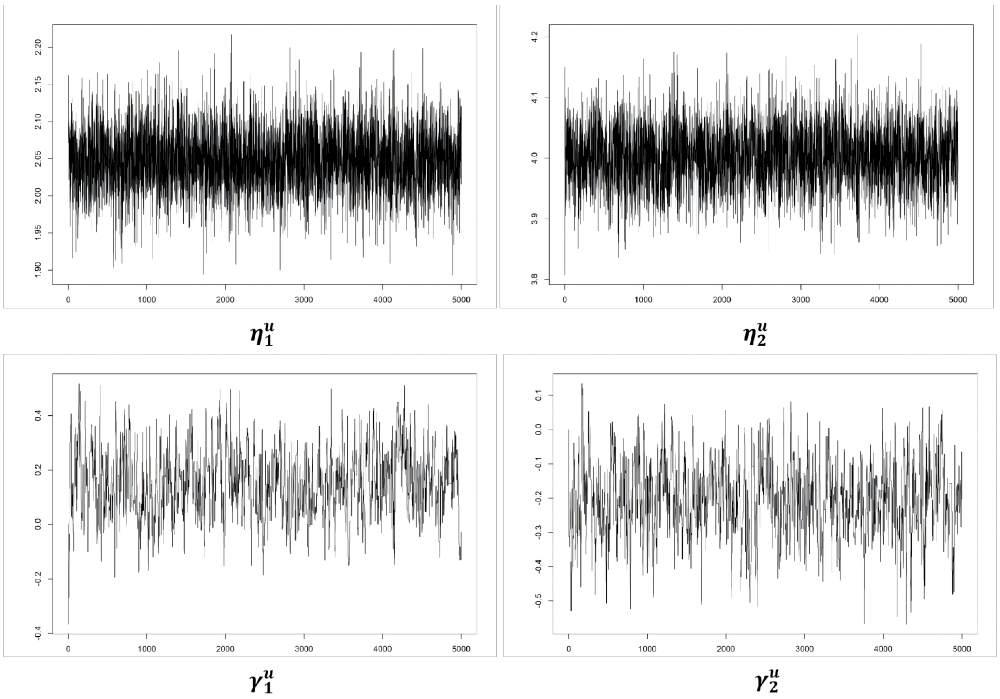


**Figure 2.** Iterative convergence plots of partial parameters in the upper-bound model with $n = 100$.

.

Based on the Bias, SD, and RMS values, the Bayesian method yields highly precise parameter estimates under all prior scenarios. The following conclusions can be drawn from the tables: (i) as the sample size $n$ increases from 100 to 200, the estimation accuracy of parameters shows a clear improvement in most cases (e.g., under the Type I prior, the SD of $\gamma_1^u$ in the upper bound model decreases from 0.1292 to 0.0814, and the RMS decreases from 0.1298 to 0.0814); (ii) the Bayesian estimation results exhibit little sensitivity to prior specifications, with no strong dependence on the choice of prior distributions; and (iii) the RMS of all parameters are less than 0.15, and the SD of the heteroscedasticity parameters $\gamma$ are less than 0.15. The estimation bias tends to decrease with an increasing sample size in most circumstances, which verifies the rationality and effectiveness of the extended method. Overall, the simulation results confirm the effectiveness of the Bayesian method to

**Table 2.** Parameter estimation of the upper bound model.

| Type | Par. | $n$ = 100 | | | $n$ = 200 | | |
|---|---|---|---|---|---|---|---|
| | | Bias | SD | RMS | Bias | SD | RMS |
| I | $\eta_1^u$ | -0.0075 | 0.0390 | 0.0397 | -0.0040 | 0.0245 | 0.0248 |
| | $\eta_2^u$ | -0.0065 | 0.0406 | 0.0411 | -0.0045 | 0.0288 | 0.0291 |
| | $\eta_3^u$ | -0.0030 | 0.0354 | 0.0355 | 0.0001 | 0.0243 | 0.0243 |
| | $\eta_4^u$ | -0.0012 | 0.0346 | 0.0347 | 0.0002 | 0.0216 | 0.0216 |
| | $\eta_5^u$ | 0.0022 | 0.0311 | 0.0312 | 0.0037 | 0.0221 | 0.0224 |
| | $\eta_6^u$ | 0.0016 | 0.0319 | 0.0319 | -0.0031 | 0.0212 | 0.0215 |
| | $\gamma_1^u$ | 0.0130 | 0.1292 | 0.1298 | -0.0029 | 0.0814 | 0.0814 |
| | $\gamma_2^u$ | 0.0125 | 0.1309 | 0.1315 | 0.0107 | 0.0773 | 0.0781 |
| | $\gamma_3^u$ | 0.0018 | 0.0953 | 0.0954 | 0.0045 | 0.0637 | 0.0639 |
| | $\gamma_4^u$ | -0.0104 | 0.0984 | 0.0989 | -0.0088 | 0.0629 | 0.0636 |
| II | $\eta_1^u$ | 0.0011 | 0.0356 | 0.0356 | 0.0022 | 0.0239 | 0.0240 |
| | $\eta_2^u$ | -0.0018 | 0.0363 | 0.0364 | -0.0019 | 0.0258 | 0.0259 |
| | $\eta_3^u$ | -0.0013 | 0.0337 | 0.0337 | -0.0007 | 0.0254 | 0.0254 |
| | $\eta_4^u$ | -0.0009 | 0.0354 | 0.0354 | 0.0001 | 0.0244 | 0.0244 |
| | $\eta_5^u$ | -0.0010 | 0.0280 | 0.0280 | -0.0026 | 0.0222 | 0.0223 |
| | $\eta_6^u$ | 0.0028 | 0.0319 | 0.0320 | 0.0034 | 0.0215 | 0.0218 |
| | $\gamma_1^u$ | 0.0066 | 0.1125 | 0.1127 | -0.0087 | 0.0881 | 0.0886 |
| | $\gamma_2^u$ | 0.0128 | 0.1224 | 0.1231 | -0.0120 | 0.0790 | 0.0799 |
| | $\gamma_3^u$ | 0.0005 | 0.0898 | 0.0898 | -0.0090 | 0.0636 | 0.0642 |
| | $\gamma_4^u$ | -0.0098 | 0.0930 | 0.0936 | 0.0087 | 0.0605 | 0.0611 |
| III | $\eta_1^u$ | 0.0016 | 0.0383 | 0.0384 | 0.0002 | 0.0224 | 0.0224 |
| | $\eta_2^u$ | 0.0037 | 0.0350 | 0.0352 | 0.0021 | 0.0285 | 0.0286 |
| | $\eta_3^u$ | 0.0030 | 0.0406 | 0.0407 | 0.0075 | 0.0276 | 0.0286 |
| | $\eta_4^u$ | -0.0066 | 0.0334 | 0.0340 | -0.0032 | 0.0250 | 0.0252 |
| | $\eta_5^u$ | -0.0027 | 0.0324 | 0.0325 | -0.0019 | 0.0212 | 0.0213 |
| | $\eta_6^u$ | 0.0080 | 0.0357 | 0.0366 | 0.0030 | 0.0228 | 0.0230 |
| | $\gamma_1^u$ | 0.0019 | 0.1334 | 0.1334 | 0.0003 | 0.0661 | 0.0661 |
| | $\gamma_2^u$ | 0.0233 | 0.1208 | 0.1230 | 0.0013 | 0.0812 | 0.0812 |
| | $\gamma_3^u$ | -0.0064 | 0.0959 | 0.0962 | 0.0095 | 0.0538 | 0.0546 |
| | $\gamma_4^u$ | 0.0010 | 0.0967 | 0.0967 | -0.0112 | 0.0561 | 0.0572 |

