Bayesian modeling with locally adaptive prior parameters in small animal imaging

Muyang Zhang[1], Robert G. Aykroyd[1]* and
Charalampos Tsoumpas[1,2]

[1]Department of Statistics, School of Mathematics, University of Leeds, Leeds, United Kingdom,
[2]Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen,
University of Groningen, Groningen, Netherlands

Medical images are hampered by noise and relatively low resolution, which create a bottleneck in obtaining accurate and precise measurements of living organisms. Noise suppression and resolution enhancement are two examples of inverse problems. The aim of this study is to develop novel and robust estimation approaches rooted in fundamental statistical concepts that could be utilized in solving several inverse problems in image processing and potentially in image reconstruction. In this study, we have implemented Bayesian methods that have been identified to be particularly useful when there is only limited data but a large number of unknowns. Specifically, we implemented a locally adaptive Markov chain Monte Carlo algorithm and analyzed its robustness by varying its parameters and exposing it to different experimental setups. As an application area, we selected radionuclide imaging using a prototype gamma camera. The results using simulated data compare estimates using the proposed method over the current non-locally adaptive approach in terms of edge recovery, uncertainty, and bias. The locally adaptive Markov chain Monte Carlo algorithm is more flexible, which allows better edge recovery while reducing estimation uncertainty and bias. This results in more robust and reliable outputs for medical imaging applications, leading to improved interpretation and quantification. We have shown that the use of locally adaptive smoothing improves estimation accuracy compared to the homogeneous Bayesian model.

KEYWORDS

Bayesian modeling, inhomogeneous parameter, image processing, Markov random field, Markov chain Monte Carlo

1 Introduction

As a non-invasive method, medical imaging is extensively used for diagnosing and monitoring various medical conditions (1). However, the loss of information during the scanning and image acquisition processes often creates an observed image that is blurred and contains noise (2). The systematic relationship between the observed and true image is often modeled linearly using a transformation matrix. However, this transformation matrix is typically large and ill-posed, so directly solving a system of linear equations to obtain the exact image is infeasible.

In medical image processing, Bayesian modeling transforms an ill-posed problem into a well-posed problem by introducing a prior distribution as a form of penalization or regularization. Moreover, this method holds potential for application in biomedical image reconstruction (3, 4). Most approaches, however, have the tendency to not only

smooth out noise but also to smooth out the signal. This raises the question of how to determine a prior distribution for smoothness in order to avoid both under- and over-smoothing. Homogeneous prior distributions have been found to be less effective in scenarios with rapid changes, such as medical images [5, 6]. Inhomogeneous Bayesian modeling aims to fully utilize the distribution's properties. Instead of employing different prior distributions, one could consider using a prior distribution with hyper-prior parameters [7]. Therefore, we integrate inhomogeneous factors into the modeling by updating our prior distribution, introducing locally adaptive hyper-prior parameters, with high dimensions, instead of a single global hyper-parameter.

2 Materials

Given the absence of real images in practical scenarios, certain statistical measurements such as mean squared error (MSE) evaluating the efficiency of statistical modeling by minimizing the difference between real and estimated values are limited in application. Hence, creating simulated data to mimic the real image is required to bridge this gap.

2.1 Designed simulation with high contrast

The process for creating simulated data is as follows: we aim to generate simulated data X that closely approaches the true image represented as "real data." We then apply random noise to create the degraded observed data Y. In this case, instead of using the projection data P in the posterior, we can obtain estimations by sampling from the posterior conditional distribution given Y. This approach also allows for a comparison between estimations and the corresponding real data X. The simulation is based on the function minimum residual sum of squares (RSS), where we can adjust the parameters to achieve simulations with different levels of noise. The general expression of the function is

$$\min_{\theta} \| Y - A(\delta)X(\mathbf{r}, \mathbf{c}, \mathbf{z}, T) - \boldsymbol{\epsilon} \|^2,$$

where $\mathbf{r}$ is the collection of four objects' radius: $\mathbf{r} = (r_1, r_2, r_3, r_4)^T$; $\mathbf{c} = (c_1, c_2, c_3, c_4)^T$ and $\mathbf{z} = (z_1, z_2, z_3, z_4)^T$ represent the parameter vectors of the objects' central position of x-axis and y-axis, respectively. $\boldsymbol{\epsilon}$ represents random errors introduced during observation. Given the prior information, the density for each cylinder is identical, denoted as T.

The elements within A represent the probabilities that pixels in X can be transformed into corresponding information in Y. In our application, a_{ij} follows a bivariate-normal distribution with zero covariance:

$$a_{ij} \sim N(\mu, \Sigma),$$

where $\mu = (0, 0)$ and $\Sigma = \delta^2 \mathbf{I}_{2 \times 2}$ ($\mathbf{I}$ is an identity matrix). The

standard deviation δ is highly dependent on the distance between the scanner and the scanner object. The transformed information from X to Y decreases with increasing distance. Ideally, when there is no distance between the scanner and the object, A is an identity matrix, allowing for complete information transformation. In medical imaging, despite the scanner's position being fixed during an examination, the distance may vary slightly depending on the size of the object. Hence, when there is a considerable distance between the scanner and scanned objects, information is missing due to the reception of limited signals. Conversely, increasing the distance introduces more blurring. Overall, it is crucial to strike a balance between maintaining information and reducing blurring simultaneously; in other words, deciding how to set δ becomes vital.

2.2 Designed simulation with high contrast

To generate simulations, we set values for parameters. Afterward, random noise is applied to create the degraded observation Y. In this case, we can obtain estimations by sampling from the posterior conditional distribution given X [8]. This approach also allows for a comparison between estimations and the corresponding real data X. The simulations are stored in pixel matrices of the size 29×58.

Observed data Y, viewed as a degraded version of the actual distributions X with blur and noise, is comparable to the projection dataset in reality. As shown in Figure 1a, the true distribution consists of four sharp regions with sharp boundaries. The red lines indicate the regions of interest, namely, the 20th row and the 36th column, which are applied in the following pixel estimations. Figure 1b depicts observed data with low contrast resolution; blur is evident around the edge of each region.

The hot region and background pixels are constant, around 1,100 and 0, respectively. In the three simulations, the noise presents in different ways. The first simulated dataset has the lowest noise compared to the other two datasets since there is a distinguished value gap between the background, where the pixels are close to zero, and the hot region, where the pixels are close to 1,100. In the second simulated dataset in Figure 2, it creates an environment where the hot regions are soft tissues with smooth edges; the average value of pixels in the background is higher than in other experiences, while the peak values are smaller than others. The red dots indicate the positions of pixels applied in the following pixel estimation comparison. Finally, the ones in Figure 3 describe the situation when there is a reduction in scanning time. There is a higher noise level than the first simulation. These simulations with artificial noise are regarded as degraded images. The proposed methods are assessed using simulated examples designed specifically to mimic real experimental data collected as part of system calibration experiments.

FIGURE 1
Simulation datasets: true information X and its correspondingly observation data Y. **(a)** Simulated image X. **(b)** Observation Y. **(c)** 3D simulated image X. **(d)** 3D observation Y.
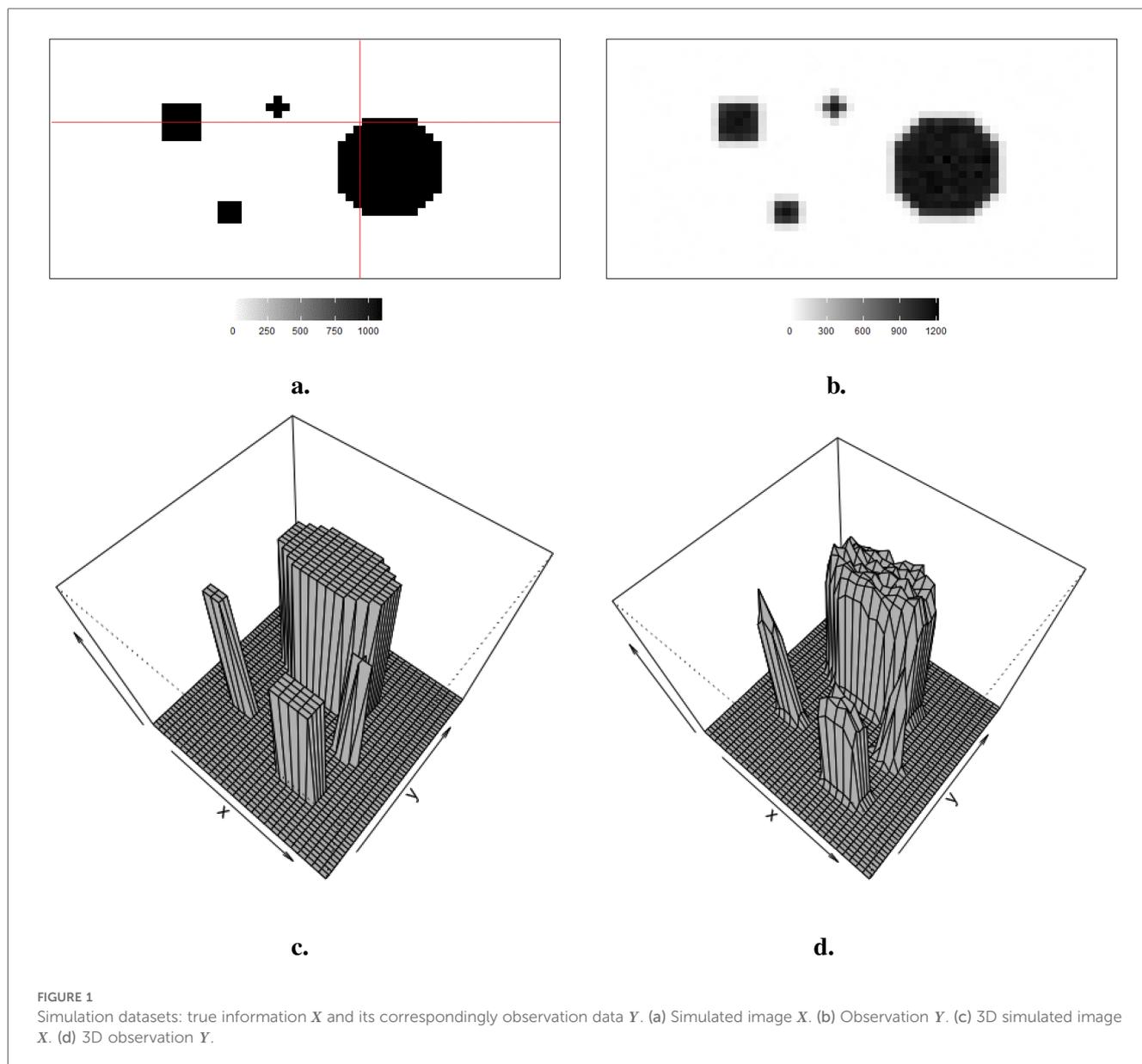
2.3 Designed simulation with smoothing edge

Supposing the objects within the circle have a soft edge instead of a hard one. In reality, the edges are likely to be considerably softer. Therefore, a smoothing pattern is obtained by applying a Gaussian kernel filter to the datasets presented in Figure 2a. If $\mathcal{G}$ is a Gaussian kernel, we say that $\mathbf{X}_1 = \mathcal{G} \times \mathbf{X}$. The high blurring around the high-contrast edge between the hot regions and the background makes it difficult to detect the original edge.

2.4 Simulated experiment with lower detected counts

Image quality improves with a reasonable extension of scanning duration. Again, supposing the scanning time is

reduced, the observation image contains blurring. Another degraded simulated image is created, as shown in Figure 3. The high-contrast edge between the hot regions and the background blurring and the observation image is expected to contain more noise than the first observation image Y.

Ultimately, these simulated Y, Y_1, and Y_2 serve as our "observed dataset" for the application sections in this paper, while the simulated X remains an unknown parameter that needs to be estimated. However, as the values of simulated X are obtained in advance, it allows for a comparison between simulated X and its estimations from our defined posterior distribution in the following step.

3 Methods

In the case of medical imaging, the aim is to estimate a discrete version of the unknown continuous emitter activity distribution X

FIGURE 2
Simulation datasets with a reduction in scanning time: true image X_1 (left) and observed image Y_1 (right). **(a)** Simulated image X_1. **(b)** Observation Y_1.

FIGURE 3
Simulation datasets with scanning time reduction: observed image Y_2 and its corresponding actual image X. **(a)** Simulated image X. **(b)** Observation image Y_2.
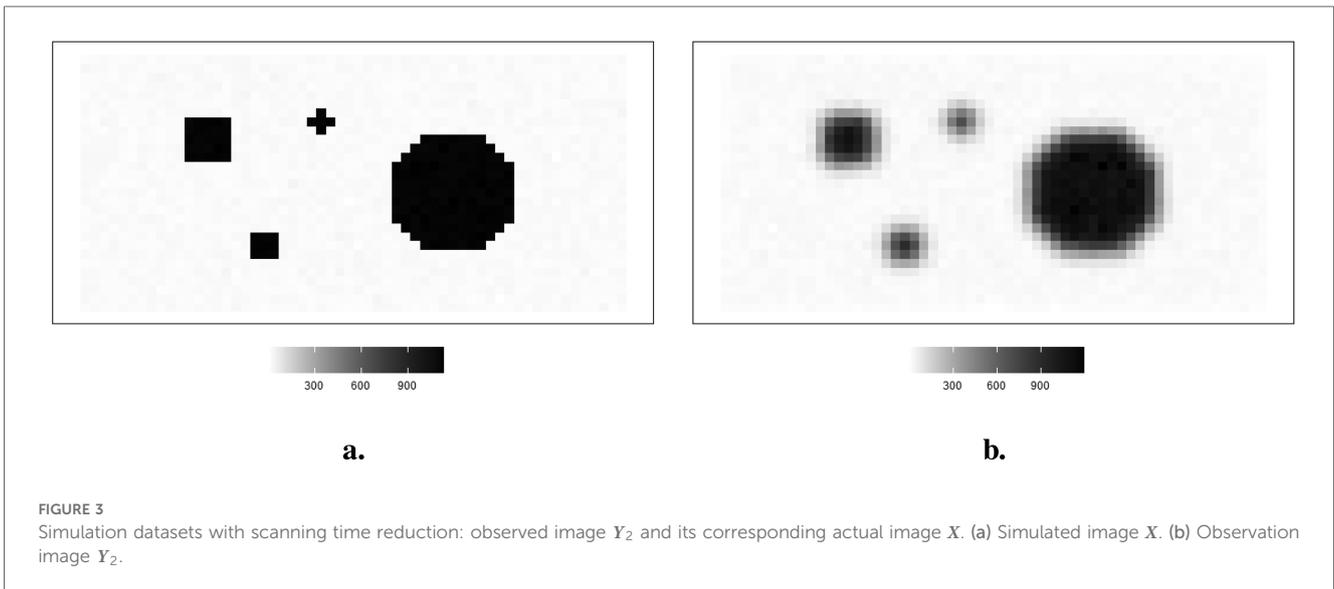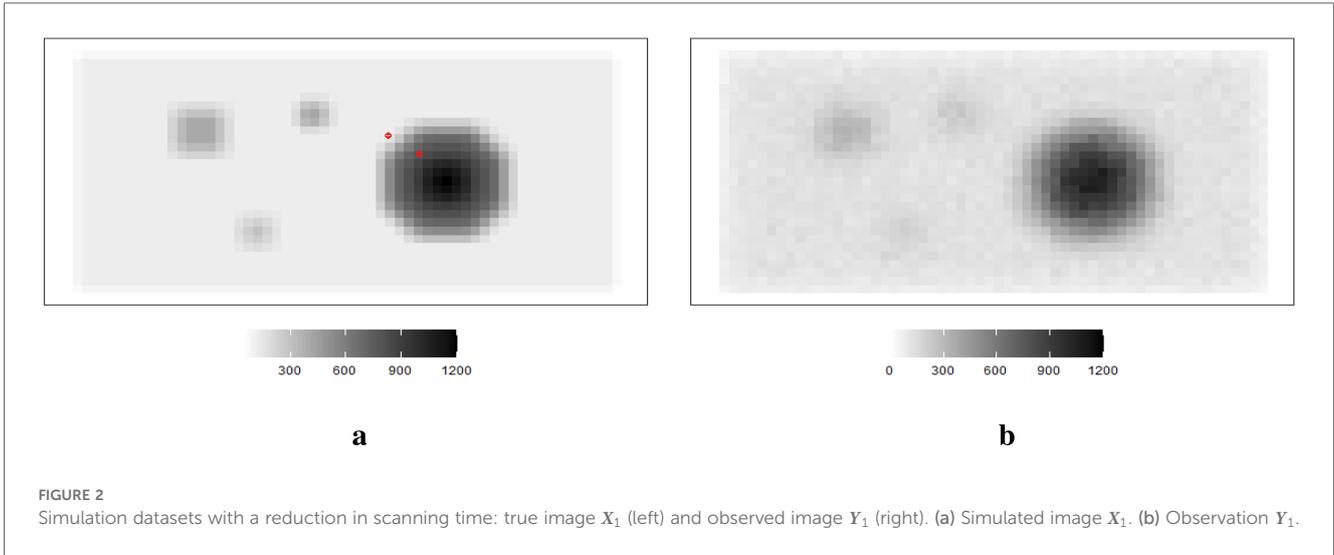
from the single projection data Y. Suppose the unknown object X is expressed as a set of m volume pixels, $X = (X_j : j = 1, 2, m)$, where X_j represents the constant value of emitter activity in the jth pixel. The data $Y = (Y_i : i = 1, 2, n)$ is related to the *actual activity* through the deterministic equation $E(Y) = f(X)$ and depending on the application being studied, $f(X)$ can become a linear function, or remain a non-linear function, especially when scanning time and multilayer factors are considered (9–11).

3.1 Likelihood function

A Poisson form, identified as suitable for various image processing with quantum noise, is particularly appropriate for the -eye$^{\text{TM}}$ camera projection data considered in our application (12–14). The first -eye$^{\text{TM}}$ scintillation camera developed by

BIOEMTECH (Athens, Greece) was used to generate two-dimensional medical images (15).

The conditional distribution for observation Y given the unknown true radionuclide distribution X is as follows:

$$f_{Y|X}(y_1, y_2, , y_n \mid x) = \prod_{i=1}^{n} \lambda_i^{y_i} \frac{\exp(-\lambda_i)}{y_i!}, \qquad (1)$$

where $E[Y_i] = \lambda_i = \sum_{j=1}^{m} a_{ij} x_j$, $j = 1, 2, , m$. In other words, each projection data value has an interaction with the whole vector X; a known transformation matrix is denoted $A = [a_{ij}]_{n \times m}$. The element a_{ij} is the probability that a gamma particle emitted from pixel location i is recorded at pixel location j. The error ϵ can be expressed as an $n \times 1$ vector with elements $(\epsilon_i : i = 1, 2, , n)$, which may come from various types of unavoidable measurement errors.

If the transformation matrix A is square and non-singular, image X can be easily realized by the least squares estimator $X \leftarrow \arg\min_X \|Y - AX\|^2$, when the squared error function is minimized:

$$\hat{X} = \left(A^T A\right)^{-1} A^T Y,$$

where $\hat{X}$ is the least square estimate of X. However, in image processing, the transformation matrix A typically has a complex structure with high dimensions and is rectangular with unavailable pseudo-inverse. This results in ill-posed and ill-conditioned issues (16).

By imposing additional constraints in terms of prior knowledge, a Bayesian approach transforms an ill-posed inverse problem into one that is well-posed (17). During the modeling process, a prior distribution is constructed to capture the statistical properties of the image, and then estimation uses a posterior distribution derived by the combination of prior and likelihood. The uncertainty between X and Y is captured by likelihood function $f_{Y|X}(y \mid x)$ and the posterior density $f_{X|Y}(x \mid y)$ is used for inference after incorporating prior knowledge $\pi_X(x)$.