**Table 3.** Parameter estimation of the lower bound model.

| Type | Par. | $n$ = 100 | | | $n$ = 200 | | |
|---|---|---|---|---|---|---|---|
| | | Bias | SD | RMS | Bias | SD | RMS |
| I | $\eta_1^l$ | -0.0087 | 0.0420 | 0.0429 | -0.0025 | 0.0263 | 0.0264 |
| | $\eta_2^l$ | -0.0007 | 0.0389 | 0.0389 | -0.0065 | 0.0297 | 0.0304 |
| | $\eta_3^l$ | -0.0043 | 0.0460 | 0.0462 | -0.0012 | 0.0267 | 0.0267 |
| | $\eta_4^l$ | 0.0047 | 0.0372 | 0.0374 | 0.0016 | 0.0212 | 0.0213 |
| | $\eta_5^l$ | 0.0018 | 0.0377 | 0.0377 | -0.0001 | 0.0251 | 0.0251 |
| | $\eta_6^l$ | -0.0056 | 0.0352 | 0.0356 | 0.0008 | 0.0235 | 0.0235 |
| | $\gamma_1^l$ | 0.0239 | 0.1129 | 0.1154 | -0.0006 | 0.0788 | 0.0788 |
| | $\gamma_2^l$ | -0.0011 | 0.0987 | 0.0987 | 0.0114 | 0.0715 | 0.0724 |
| | $\gamma_3^l$ | -0.0195 | 0.0874 | 0.0896 | 0.0107 | 0.0592 | 0.0602 |
| | $\gamma_4^l$ | 0.0116 | 0.0870 | 0.0878 | -0.0148 | 0.0593 | 0.0611 |
| II | $\eta_1^l$ | -0.0018 | 0.0422 | 0.0422 | -0.0008 | 0.0296 | 0.0296 |
| | $\eta_2^l$ | -0.0042 | 0.0414 | 0.0416 | 0.0043 | 0.0280 | 0.0284 |
| | $\eta_3^l$ | -0.0030 | 0.0451 | 0.0452 | -0.0033 | 0.0310 | 0.0312 |
| | $\eta_4^l$ | -0.0016 | 0.0359 | 0.0359 | -0.0018 | 0.0244 | 0.0244 |
| | $\eta_5^l$ | -0.0002 | 0.0348 | 0.0348 | -0.0019 | 0.0248 | 0.0249 |
| | $\eta_6^l$ | 0.0037 | 0.0328 | 0.0330 | 0.0036 | 0.0234 | 0.0237 |
| | $\gamma_1^l$ | 0.0152 | 0.1035 | 0.1046 | 0.0038 | 0.0732 | 0.0733 |
| | $\gamma_2^l$ | -0.0068 | 0.1207 | 0.1209 | -0.0064 | 0.0868 | 0.0871 |
| | $\gamma_3^l$ | -0.0182 | 0.0919 | 0.0936 | -0.0031 | 0.0602 | 0.0603 |
| | $\gamma_4^l$ | 0.0136 | 0.0924 | 0.0934 | 0.0021 | 0.0584 | 0.0585 |
| III | $\eta_1^l$ | 0.0137 | 0.0447 | 0.0468 | 0.0066 | 0.0278 | 0.0286 |
| | $\eta_2^l$ | 0.0125 | 0.0468 | 0.0484 | 0.0046 | 0.0267 | 0.0271 |
| | $\eta_3^l$ | 0.0067 | 0.0378 | 0.0384 | 0.0015 | 0.0335 | 0.0335 |
| | $\eta_4^l$ | -0.0099 | 0.0332 | 0.0347 | -0.0024 | 0.0233 | 0.0235 |
| | $\eta_5^l$ | -0.0024 | 0.0346 | 0.0346 | -0.0041 | 0.0242 | 0.0245 |
| | $\eta_6^l$ | 0.0075 | 0.0338 | 0.0346 | 0.0047 | 0.0238 | 0.0242 |
| | $\gamma_1^l$ | -0.0103 | 0.1155 | 0.1159 | 0.0047 | 0.0728 | 0.0730 |
| | $\gamma_2^l$ | -0.0178 | 0.1203 | 0.1216 | -0.0029 | 0.0756 | 0.0756 |
| | $\gamma_3^l$ | 0.0131 | 0.0815 | 0.0825 | 0.0006 | 0.0586 | 0.0586 |
| | $\gamma_4^l$ | -0.0138 | 0.0814 | 0.0826 | -0.0033 | 0.0607 | 0.0608 |

estimate the unknown parameters of the BHPM.

### *5.2. Model comparison*

In this subsection, we conduct two synthetic simulation experiments based on the PM and CRM models. According to the types of interval-valued symbolic data generated in the simulation process and the magnitude of the dataset variance, we design four sets of simulation studies with sample sizes $n = 100, 200$, and $p = 3$.

**Simulation 1 (Configurations based on PM).** Two configurations, C1 and C2, are generated by the following steps ($i = 1, 2, \ldots, n$; $j = 1, 2, 3$):

1) First, we generate $x_{ij}^l \sim U[2, 4]$, $x_{ij}^u \sim U[8, 10]$.