3.2 Prior distribution

Discrete images comprise elements of finite product spaces, and the probability distributions on such sets of images as prior information are a critical part of image processing. The efficiency of prior distributions depends on the available first-hand information. Regarding informative priors, it is generally expressed as the Gibbs measure, which was borrowed from statistical physics (18). The primary goal of introducing this type of probability is to describe the features relative to "macroscopic" physical samples, such as an "infinite system" (19, 20).

The Gibbs probability distribution has gradually found applications in various fields, including "Gibbs Sampling" in

Bayesian modeling. The Gibbs distribution is defined as

$$\pi_X(x \mid B) = \mathcal{Z}^{-1} \exp(-B\kappa(x)), \quad \mathcal{Z} = \int_x \exp(-B\kappa(x))\mathrm{d}x,$$
$$X \in R^m, B > 0, \tag{2}$$

where $\mathcal{Z}$ is the normalization for the Gibbs distribution; the energy function is κ, representing the energy of the configuration of pixels; and B is a non-negative smoothing parameter. Furthermore, the energy function can be rewritten as the sum of local energy functions $\Phi(\cdot)$:

$$\kappa(x) = \sum_{j=1}^m \Phi_j(x), \tag{3}$$

where $\Phi_j(\cdot)$ represents the local energy function to corresponding $X_j = x_j$.

3.2.1 Markov random field for pixel differences

Briefly, the primary assumption for Markov random field (MRF) models is that a variable is only related to its adjacent variables while being conditionally independent of the others (21, 22). Specifically, the clique-based structure makes MRF models well-suited for capturing local pixel relations in images. It proposes a lattice system, denoted as $G = (V, E)$, to represent the connections between pixels (as illustrated in Figure 4). For instance, assuming the yellow node represents the object under analysis in the first-order system, its four closest neighbors are located to its left, right, bottom, and top sides, as indicated by the black solid lines. In the second-order system, an additional four neighborhoods are considered, located at the top-left, top-right, bottom-left, and bottom-right corners around the yellow node, as indicated by the dashed lines. In this system, pixels are represented as nodes (V), and edges (E) connect all the nodes. While the shape of the grid is not required to be rectangular, it is the most common in applications.

FIGURE 4
Two-dimensional rectangular grid G. The blue and yellow nodes in the lattice represent pixels, while the solid lines and dashed lines describe first-order and second-order neighborhoods, respectively.
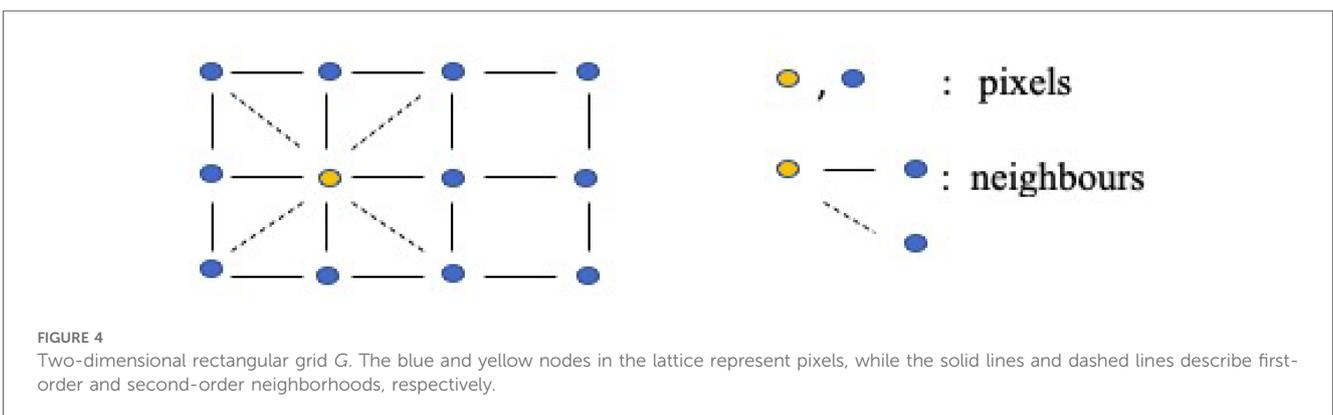
Once we define the MRF prior, the local function in Equations 2, 3 is rewrite a linear combination of differences between the pixel and its neighborhoods:

$$\Phi_j(x) = \sum_{t \in \partial(j)} w_{jt} \phi(x_j - x_t), \tag{4}$$

where w_{jt} represents the weight for each paired comparison, as the increased order in MRF, and w_{jt} may change according to the interaction within the neighborhood. For example, for the first-order MRF, $w_{jt} = 0.5$ if there is only four neighbors to consider. The set of nodes $\partial(j)$ in a finite graph X with edges $j \sim t$. Finally, after employing the Markov random field for pixel difference, the updated prior distribution is

$$\pi_X(\boldsymbol{x} \mid B) = \mathcal{Z}^{-1} \exp\left(-B \sum_{j=1}^{m} \sum_{t \in N_{\partial(j)}} w_{jt} \phi(x_j - x_t)\right). \tag{5}$$

Here, only local characteristics are considered when it comes to estimating the individual pixel, which can be briefly divided into two properties. First, in comparison to the global character which includes all the pixels, we only study the smaller number of pixel's neighbors; for instance, the first-order neighborhood is adopted in the prior distribution. Second, given the local property of the prior distribution, we assume that the posterior is sensitive to the local property of the prior distribution. If we define the potential functions as absolute value function, then $\phi(\mu) = |\mu|$ respectively. Thereby, the corresponding priors are quantified Markov random fields with absolute function (LMRF) combined with Equations 4, 5:

$$\pi_{X|\tau}(\boldsymbol{x} \mid \boldsymbol{\tau}) = \prod_{j=1}^{m} \frac{1}{(2\tau_j)} \exp\left(-\frac{\sum_{t \in \partial(j)} |x_j - x_t|}{\tau_j}\right), \tag{6}$$
$$x_j \geq 0, \tau_j > 0,$$

where $X = \{x_j, j = 1, 2, , m\}$ and x_j is conditional based on the neighbor's values and $\boldsymbol{\tau} = \{\tau_j, j = 1, 2, , m\}$ is the local conditional variance in the prior distribution, which accounts for the value variances among individual pixels and its four neighbors. As a comparison to Equation 6, we define a homogeneous prior variance τ to capture the global variance between the pixel differences:

$$\pi_{X|\tau}(\boldsymbol{x} \mid \boldsymbol{\tau}) = \prod_{j=1}^{m} \frac{1}{(2\tau)} \exp\left(-\frac{\sum_{t \in \partial(j)} |x_j - x_t|}{\tau}\right),$$
$$x_j \geq 0, \tau > 0.$$

The inhomogeneous prior relies on the contrast levels in the image's segmentation, making it locally adaptive to various image densities. In order to cover all scenarios among pixel neighborhoods, the modeling with inhomogeneous hyper-parameter τ can tolerate a large value fluctuation and detect a small variation within the smoothing areas.

3.2.2 High-dimensional Markov random field

The image dataset presents a two-dimensional image in our application. Hence, we only consider the first-order system in Markov random field priors, where the pixel and its four closest neighbors are on the same planet. However, if the application dataset from the projection data to the tomography image, where pixels within two-dimensional space transfer into voxels within three-dimensional space, we can introduce another two neighbors of pixels based on the first-order system. As seen in Figure 5, the left side, as a comparison to the right side, displays the Markov random field within two-dimensional space. There are two additional red nodes applied based on the first-order system.

The Markov random field prior distribution can still be written as a general form of Gibbs distribution:

$$\pi_X(\boldsymbol{x} \mid B) = kB^m \exp\left(-B \sum_{j=1}^{m} \sum_{t \in N_{\partial(j)}} w_{jt} \phi(x_j - x_t)\right),$$
$$x_j \geq 0, \tau > 0,$$

where the number of neighbors in $N_{\partial(j)}$ increases from four to six neighbors in three-dimensional space.

3.2.3 Hyper-prior distribution

In the inhomogeneous model, $\boldsymbol{\tau}$ is a vector of unknown local parameters which can either be defined by allocating a series of artificial values or by introducing another level of modeling that correctly incorporates the additional uncertainty. The first solution requires advanced information for the assignment which is not usually available. Alternatively, the second action employs a hyper-prior distribution, $\pi_\tau(\boldsymbol{\tau})$, which is diffuse unless there is
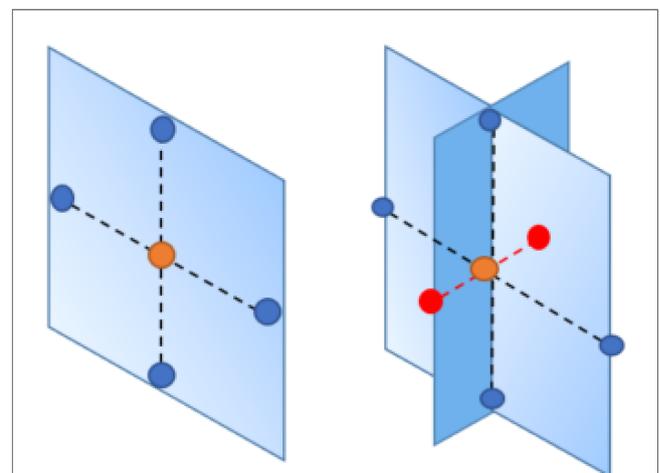
FIGURE 5
Markov random field in three-dimensional space. Left image: Markov random field in two-dimensional space. Right image: Markov random field in three-dimensional space. The nodes with different colors in the lattice represent pixels, while the dashed lines describe the connection between the objective pixel (orange node) and its neighbors.
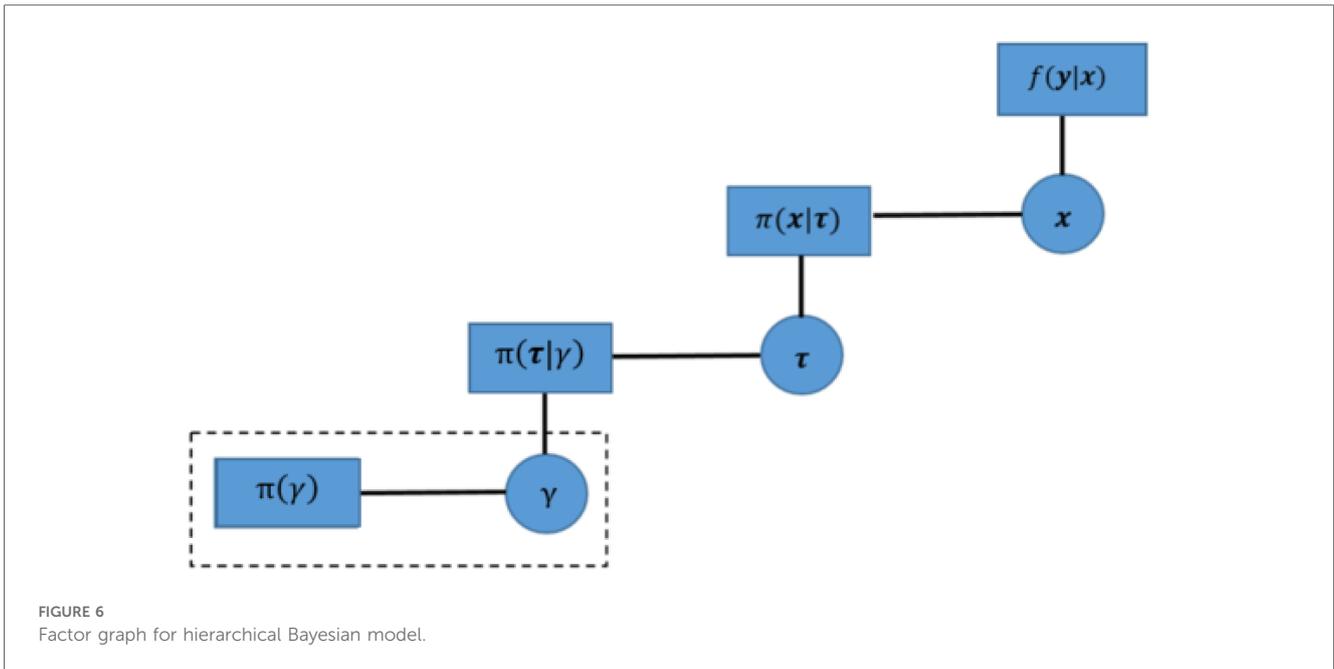
FIGURE 6
Factor graph for hierarchical Bayesian model.

more informative information available. Here, an exponential distribution is used for each τ_j independently but with a common rate parameter, γ, that is, $\tau_j|\gamma \sim \text{Exp}(\gamma)$. The value of γ is chosen to produce a long-tailed distribution to cover varied scenarios. As well as promoting smaller values of τ_j compared to a non-informative uniform distribution, it also avoids the need to impose an arbitrary upper range limit. The distribution for $\pi_{\tau|\gamma}(\boldsymbol{\tau}|\gamma)$:

$$\pi_{\tau|\gamma}(\boldsymbol{\tau}|\gamma) = \prod_{j=1}^{m} \gamma \exp(-\gamma \tau_j), \quad \gamma > 0. \tag{7}$$

The elements in $\boldsymbol{\tau} = \{\tau_j; j = 1, 2, , m\}$ are the collection of local prior variances and γ refers to the rate parameter in the hyper-prior distribution, a smaller value for the rate parameter indicates the more flat hyper-prior distribution.

3.2.4 Rate parameter in the hyper-prior distribution

The fundamental strategy for establishing a hierarchical Bayesian multilevel model is specifying prior distributions for each unknown parameter, which enables estimation for each parameter based on the other prior distributions from different levels. Therefore, it can incorporate more prior knowledge and hence improve estimation accuracy. However, once additional prior levels are involved in the model, it can result in a prolonged computation time and a high demand for supportive information. Figure 6 explains the multilevel structure in our hierarchical Bayesian model.

Now that an additional hyper-parameter, γ, has been introduced, further modeling must be considered. Due to a lack of supportive prior knowledge, it is common to utilize a uniform distribution or a Jacobian transformation of a uniform distribution. Another option is to introduce a conjugate prior

distribution which here would be the Gamma distribution with shape and scale parameters.

For the first method, the uniform prior allocates the same probability to each value within its defined range; it is subjective but hard to estimate values when the value is extremely small (23, 24). However, overly complex models have a high risk of poor performance (25). The use of a Gamma distribution unavoidably expands the number of unknown parameters. We expect, however, that the density peak of the Gamma distribution to be a small value and therefore, an exponential distribution with a small value for the rate parameter, θ, should be adequate: $\rho \exp(\Theta)$.

The result of these multilevel models, combining likelihood function in Equation 1, prior distribution in Equation 6, and hyper-prior distribution in Equation 7, produces the following posterior distribution:

$$
\begin{aligned}
f_{X,\tau,\gamma|Y}&(\boldsymbol{x}, \boldsymbol{\tau}, \gamma \mid \boldsymbol{y}) \\
&= f_{Y|X}(\boldsymbol{y} \mid \boldsymbol{x}) \, \pi_{X|\tau}(\boldsymbol{x}|\boldsymbol{\tau}) \, \pi_{\tau|\gamma}(\boldsymbol{\tau}|\gamma) \, \pi_\gamma(\gamma) \\
&\propto \prod_{i=1}^{n} \lambda_i^{p_i} \frac{\exp(-\lambda_i)}{p_i!} \left[\prod_{j=1}^{m} \frac{1}{(2\tau_j)} \exp\left(-\frac{\sum_{t \in \partial(j)} |x_j - x_t|}{\tau_j}\right) \cdot \gamma \exp(-\gamma \tau_j) \right] \\
&\quad \cdot \exp(-\theta \gamma),
\end{aligned}
\tag{8}
$$

where $\boldsymbol{X} = \{X_j, j = 1, 2, , m\}$ represents the unknown radionuclide distribution, $\boldsymbol{\tau} = \{\tau_j, j = 1, 2, , m\}$ are the locally adaptive prior variance parameters, γ is the hyper-parameter modeling $\boldsymbol{\tau}$, and $\boldsymbol{Y}: \{Y_i, i = 1, 2, , n\}$ is the observed data.

There is a concern about the potentially unreliable posterior estimations from the hierarchical model, especially for estimating small values. Hence, instead of estimating the hyper-prior parameter γ, we consider a calibration experiment where a series of values for γ, from 10^{-4} to 10^3, are used to investigate the trend in the mean squared error under different γ. In other words, instead of bringing another level of hyper-prior
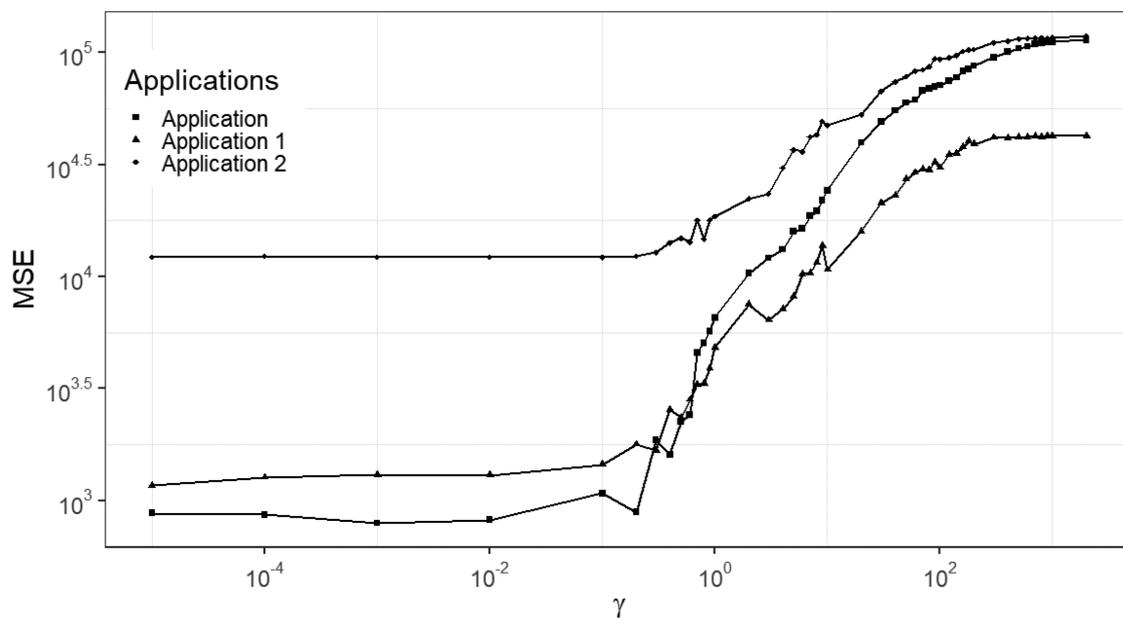
FIGURE 7
MSE of posterior estimation with fixed hyper prior parameter γ. Applications refer to the employment of simulation datasets (Y, Y_1, and Y_2), respectively.

distribution, γ is fixed, but based on the calibration experiments rather than requiring prior knowledge.

As shown in Figure 7, a similar increased pattern of MSE occurs among the three data applications with different levels of noise and blurring. The applications are the three designed examples. The estimation using small γ, smaller than 0.1 say, is robust and based on mean squared error performance. Combining the posterior estimation for γ with the hierarchical Bayesian model, we can conclude that a rate parameter γ around 10^{-2} is a robust hyper-prior parameter value for our current application and hence in what follows $\gamma = 10^{-2}$.