2) We set the regression coefficients as fixed constant vectors throughout the simulation:$\boldsymbol{\alpha^l} = (6, 3, 4)^T$, $\boldsymbol{\omega^l} = (3, 2, 6)^T$, $\boldsymbol{\alpha^u} = (2, 4, 6)^T$ , and $\boldsymbol{\omega^u} = (3, 6, 9)^T$.

3) The response variables $y_i^l$ and $y_i^u$ are calculated by the following:

$$\begin{aligned} y_i^l &= \mathbf{x}_i^l \boldsymbol{\alpha}^l + \mathbf{x}_i^u \boldsymbol{\omega}^l + \epsilon_i^l, \\ y_i^u &= \mathbf{x}_i^l \boldsymbol{\alpha}^u + \mathbf{x}_i^u \boldsymbol{\omega}^u + \epsilon_i^u, \end{aligned}$$

where $\epsilon_i^b \sim N(0, (\sigma_i^b)^2)$, and $\log\left((\sigma_i^b)^2\right) = \mathbf{Z}_i^{lu} \boldsymbol{\gamma}^b$.

4) We set the heteroscedasticity parameters as fixed constant vectors:

$$\begin{aligned} \boldsymbol{\gamma}^l &= (0.2,\ -0.1,\ 0.1,\ -0.2), \\ \boldsymbol{\gamma}^u &= (0.3,\ -0.2,\ 0.2,\ -0.3). \end{aligned}$$

5) For the configuration C1, we use the heteroscedasticity parameters $\boldsymbol{\gamma}^l$ and $\boldsymbol{\gamma}^u$ defined in Step 4. For the configuration C2, we scale the parameters to $3\boldsymbol{\gamma}^l$ and $3\boldsymbol{\gamma}^u$ to vary the dataset variance.

6) We generate a total of $n$ samples. Then, we choose $4n/5$ data as a training set and the remaining $n/5$ data as a test set.

**Simulation 2 (Configurations based on CRM).** Two configurations, C3 and C4, are generated by the following steps ($i = 1, 2, \ldots, n$; $j = 1, 2, 3$):

1) First, we generate $x_{ij}^c \sim U[5, 10]$, $x_{ij}^r \sim U[4, 7]$.

2) We set the regression coefficients as fixed constant vectors throughout the simulation: $\boldsymbol{\beta^c} = (3, 5, 7)^T$ and $\boldsymbol{\beta^r} = (2, 1.5, 1)^T$ .

3) The response variables $y_i^c$ and $y_i^r$ are calculated by the following:

$$\begin{aligned} y_i^c &= \beta_1^c x_{i1}^c + \beta_2^c x_{i2}^c + \beta_3^c x_{i3}^c + \epsilon_i^c, \\ y_i^r &= \beta_1^r x_{i1}^r + \beta_2^r x_{i2}^r + \beta_3^r x_{i3}^r + \epsilon_i^r, \end{aligned}$$

where $\epsilon_i^s \sim N(0, (\sigma_i^s)^2)$, and $\log\left((\sigma_i^s)^2\right) = \mathbf{Z}_i^{lu} \boldsymbol{\gamma}^s$ $(s \in \{c, r\})$.

4) We set the heteroscedasticity parameters as fixed constant vectors:

$$\boldsymbol{\gamma}^c = (0.3,\ -0.2,\ 0.2,\ -0.3),$$
$$\boldsymbol{\gamma}^u = (0.2,\ -0.1,\ 0.1,\ -0.2).$$

5) For the configuration C3, we use the heteroscedasticity parameters $\boldsymbol{\gamma}^c$ and $\boldsymbol{\gamma}^r$ defined in Step 4. For the configuration C4, we scale the parameters to $3\boldsymbol{\gamma}^c$ and $3\boldsymbol{\gamma}^r$ to vary the dataset variance.

6) A set of $n$ samples are generated. We choose $4n/5$ data as a training set and the remaining $n/5$ data as a test set.

We define the generated interval-valued dependent variable $Y^I$ by the samples $\{[y_i^l, y_i^u]\}_{i=1}^n$, where $y_i^l = y_i^c - y_i^r$ and $y_i^u = y_i^c + y_i^r$. Similarly, we define the first interval-valued independent variable $\{X_j^I\}_{j=1}^3$ by the samples $[x_{ij}^l, x_{ij}^u]$ for $i = 1, 2, \ldots, n$, where $x_{ij}^l = x_{ij}^c - x_{ij}^r$ and $x_{ij}^u = x_{ij}^c + x_{ij}^r$. The detailed settings of the four configurations are presented in Table 4.

**Table 4.** Interval-valued data configurations C1–C4 ($i = 1, \ldots, n$, $j = 1, 2, 3$).

| | | | | |
|---|---|---|---|---|
| C1 | $x_{ij}^l \sim U[2,4]$ | $x_{ij}^u \sim U[8,10]$ | $\boldsymbol{\gamma}^l = (0.2, -0.1, 0.1, -0.2)^T$ | $\boldsymbol{\gamma}^u = (0.3, -0.2, 0.2, -0.3)^T$ |
| C2 | $x_{ij}^l \sim U[2,4]$ | $x_{ij}^u \sim U[8,10]$ | $\boldsymbol{\gamma}^l = 3 \times (0.2, -0.1, 0.1, -0.2)^T$ | $\boldsymbol{\gamma}^u = 3 \times (0.3, -0.2, 0.2, -0.3)^T$ |
| C3 | $x_{ij}^c \sim U[5,10]$ | $x_{ij}^r \sim U[4,7]$ | $\boldsymbol{\gamma}^c = (0.3, -0.2, 0.2, -0.3)^T$ | $\boldsymbol{\gamma}^r = (0.2, -0.1, 0.1, -0.2)^T$ |
| C4 | $x_{ij}^c \sim U[5,10]$ | $x_{ij}^r \sim U[4,7]$ | $\boldsymbol{\gamma}^c = 3 \times (0.3, -0.2, 0.2, -0.3)^T$ | $\boldsymbol{\gamma}^r = 3 \times (0.2, -0.1, 0.1, -0.2)^T$ |

For comparison, we conduct a comparative analysis of the four models: BHPM, HPM, HMM, and HCRM. More concretely, we first construct the models using training datasets, then employ the fitted models to predict new interval-valued data on the test sets, and finally compute the corresponding evaluation measurements, including $RMSE_L$, $RMSE_U$, $RMSE_H$, and $RI$. The definitions of these measurements are as follows:

1) Following Lima-Neto et al. [5], we separately compute the RMS errors for the lower and upper bounds ($RMSE_L$ and $RMSE_U$):

$$RMSE_L = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^l - \hat{y}_i^l)^2}, \quad RMSE_U = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^u - \hat{y}_i^u)^2}, \quad \text{respectively,}$$

where $y_i^l$ and $y_i^u$ are the observed lower and upper bounds, respectively, and $\hat{y}_i^l$ and $\hat{y}_i^u$ are their corresponding predictions.