4 Results

4.1 Estimation strategy

Once the posterior distribution involving likelihood, prior, and hyper prior distribution is defined, as seen in Equation 8. A metropolis Hasting algorithm is used for the estimation. This is an example of the general Markov chain Monte Carlo (MCMC) approach that is able to handle complex distributions where other estimation methods fail. Details of the estimation process of our application can be found in Estimation algorithm of Markov chain Monte Carlo, Appendix Table A1. Apart from interest in the posterior estimation of the unknown radionuclide distribution, X, the locally-adaptive hyper-prior parameters τ must be estimated simultaneously. In general, the single global prior variance τ should capture the global variance between pixels. However, in the locally-adaptive model the hyper-prior

variances τ measure each pixel's variation within the corresponding neighbourhood, and hence we expect the elements $\{\tau = \tau_j, j = 1, 2, \ldots, m\}$ to be non-identical.

4.2 Posterior estimation

Figure 8 shows the posterior estimates for the three scenarios, with X, X_1, and X_2. For comparison, posterior estimates using the homogeneous model (left) are shown along with the estimates using the inhomogeneous, locally adaptive prior model (right). After incorporating prior flexibility, estimation improvements are apparent for deblurring and denoising, particularly in the second and third applications of simulation datasets with high-level blurring.

In the first application, where more accurate and sufficient first-hand information Y is available, the results from both models (the homogeneous and inhomogeneous models, respectively) are similar. However, in the third application, despite the foundational truth being the same as in the first example, the first-hand information Y_2 contains high levels of noise and blurring. In this case, the hierarchical Bayesian modeling successfully produces an image based on Y_2 that is closer to the truth compared to the outcomes from the homogeneous modeling.

In addition to posterior estimates of X, standard deviation and bias values are shown in Figures 9, 10.

For standard deviation, the results from the homogeneous models (left) show higher values than the inhomogeneous models (right), especially for the hot regions and smoothing background. The variation can still be found in high-contrast
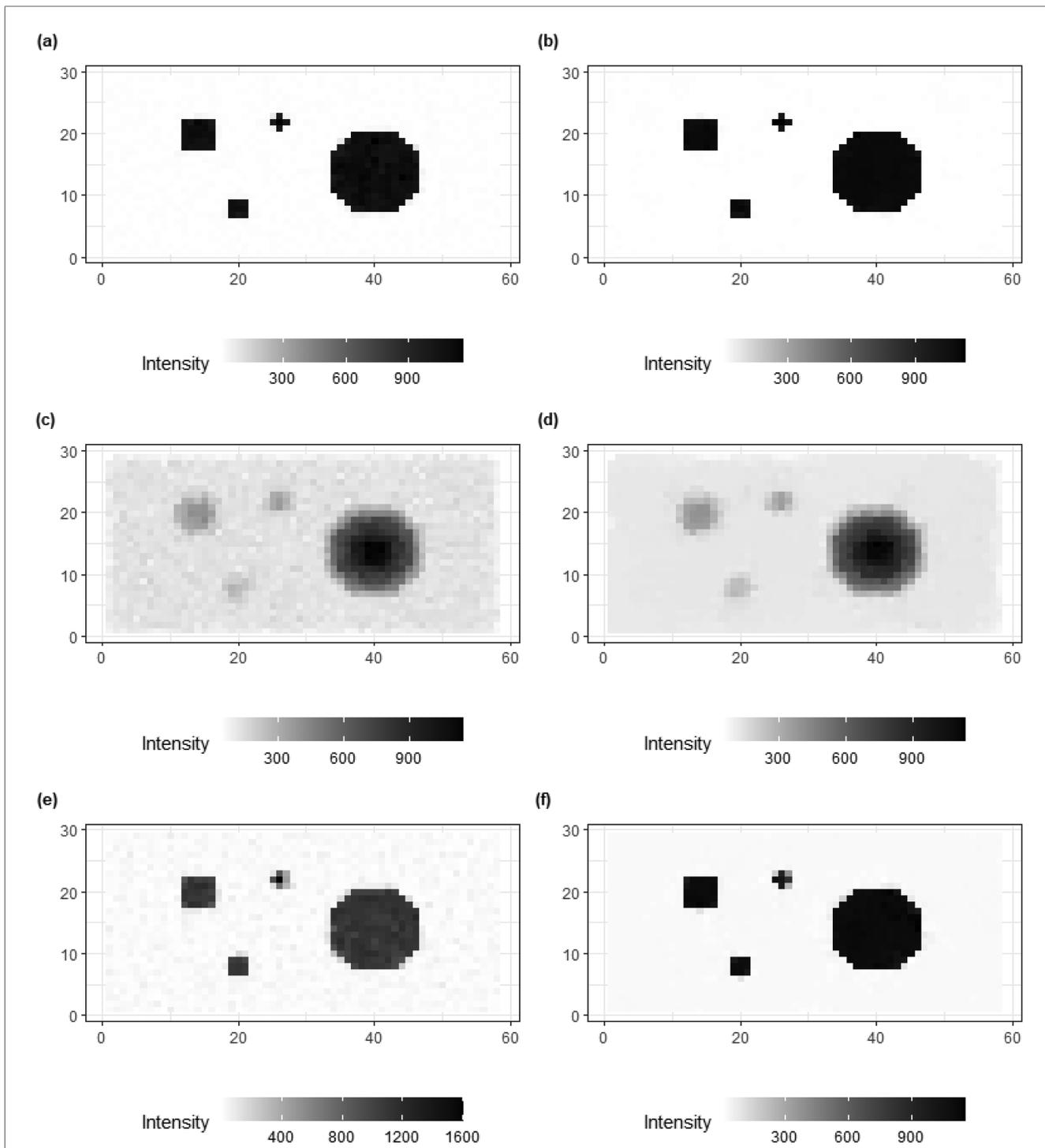
FIGURE 8
Comparisons of image processing for three simulations, X, X_1 and X_2, respectively. From the right side, (a,c,e) present the image processing with a global hyper-prior variance from the homogeneous Bayesian modeling. While (b,d,f) show the image processing under hierarchical Bayesian modeling with local hyper-prior variances.

edges while there is a reduction within the hot regions and smooth backgrounds where the pixel differences are assumed to be small.

The bias in Figure 10 provides clear additional evidence that the estimation accuracy of the locally adaptive model is improved compared to that of the homogeneous model. For instance, in Figure 10f, the bias around the high-contrast edge and

smoothing hot regions is hardly detected in the right image as compared to the image on the left side. Generally, the image patterns of standard deviation and bias on the right side have less information observed compared to the left side. In other words, the adjusted model captures the variation and bias within the posterior estimations.

FIGURE 9
Standard deviation of posterior estimations for true X, X_1 and X_2 in image pattern. (a,c,e) present the image patterns of standard deviation that come from the homogeneous Bayesian model with a global hyper-prior variance. (b,d,f) show the image patterns of standard deviation originate from hierarchical Bayesian modeling with local hyper-prior variances.
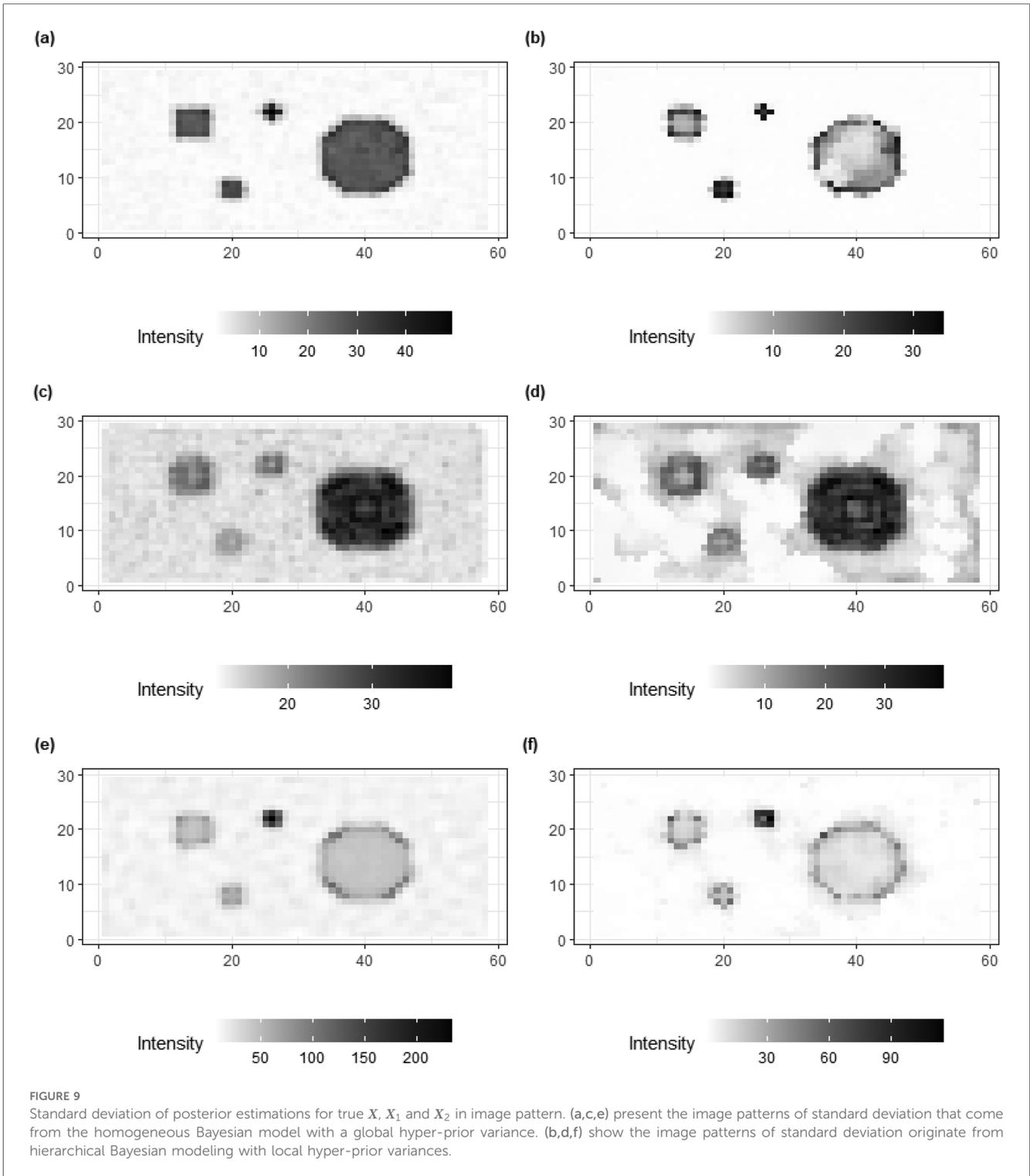
4.3 Comparisons of posterior pixel estimation

Figures 11a,b show box plots of three pixels from the background and three from the hot regions, respectively. The distribution in the blue box plots (left of each pair) is close to the red dashed line, representing the truth. In contrast, the ones

in the orange box plots (right of each pair) have significant bias. Overall, the Bayesian model with locally adaptive hyper-prior variance introduces estimation flexibility to realize a more accurate outcome in each application.

Figures 12–14 present comparisons of estimation pixels between the posterior containing homogeneous and inhomogeneous hyper-prior parameters in each simulation

FIGURE 10
Bias of posterior estimations for truth X, X_1, and X_2 in image pattern. (a,c,e) present the image patterns of bias that come from the homogeneous Bayesian model with a global hyper-prior variance. (b,d,f) show the image patterns of standard deviation originating from hierarchical Bayesian modeling with local hyper-prior variances.
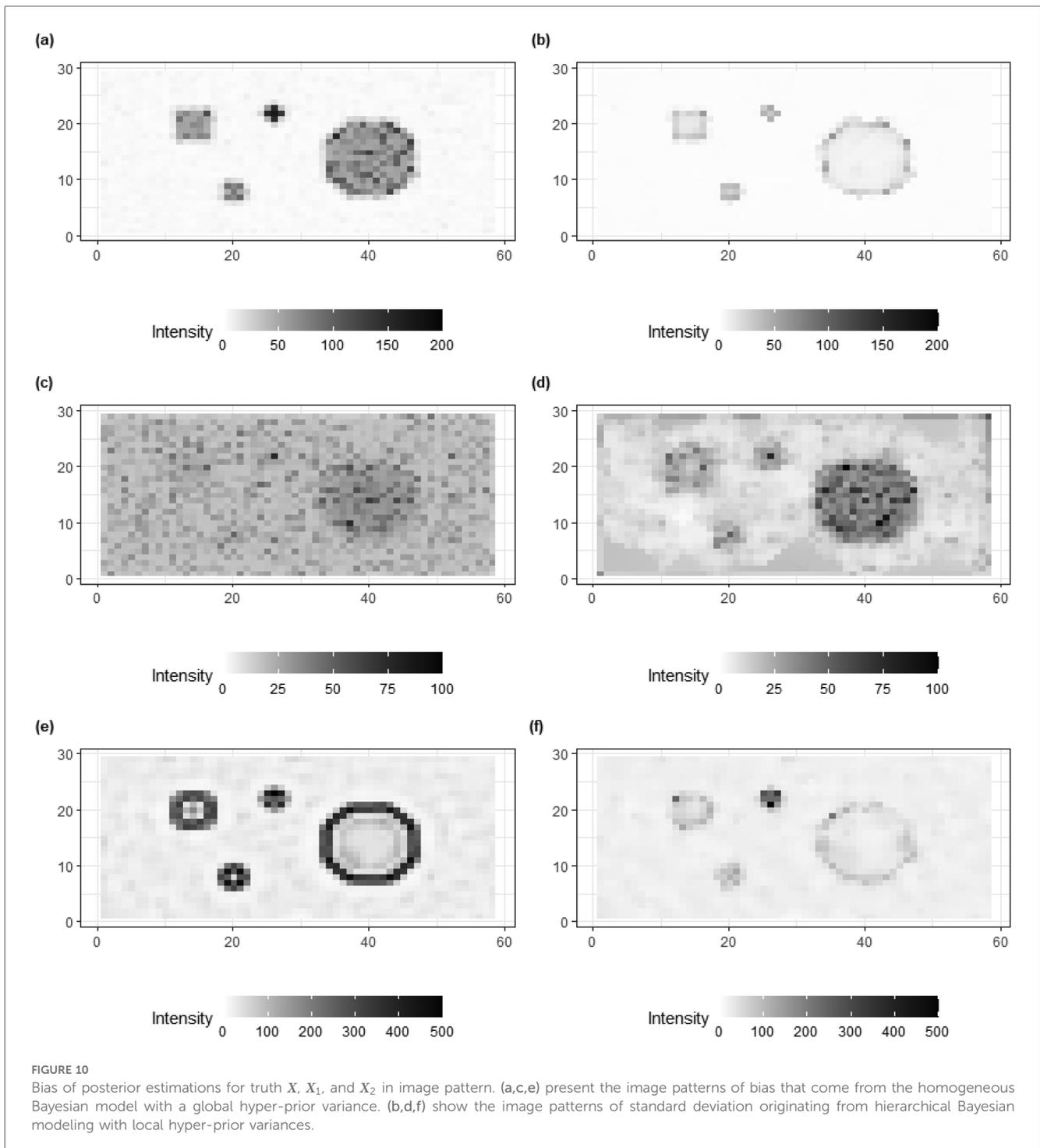
application (X, X_1, and X_2, respectively). In general, the estimations with a global hyper-parameter (left side) tend to have higher variations and broader credible intervals than those with locally adaptive hyper-parameters (right side), especially for the hot regions. Although the wider credible interval is more capable of covering variation in the values in comparison to the narrow credible interval, we noticed that there are some severe estimation fluctuations, especially within the smoothing pixel region, for example, in Figures 12, 13.

Furthermore, the value difference between the estimations from the locally adaptive model and the truth is smaller than the homogeneous ones. For instance, without sufficient first-hand information in the third simulation experience, as shown in Figure 14, the modeling with local hyper-prior variance produces a more accurate estimation as opposed to the one with global hyper-prior variance. Table 1 shows the corresponding estimation measurements of eight selected pixels in the 20th row in the third simulation application.
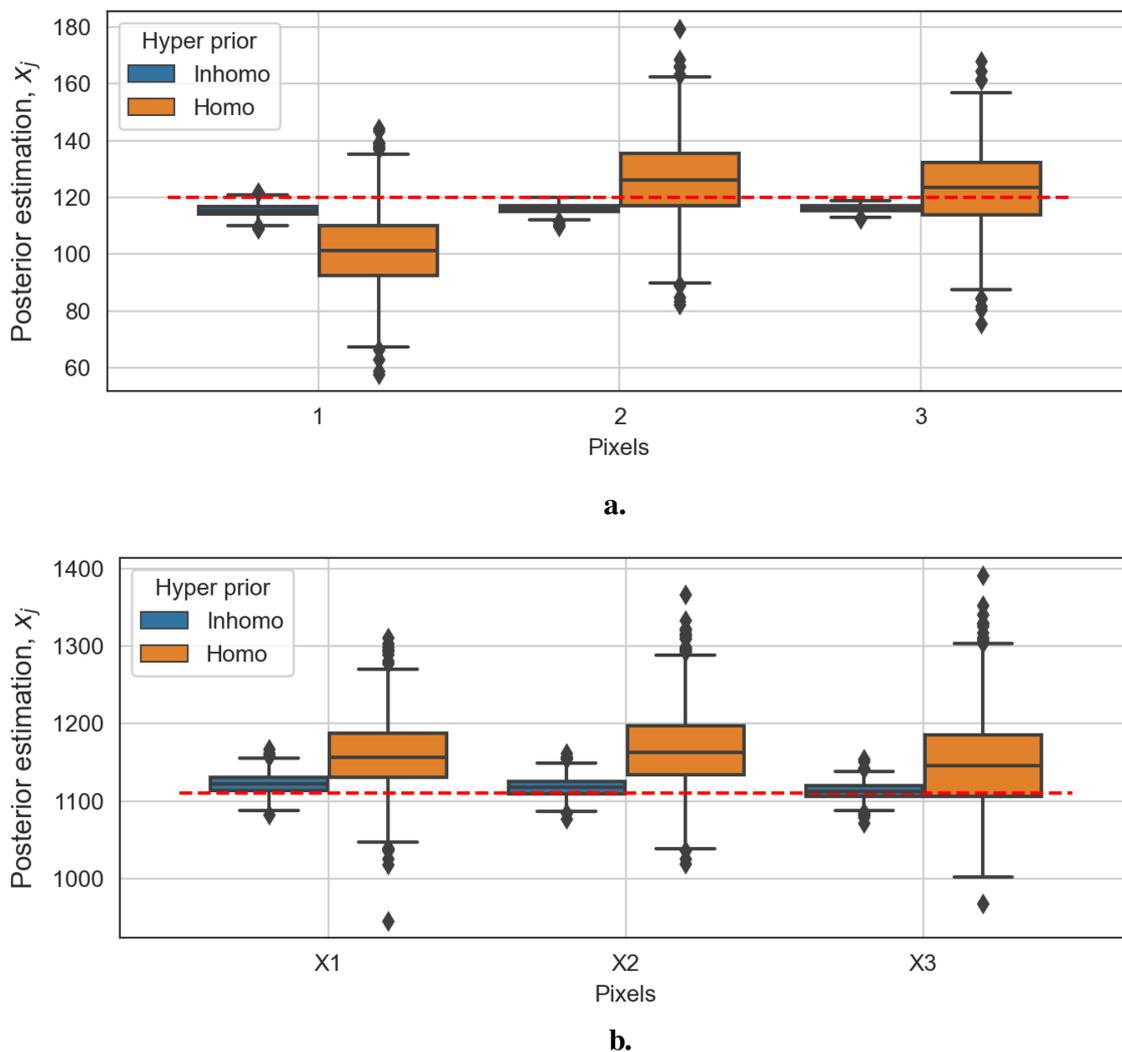
FIGURE 11
Box plots of posterior distributions. The box plots in blue show the results of the model with locally adaptive hyper-prior parameters, and the box plots in orange represent estimations from the homogeneous model. The red dashed line is the true pixel value. (a) Box plots of three background posterior distributions from the second example. (b) Box plots of three hot-region posterior distributions from the second example.