2) The $RMSE_H$ measurement, introduced by Carvalho et al. [31], captures the overall prediction accuracy by considering both the center and range of each interval:

$$RMSE_H = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(|y_i^c - \hat{y}_i^c| + |y_i^r - \hat{y}_i^r|\right)^2}.$$

3) The rate of interval agreement (*RI*), introduced by Hu and He [32], assesses how well the predicted intervals overlap with the observed ones:

$$RI = \frac{1}{n}\sum_{i=1}^{n}\frac{w(y_i^I \cap \hat{y}_i^I)}{w(y_i^I \cup \hat{y}_i^I)},$$

where $w(\cdot)$ represents the interval width. The *RI* ranges from 0 to 1, with values closer to 1 indicating better agreement.

Similar to Section 5.1, each setting is replicated 100 times. After specifying no prior information, we estimated the unknown parameters in the BHPM via the Bayesian approach with 5000 iterations, where the burn-in period was set to 3000. For the other three models, we adopted the Weighted Least Squares [21] (WLS) method for the parameter estimation. The average values of the model performance measurements across the four experimental groups were recorded in Tables 5 and 6.

Table 5 presents the performance comparison of the four methods (BHPM, HPM, HMM, and HCRM) across four evaluation measurements under simulated data based on the PM mechanism. The results demonstrate that the BHPM significantly outperforms the other models under most experimental settings. For instance, in the C1 experiment with $n = 100$, the RI values of the BHPM, HPM, HMM, and HCRM are 0.9913, 0.9913, 0.7842, and 0.9333, respectively. When $\boldsymbol{\gamma}^b$ changes in the C2 experiment with $n = 100$, the RI value of the BHPM reaches 0.9989, which is higher than those of the HPM (0.9988), HMM (0.7830), and HCRM (0.9348). Regardless of the variation in $\boldsymbol{\gamma}^b$ across experimental settings, the BHPM consistently outperforms the other models. As the sample size increases, the prediction accuracy of the four models generally improves, while the BHPM consistently maintains its optimal performance.

Table 6 compares the four methods under the CRM data generation mechanism. The results indicate that the HCRM has a slight overall advantage in this scenario, while the BHPM performs very closely behind it. For example, in the C3 experiment with $n = 100$, the RI values of the HCRM and BHPM are 0.9301 and 0.9290, respectively. In the C4 experiment with $n = 200$, the RI values of the HCRM and BHPM are 0.8063 and 0.8079, respectively, showing nearly identical performance. Both models significantly outperform the HPM and HMM in this scenario.

Collectively, the results from both tables highlight that the BHPM exhibits a definitive advantage under the PM data generation mechanism. Under the CRM mechanism, although the BHPM is marginally inferior to the HCRM, the performance gap is minimal. Taking the RI as an example, we further compare the differences among the four methods using box plots, as illustrated in Figures 3 and 4.

## 6. Real data analysis

It has been well-documented that the prevalence of respiratory illnesses, cardiovascular diseases, and stroke-related hospitalizations has significantly risen in China in recent years, which has drawn widespread public attention in the severe adverse health impacts of ambient air pollution. This subsection describes the air quality dataset used in the empirical analysis. The dataset employed in this study is obtained from official air quality monitoring records released by the Ministry of Ecology and Environment (formerly the Ministry of Environmental Protection, MEP) of the People's Republic of China. The study focuses on Shanghai, thereby covering the period from 2014.4.1 to 2021.3.24. We aggregate daily monitoring data into biweekly interval observations. Each biweekly interval is

**Table 5.** Mean and standard deviation (in parentheses) of measurements in C1, C2.

| Config. | $n$ | $RMSE_L$ | | | | $RMSE_U$ | | | | $RMSE_H$ | | | | $RI$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | BHPM | HPM | HMM | HCRM | BHPM | HPM | HMM | HCRM | BHPM | HPM | HMM | HCRM | BHPM | HPM | HMM | HCRM |
| C1 | 100 | 0.4177 | **0.4163** | 12.4634 | 2.8380 | **0.2358** | 0.2366 | 4.5823 | 2.1930 | **0.5977** | **0.5977** | 16.7184 | 4.7536 | **0.9913** | **0.9913** | 0.7842 | 0.9333 |
| | | (0.0741) | (0.0742) | (1.9872) | (0.4548) | (0.0368) | (0.0380) | (0.6354) | (0.3668) | (0.0806) | (0.0823) | (2.4913) | (0.5731) | (0.0012) | (0.0012) | (0.0355) | (0.0079) |
| | 200 | **0.4150** | 0.4152 | 12.5760 | 2.7876 | **0.2327** | 0.2330 | 4.5321 | 2.1098 | **0.5945** | 0.5948 | 16.7866 | 4.6001 | **0.9913** | **0.9913** | 0.7850 | 0.9354 |
| | | (0.0392) | (0.0392) | (1.3438) | (0.2633) | (0.0261) | (0.0263) | (0.4417) | (0.2406) | (0.0458) | (0.0459) | (1.7264) | (0.3644) | (0.0007) | (0.0007) | (0.0251) | (0.0053) |
| C2 | 100 | **0.0739** | 0.0740 | 12.7312 | 2.7910 | **0.0135** | 0.0136 | 4.5953 | 2.1284 | **0.0833** | 0.0835 | 17.0301 | 4.6380 | **0.9989** | 0.9988 | 0.7830 | 0.9348 |
| | | (0.0135) | (0.0136) | (2.0089) | (0.3897) | (0.0028) | (0.0028) | (0.5888) | (0.3158) | (0.0138) | (0.0138) | (2.5197) | (0.5199) | (0.0002) | (0.0002) | (0.0358) | (0.0076) |
| | 200 | **0.0730** | 0.0731 | 12.5029 | 2.7858 | **0.0133** | **0.0133** | 4.5481 | 2.0846 | **0.0820** | 0.0821 | 16.7622 | 4.5960 | **0.9989** | **0.9989** | 0.7864 | 0.9359 |
| | | (0.0082) | (0.0082) | (1.2716) | (0.2680) | (0.0020) | (0.0020) | (0.4246) | (0.2380) | (0.0085) | (0.0084) | (1.6282) | (0.3713) | (0.0001) | (0.0001) | (0.0245) | (0.0054) |

**Table 6.** Mean and standard deviation (in parentheses) of measurements in C3, C4.

| Config. | $n$ | $RMSE_L$ | | | | $RMSE_U$ | | | | $RMSE_H$ | | | | $RI$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | BHPM | HPM | HMM | HCRM | BHPM | HPM | HMM | HCRM | BHPM | HPM | HMM | HCRM | BHPM | HPM | HMM | HCRM |
| C3 | 100 | 2.3476 | 2.3558 | 30.1709 | **2.3071** | 2.2733 | 2.2865 | 5.9460 | **2.2397** | 4.2271 | 4.2413 | 35.4493 | **4.1481** | 0.9290 | 0.9285 | 0.6109 | **0.9301** |
| | | (0.4652) | (0.4685) | (4.6392) | (0.4522) | (0.4561) | (0.4830) | (0.9451) | (0.4381) | (0.6822) | (0.7019) | (5.3107) | (0.6610) | (0.0105) | (0.0109) | (0.0491) | (0.0105) |
| | 200 | 2.3204 | 2.3204 | 29.7852 | **2.2948** | 2.3702 | 2.3747 | 5.9565 | **2.3530** | 4.2866 | 4.2897 | 35.1186 | **4.2488** | 0.9288 | 0.9287 | 0.6094 | **0.9293** |
| | | (0.2802) | (0.2839) | (3.0394) | (0.2846) | (0.2651) | (0.2687) | (0.5896) | (0.2657) | (0.3711) | (0.3801) | (3.5064) | (0.3800) | (0.0056) | (0.0058) | (0.0308) | (0.0058) |
| C4 | 100 | **8.5447** | 8.6061 | 30.6038 | 8.6315 | 8.3424 | 8.3782 | 38.4810 | **8.3245** | 16.0301 | 16.1371 | 38.4810 | **16.0244** | **0.8120** | 0.8095 | 0.5637 | 0.8065 |
| | | (3.1080) | (3.0984) | (5.0271) | (3.0671) | (3.0629) | (3.0459) | (5.9546) | (3.0249) | (5.7603) | (5.7499) | (5.9546) | (5.7083) | (0.0457) | (0.0459) | (0.0510) | (0.0457) |
| | 200 | **9.6504** | 9.6513 | 31.4446 | 9.6831 | 9.4554 | 9.4632 | 11.2102 | **9.4330** | **18.3486** | 18.3608 | 39.9233 | 18.4023 | **0.8079** | 0.8073 | 0.5715 | 0.8063 |
| | | (3.0191) | (2.9970) | (3.0045) | (2.9991) | (2.9359) | (2.9279) | (2.8683) | (2.8786) | (5.7958) | (5.7584) | (4.0429) | (5.7796) | (0.0280) | (0.0281) | (0.0334) | (0.0281) |

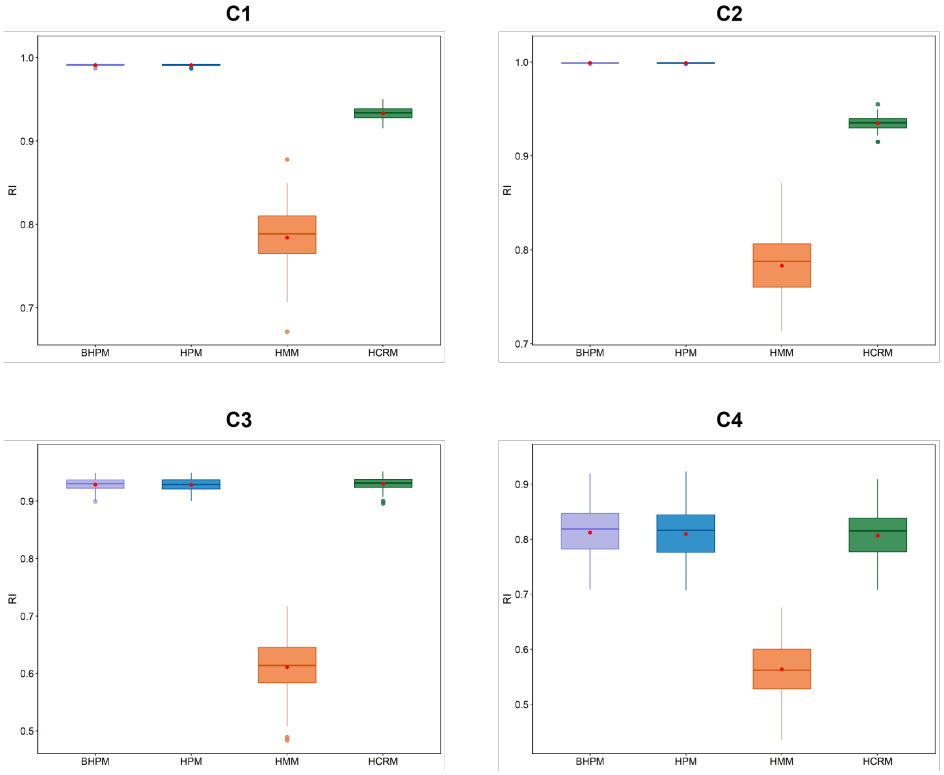


**Figure 3.** The box plots of *RI* calculated from the HCRM, HMM, and BHPM methods with different data set configurations (C1–C4) when $n = 100$ and $p = 3$.