Overall, the Bayesian model with locally adaptive hyper-prior variance introduces estimation flexibility to realize a more accurate outcome in each application.

5 Modeling application in small animal imaging

In the previous simulation examples, locally adaptive Bayesian modeling proves the advantages of estimation accuracy. We now aim to apply this technique to images obtained from mouse scanning by using -eye^{TM} to confirm the conclusion obtained from the former sections. Figure 15a shows the image of a mouse injected with a technetium-99m labeled radiotracer acquired with -eye^{TM}, and Figure 15b

presents the correspondingly designed dataset for estimation application.

The results are presented in Figures 16b,c. Here, we assign the rate parameter $\gamma = 10^{-2}$ in the hyper-prior distribution. As in the previous examples, the estimation with the locally adaptive hyper-prior variance performs better compared to the homogeneous model in terms of deblurring and denoising; for instance, the smoothing edge (red circle) between the background and hot region is clearer in Figure 16c than in Figure 16b. In Figure 16d, the estimation of hyper-prior τ is displayed, showing the clear non-identical value distribution of τ. The high-dimension τ introduces flexibility when estimating pixel variance.

The locally adaptive Bayesian model with inhomogeneous hyper-prior parameters can specifically describe the
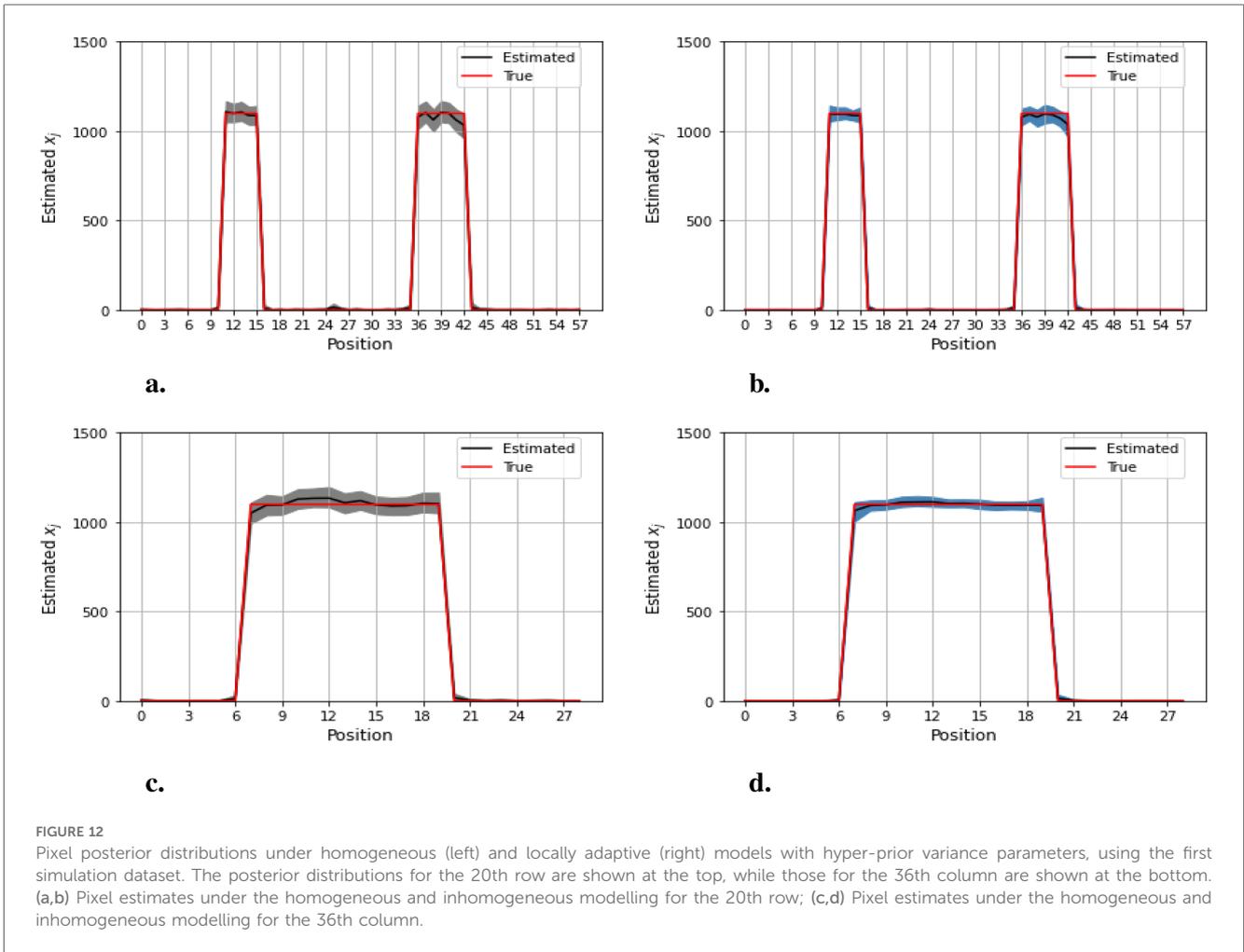
FIGURE 12
Pixel posterior distributions under homogeneous (left) and locally adaptive (right) models with hyper-prior variance parameters, using the first simulation dataset. The posterior distributions for the 20th row are shown at the top, while those for the 36th column are shown at the bottom. (a,b) Pixel estimates under the homogeneous and inhomogeneous modelling for the 20th row; (c,d) Pixel estimates under the homogeneous and inhomogeneous modelling for the 36th column.

probabilistic distribution of each unknown pixel x_j. Beyond pixel-wise posterior estimation, these inhomogeneous hyper-prior variances enabled a more detailed exploration of the outcomes. For instance, plotting the estimated hyper-prior variances directly reveals spatial information about the pixels. In conclusion, the locally adaptive Bayesian modeling constructs a hierarchical network that encompasses multiple levels of parameters. This network effectively integrates information from estimated parameters across different levels.

6 Tikhonov regularization and real image application

The application of Bayesian modeling in cylinder simulation datasets demonstrates advantages in terms of deblurring and denoising. Therefore, to confirm the applicability of Bayesian modeling, another real-world data application is required. Tikhonov regularization has been identified as useful as it introduces the homogeneous regularization term into ill-conditioned problems, specifically in the context of inverse problems (26–28). Therefore, a comparison between the

estimations from Bayesian modeling and Tikhonov regularization is necessary.

6.1 Tikhonov regularization comparison

Tikhonov regularization for medical image processing, which holds the linear relationship between observation image and real image, can be written as

$$\min\big(\|Y - AX\|_2^2 + \lambda\|LA\|_2^2\big),$$

where A is the transformation matrix, and Y and X are the observation dataset and real unknown dataset, respectively. The regularization parameter λ controls the trade-off between the model fitness and the regularization term. The regularization matrix L contains the prior information about the solutions. Here, we employ the identity matrix as the regularization matrix L because of the lack of supportive prior information.

In the context of the regularization parameter λ, the criterion of cross-validation has been applied in various regularization algorithms, including Tikhonov regularization. Cross-validation
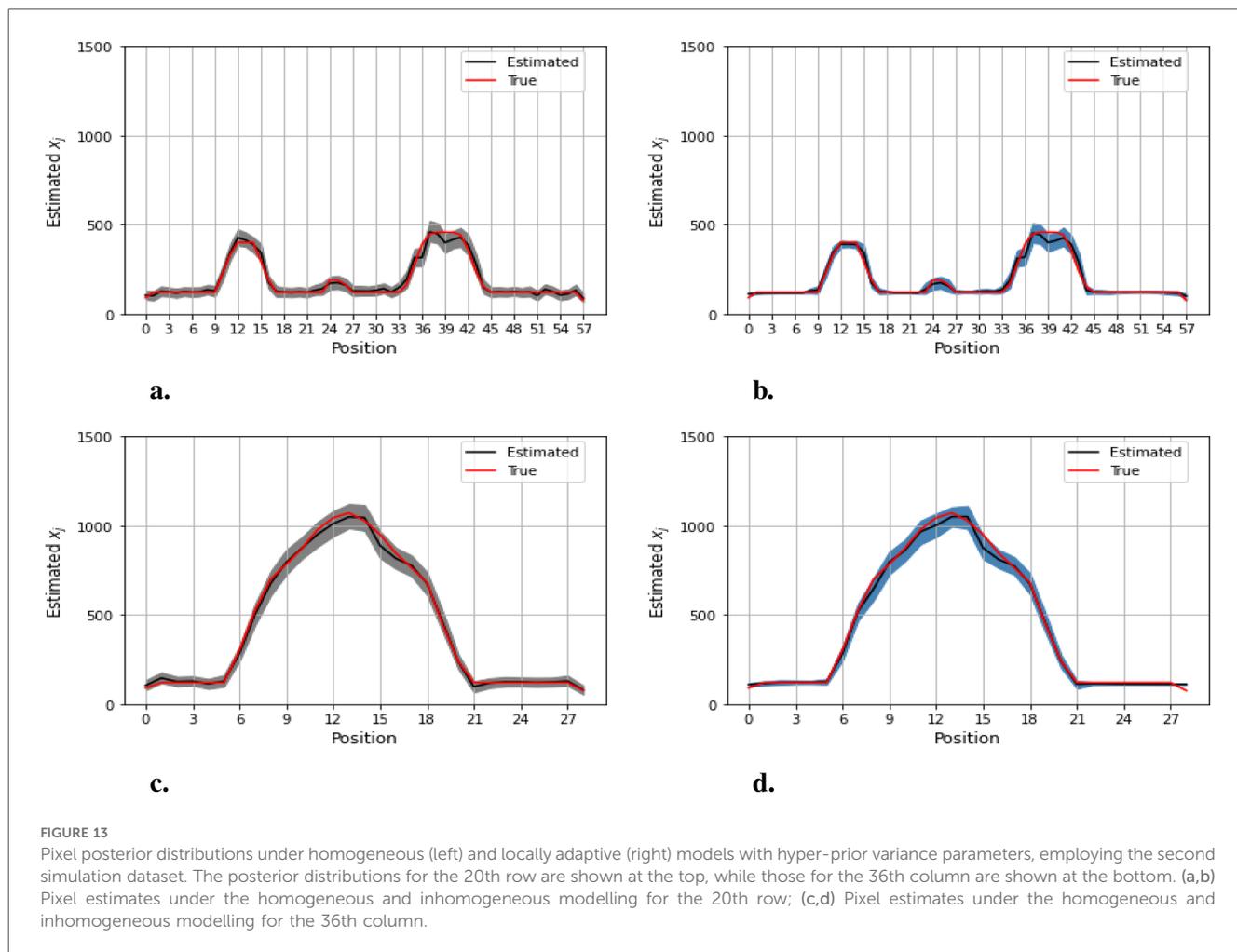
FIGURE 13
Pixel posterior distributions under homogeneous (left) and locally adaptive (right) models with hyper-prior variance parameters, employing the second simulation dataset. The posterior distributions for the 20th row are shown at the top, while those for the 36th column are shown at the bottom. **(a,b)** Pixel estimates under the homogeneous and inhomogeneous modelling for the 20th row; **(c,d)** Pixel estimates under the homogeneous and inhomogeneous modelling for the 36th column.