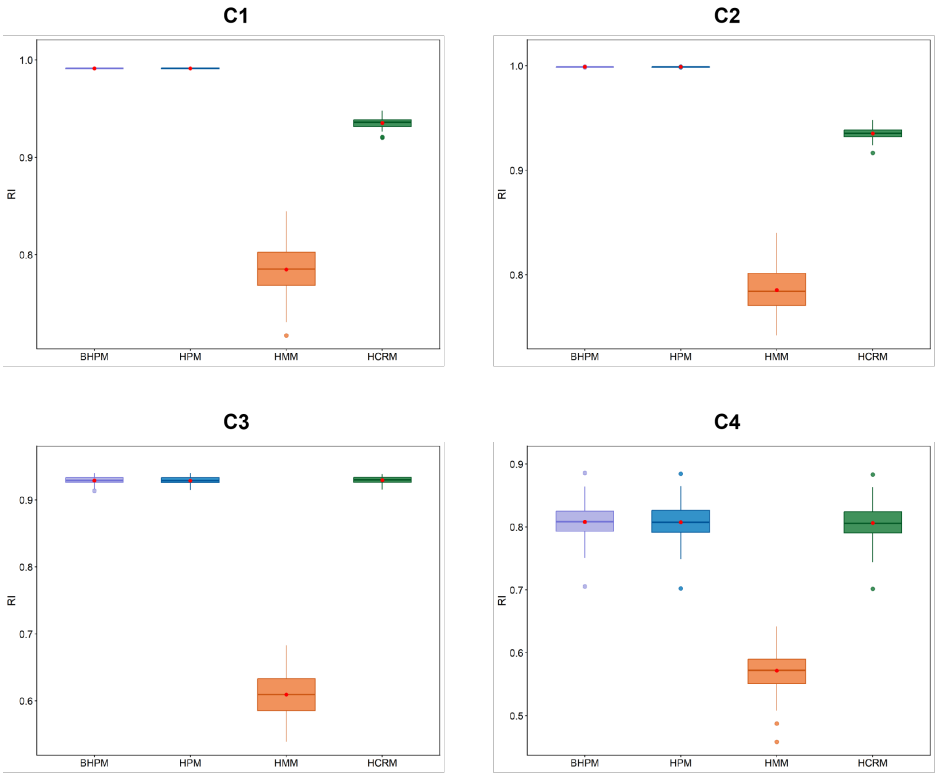


**Figure 4.** The box plots of *RI* calculated from the HCRM, HMM, and BHPM methods with different data set configurations (C1–C4) when $n = 200$ and $p = 3$.

constructed using the minimum and maximum values of the corresponding daily indicators within the 15-day period, thus yielding a total of 170 interval-valued observations. The dataset contains interval data for seven variables: the response variable $Y$ representing air quality index (AQI), and the explanatory variables $X_1, X_2, \ldots, X_6$ which correspond to $PM_{2.5}$, $PM_{10}$, $SO_2$, $NO_2$, CO, and $O_3$, respectively. Partial data are presented in Table 7.

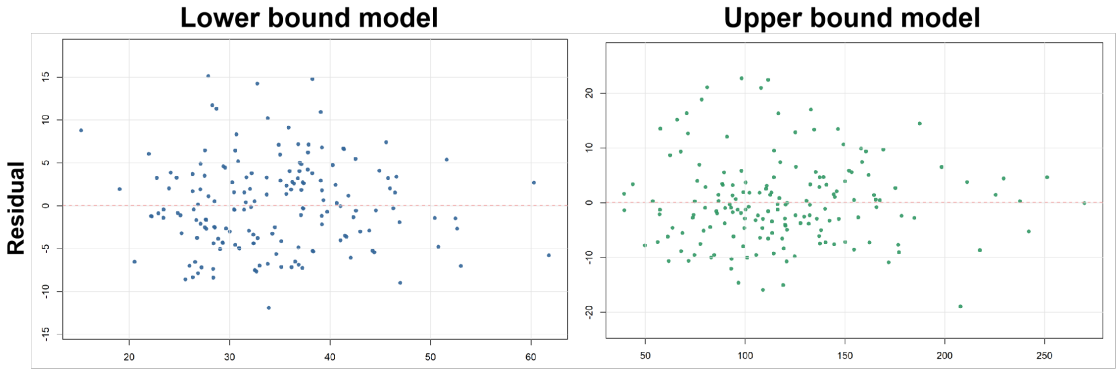


**Figure 5.** Scatter plot of residuals for air dataset.

Prior to estimating the proposed model, we perform the following specification tests. The ugmented Dickey-Fuller (ADF) stationarity test is conducted for all variables, thus confirming that all time series are stationary. Variance Inflation Factor (VIF) tests are further performed on the explanatory variables

**Table 7.** Partial interval-valued air quality dataset for Shanghai.

| NO. | AQI | $PM_{2.5}$ | $PM_{10}$ | $SO_2$ | $NO_2$ | CO | $O_3$ |
|---|---|---|---|---|---|---|---|
| 1 | [50, 182] | [18, 140] | [28, 157] | [9, 30] | [27, 87] | [0.37, 1.5] | [15, 101] |
| 2 | [41, 102] | [16, 73] | [21, 101] | [7, 32] | [22, 73] | [0.51, 1.05] | [43, 105] |
| 3 | [57, 116] | [30, 83] | [43, 112] | [6, 23] | [23, 58] | [0.53, 1.14] | [53, 115] |
| 4 | [45, 189] | [22, 144] | [37, 239] | [7, 34] | [19, 68] | [0.47, 1.06] | [55, 122] |
| 5 | [39, 139] | [21, 94] | [34, 139] | [7, 21] | [20, 56] | [0.52, 0.92] | [57, 131] |
| 6 | [36, 130] | [16, 97] | [21, 122] | [6, 24] | [23, 58] | [0.55, 1.18] | [53, 118] |
| 7 | [41, 169] | [24, 127] | [34, 176] | [8, 30] | [20, 55] | [0.56, 1.24] | [40, 110] |
| 8 | [33, 133] | [14, 101] | [32, 117] | [5, 18] | [10, 40] | [0.45, 0.99] | [33, 96] |
| 9 | [37, 115] | [20, 73] | [29, 104] | [7, 17] | [17, 43] | [0.52, 1.08] | [41, 125] |
| 10 | [37, 101] | [19, 68] | [22, 92] | [8, 21] | [26, 50] | [0.55, 1.08] | [46, 102] |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 170 | [37, 92] | [19, 62] | [33, 94] | [4, 7] | [30, 52] | [0.48, 0.8] | [37, 85] |

to assess multicollinearity. All VIF values are well below the conventional threshold of 10, thus confirming the absence of significant multicollinearity among the predictors. The Breusch–Pagan test for heteroscedasticity was separately applied to the upper bound and lower bound models. The test statistics are $BP^u = 35.743$ and $BP^l = 23.358$, both with $p < 0.05$. Furthermore, the residual analysis presented in Figure 5 confirms the presence of clear heteroscedasticity. Considering the potential heteroscedasticity in the air dataset, we separately establish models using the BHPM, HPM, HMM, and HCRM. All data are divided into training and test sets with an 8:2 proportion. Each model is repeatedly fitted 50 times to obtain the final results. Table 8 presents the comparison results of different methods in terms of the measurements $RMSE_L$, $RMSE_U$, $RMSE_H$, and $RI$. Overall, the BHPM achieves better fitting performance than the other models. For instance, in terms of the RI, the average value of the BHPM is 0.8584, which is larger than those of the other methods (0.8560 for the HPM, 0.8562 for the HMM, and 0.8575 for the HCRM). The results of the linear regression equation established by the BHPM model are shown as follows:

$$\begin{aligned}\hat{y}^l = {} & 0.0020 + 0.3121x_1^l + 0.0217x_1^u + 0.2719x_2^l + 0.0142x_2^u + 0.0499x_3^l + 0.0696x_3^u \\ & + 0.0789x_4^l + 0.0700x_4^u + 0.0024x_5^l + 0.0014x_5^u + 0.1222x_6^l + 0.0737x_6^u \end{aligned} \tag{6.1}$$

$$\begin{aligned}\hat{y}^u = {} & 0.0070 + 0.0335x_1^l + 0.9968x_1^u - 0.0682x_2^l + 0.1704x_2^u - 0.0167x_3^l - 0.0929x_3^u \\ & + 0.0776x_4^l + 0.0731x_4^u + 0.0044x_5^l + 0.0110x_5^u + 0.0020x_6^l + 0.1305x_6^u \end{aligned} \tag{6.2}$$

Based on regression Eqs (6.1) and (6.2), AQI is primarily influenced by $PM_{2.5}$ and $PM_{10}$. In the lower bound model, the coefficients of $PM_{2.5}$ and $PM_{10}$ are 0.3121 and 0.2719, respectively, thus dominating the lower bound of AQI. In the upper bound model, the coefficient of $PM_{2.5}$ reaches 0.9968, which is significantly larger than other variables, thus indicating its dominant role in shaping the

upper bound of AQI. Although the coefficients are theoretically positive, several negative values are observed. A correlation analysis confirms certain negative correlations among explanatory variables, which, combined with interval-level pollutant peak asynchrony, explain these negative coefficients.

**Table 8.** Comparison of different methods in air dataset (standard deviations in parentheses).

| | $\text{RMSE}_L$ | $\text{RMSE}_U$ | $\text{RMSE}_H$ | RI |
|---|---|---|---|---|
| BHPM | **5.7825** | 8.4877 | **12.7244** | **0.8584** |
| | (0.5944) | (1.3708) | (1.3545) | (0.0192) |
| HPM | 5.8183 | 8.6243 | 13.0237 | 0.8560 |
| | (0.6128) | (1.4577) | (1.3850) | (0.0185) |
| HMM | 5.8040 | 8.5206 | 12.9280 | 0.8562 |
| | (0.5668) | (1.5534) | (1.5632) | (0.0180) |
| HCRM | 5.8482 | **8.3577** | 12.7912 | 0.8575 |
| | (0.6336) | (1.2886) | (1.2347) | (0.0174) |

## 7. Conclusions

This paper proposed a novel Bayesian heteroscedastic parametrized method designed for interval-valued data. The core innovation of the BHPM lies in integrating the reference point of the traditional PM model with heteroscedasticity, thereby constructing a more robust regression model. Bayesian estimates of the extended BHPM were subsequently derived via the combination of Gibbs sampling and MH algorithms. Compared with conventional interval-valued regression approaches, the BHPM not only introduces heteroscedastic structures but also enhances model interpretability and a generalization capability through the Bayesian inference framework.

The results from both simulation studies and real data application demonstrated the superiority of the BHPM in linear interval regression. The simulation results indicated that the BHPM remains the optimal choice to analyze linear relationships in the interval-valued data. The analysis of real dataset further validated the effectiveness and practical applicability.

For future research, we will extend the scope of the BHPM in two key directions. First, we will generalize the current linear model to partially linear or nonlinear models to better capture the complex nonlinear relationships in interval-valued data. Second, to meet the growing demand for panel interval data analysis, we will incorporate fixed effects (e.g., time-specific effects) into the Bayesian framework for the parameter estimation. This extension will enable the BHPM to be widely applied in empirical fields such as economics and finance.

## Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

## Acknowledgments

This research was funded by the National Social Science Fund of China (Grant No.23BTJ069). We would like to thank the anonymous reviewers and the journal editors for their valuable and constructive suggestions, which have significantly enhanced the quality of this paper.

## Conflict of interest

The authors declare there is no conflicts of interest.

## References

1. J. P. Liu, P. Wang, H. Y. Chen, J. M. Zhu, A combination forecasting model based on hybrid interval multi-scale decomposition: application to interval-valued carbon price forecasting, *Expert Syst. Appl.*, **191** (2022), 116267. https://doi.org/10.1016/j.eswa.2021.116267
2. X. Wang, S. M. Li, T. Denoeux, Interval-valued linear model, *Int. J. Comput. Intell. Syst.*, **8** (2015), 114–127. https://doi.org/10.1080/18756891.2014.967010
3. L. Billard, E. Diday, Regression analysis for interval-valued data, in *Data Analysis, Classification, and Related Methods* (eds. H. A. L. Kiers, J. P. Rasson, P. J. F. Groenen, M. Schader), Springer, (2000), 369–374. https://doi.org/10.1007/978-3-642-59789-3_58
4. L. T. Kong, X. W. Gao, A regularized MM estimate for interval-valued regression, *Expert Syst. Appl.*, **238** (2024), 122044. https://doi.org/10.1016/j.eswa.2023.122044
5. E. A. Lima Neto, F. A. T. de Carvalho, Centre and range method for fitting a linear regression model to symbolic interval data, *Comput. Stat. Data Anal.*, **52** (2008), 1500–1515. https://doi.org/10.1016/j.csda.2007.04.014
6. P. Hao, J. P. Guo, Constrained center and range joint model for interval-valued symbolic data regression, *Comput. Stat. Data Anal.*, **116** (2017), 106–138. https://doi.org/10.1016/j.csda.2017.06.005
7. M. Xu, Z. F. Qin, A bivariate Bayesian method for interval-valued regression models, *Knowl.-Based Syst.*, **235** (2022), 107396. https://doi.org/10.1016/j.knosys.2021.107396
8. L. C. Souza, R. M. C. R. de Souza, G. J. A. Amaral, T. M. S. Filho, A parametrized approach for linear regression of interval data, *Knowl.-Based Syst.*, **131** (2017), 149–159. https://doi.org/10.1016/j.knosys.2017.06.012
9. M. Xu, Z. F. Qin, A Bayesian parametrized method for interval-valued regression models, *Stat. Comput.*, **33** (2023), 67. https://doi.org/10.1007/s11222-023-10234-2
10. A. Duary, M. S. Rahman, A. K. Manna, A. A. Shaikh, Optimal policy of an interval production problem with variable demand and warranty policy via 0-1 parametrized optimization technique, *Int. J. Syst. Assur. Eng. Manage.*, **16** (2025), 2967–2982. https://doi.org/10.1007/s13198-025-02808-2
11. S. M. Aljeddani, F. E. Almuhayfith, M. S. Rahman, Parametric optimization of an interval sustainable supply chain with green-thermal-linked demand rate, *Mathematics*, **14** (2025), 91. https://doi.org/10.3390/math14010091

12. T. T. Huang, A spatial Durbin model for interval-valued data with t-distribution, *J. Stat. Comput. Simul.*, **94** (2024), 4037–4071. https://doi.org/10.1080/00949655.2024.2408654

13. L. T. Kong, X. J. Song, X. M. Wang, Nonparametric regression for interval-valued data based on local linear smoothing approach, *Neurocomputing*, **501** (2022), 834–843. https://doi.org/10.1016/j.neucom.2022.06.073

14. M. Xu, Z. F. Qin, Bayesian framework for interval-valued data using Jeffreys' prior and posterior predictive checking methods, *Commun. Stat. Simul. Comput.*, **53** (2024), 2425–2443. https://doi.org/10.1080/03610918.2022.2076869

15. A. B. Ji, J. J. Zhang, Y. Cao, Bivariate Bayesian regression method for fixed effects panel interval-valued data models, *Commun. Stat. Theory Methods*, **54** (2025), 5545–5565. https://doi.org/10.1080/03610926.2024.2440003

16. J. Zhang, M. Liu, M. Dong, Variational Bayesian inference for interval regression with an asymmetric Laplace distribution, *Neurocomputing*, **323** (2019), 214–230. https://doi.org/10.1016/j.neucom.2018.09.083

17. P. O. Goffard, P. J. Laub, Approximate Bayesian computations to fit and compare insurance loss models, *Insur. Math. Econ.*, **100** (2021), 350–371. https://doi.org/10.1016/j.insmatheco.2021.06.002

18. K. Shimizu, Asymptotic properties of Bayesian inference in linear regression with a structural break, *J. Econ.*, **235** (2023), 202–219. https://doi.org/10.1016/j.jeconom.2022.03.006

19. F. Q. Li, M. T. Zhao, K. S. Zhang, Bayesian adaptive Lasso estimation of large graphical model based on modified Cholesky decomposition, *Stat. Probab. Lett.*, **206** (2024), 110004. https://doi.org/10.1016/j.spl.2023.110004

20. K. Fan, S. Subedi, V. Dissanayake, C. Wu, Robust Bayesian high-dimensional variable selection and inference with the horseshoe family of priors, *Comput. Stat. Data Anal.*, **219** (2026), 108358. https://doi.org/10.1016/j.csda.2026.108358

21. H. White, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, **48** (1980), 817–838. https://www.jstor.org/stable/1912934

22. G. S. Maddala, Generalized least squares with an estimated variance covariance matrix, *Econometrica*, **39** (1971), 23–33. https://doi.org/10.2307/1909137

23. D. K. Xu, Z. Z. Zhang, A semiparametric Bayesian approach to joint mean and variance models, *Stat. Probab. Lett.*, **83** (2013), 1624–1631. https://doi.org/10.1016/j.spl.2013.02.023

24. M. Pourahmadi, Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation, *Biometrika*, **86** (1999), 677–690. http://www.jstor.org/stable/2673662

25. W. Zhang, C. Leng, C. Y. Tang, A joint modelling approach for longitudinal studies, *J. R. Stat. Soc. Ser. B*, **77** (2015), 219–238. http://www.jstor.org/stable/24774732

26. Y. Zhong, Z. Z. Zhang, S. M. Li, A constrained interval-valued linear regression model: a new heteroscedasticity estimation method, *J. Syst. Sci. Complex.*, **33** (2020), 2048–2066. https://doi.org/10.1007/s11424-020-9075-2

27. G. O. Roberts, Markov chain concepts related to sampling algorithms, in *Markov Chain Monte Carlo in Practice* (eds. W. R. Gilks, S. Richardson, D. J. Spiegelhalter), Chapman and Hall/CRC, (1996), 45–58. https://doi.org/10.1201/b14835-8

28. S. Y. Lee, H. T. Zhu, Statistical analysis of nonlinear structural equation models with continuous and polytomous data, *Br. J. Math. Stat. Psychol.*, **53** (2000), 209–232. https://doi.org/10.1348/000711000159303

29. A. Gelman, G. O. Roberts, W. R. Gilks, Efficient Metropolis jumping rules, in *Bayesian Statistics 5* (eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, A. F. M. Smith), Oxford University Press, (1996), 599–608. https://doi.org/10.1093/oso/9780198523567.003.0038

30. C. J. Geyer, Practical Markov chain Monte Carlo, *Stat. Sci.*, **7** (1992), 473–483. https://doi.org/10.1214/ss/1177011137

31. F. A. T. de Carvalho, R. M. C. R. de Souza, M. Chavent, Y. Lechevallier, Adaptive Hausdorff distances and dynamic clustering of symbolic interval data, *Pattern Recognit. Lett.*, **27** (2006), 167–179. https://doi.org/10.1016/j.patrec.2005.08.014

32. C. Y. Hu, L. T. He, An application of interval methods to stock market forecasting, *Reliab. Comput.*, **13** (2007), 423–434. https://doi.org/10.1007/s11155-007-9039-4


AIMS Press