selects the optimum regularization parameter λ by identifying the minimum estimation residuals.[1] The estimation outcome from Tikhonov regularization, using observation Y from within the first simulation dataset, is shown in Figure 17. The left side displays the observation of the first simulation dataset, while the right side presents the estimation from Tikhonov regularization.

Compared to the Bayesian application shown in Figure 8, the blurring in Figure 17b can still be detected around the high-contrast area between the background and hot region. Hence, the estimation for real image X is not accurate, since regularization applies to the whole information not only noise but also pixel values. In other words, the smoothing effect from regularization applies globally to pixels within both background and hot regions simultaneously, regardless of varied pixel densities. The estimated pixels with a large value of regularization (represented in green) are smoother than those with small regularization

(represented in blue). Furthermore, some non-negative pixels from the background are unavoidably transformed into negatives after applying regularization.

Similarly, the estimations of specific columns and rows within the pixel matrix from Bayesian modeling and Tikhonov regularization are presented in Figure 18. Compared to the Bayesian modeling, Tikhonov regularization is only based on the pixel point estimation without consideration of the pixels' environment. Furthermore, unlike estimations from Markov chain Monte Carlo within the Bayesian framework, the distribution of estimated pixels and the quantified information, including the confidence intervals of estimations, are not available.

6.2 Real medical image application

Here, the employed medical image for the mouse kidney was obtained by using a dimercaptosuccinic acid scan (DMSA). Compared to the -eye^{TM} camera, the DMSA scan with technetium-99 m labeled radiotracer is well-known for its valuable capability in identifying patients' kidney shape and location.

The information of the region of interest, where the kidneys are located, is clearer in the reprocessed image in Figure 19b compared to the observed image in Figure 19a.

---

[1]The estimation algorithm for regularization parameter "cv.glmnet" and the application of Tikhonov regularization "glmnet" can be find in R package "glmnet."

FIGURE 14
Pixel posterior distributions under homogeneous (left) and locally adaptive (right) models with hyper-prior variance parameters, employing the third simulation dataset. The posterior distributions for the 20th row are shown at the top, while those for the 36th column are shown at the bottom. (a), (b) Pixel estimates under the homogeneous and inhomogeneous modelling for the 20th row; (c), (d) Pixel estimates under the homogeneous and inhomogeneous modelling for the 36th column.

TABLE 1 The list includes estimation measurements.

Position	ucl.H	ucl.In	lci.H	lci.In	mean.H	mean.In
1	45.36	20.60	0.90	18.47	15.80	19.58
4	32.58	20.09	0.48	16.81	11.46	18.67
8	35.69	23.06	0.59	13.86	12.18	18.36
14	1,219.05	1,136.55	1,045.48	1,086.63	1,132.85	1,111.66
18	79.04	30.05	2.13	15.27	29.53	21.53
22	62.27	24.30	1.66	16.50	23.93	20.46
24	48.84	25.52	0.55	16.01	16.94	20.55
26	123.85	43.67	3.40	9.73	44.13	24.86

"H" indicates the posterior estimation from Bayesian modeling with the global LMRF, while "In" indicates the posterior estimation from Bayesian modeling with locally adaptive LMRF. The outcomes are stored to two decimal places.

FIGURE 15
Scan of a mouse using -eye[TM]: (a) real scan and (b) a simulated dataset.

a. Degraded observation image from the simulated dataset

b. Posterior mean using a global prior variance

c. Posterior mean of pixels using locally-adaptive Bayesian modelling

d. Posterior mean of locally adaptive prior parameters

FIGURE 16

Image processing under the Bayesian modeling. (a) shows the simulation image with noise and blurring. (b) displays image processing under the Bayesian modeling with homogeneous hyper-prior parameter τ. (c) displays image processing under the Bayesian modeling with inhomogeneous hyper-prior parameter τ. (d) shows the posterior estimation of locally adaptive hyper-prior variances in the image pattern.

a.

b.

FIGURE 17

Observation image and the corresponding estimation from Tikhonov regularization with the optimum regulation. (a) Observation Y. (b) Tikhonov regularization.

FIGURE 18
Estimation comparison between Bayesian modeling and Tikhonov regularization. The estimation from Bayesian modeling highlighted in the dark color has a credible interval in gray. The estimations from Tikhonov regularization are presented with two regularization options: the green line represents the optimum regularization defined by the cross-validation method, while the blue line represents manual regularization with $\lambda=0.1$ applied. Here, the pixel estimations for the 20th row are shown at the top, while those for the 36th column are shown at the bottom. (a) Estimation comparison between Bayesian modeling and Tikhonov regularization I. (b) Estimation comparison between Bayesian modeling and Tikhonov regularization II.

FIGURE 19
A real application in a medical image using DMSA (left) and the corresponding posterior estimations from Bayesian modeling with LMRF prior distribution (right). **(a)** Observed image from DMSA. **(b)** Estimation from Bayesian modeling.
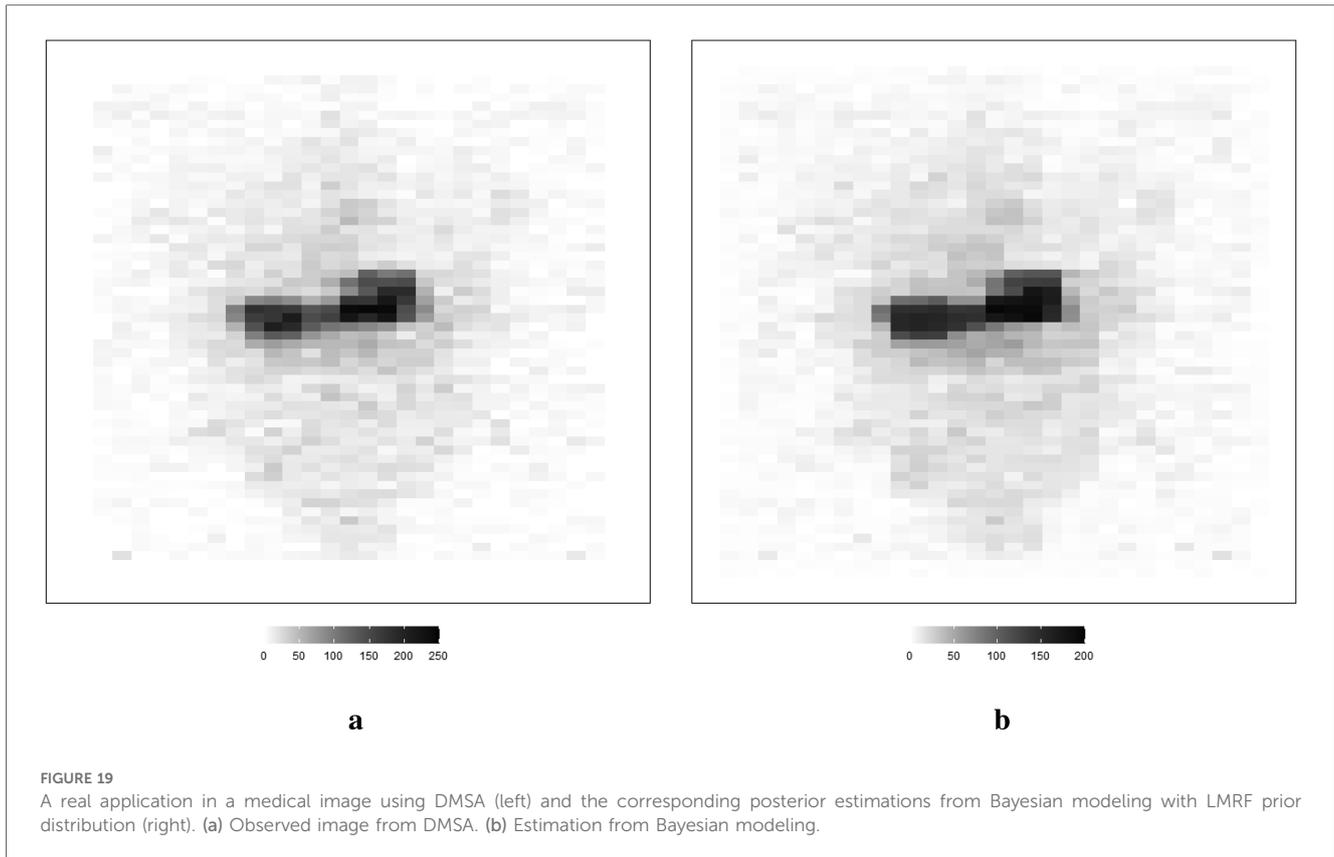
7 Conclusions

We extended the hierarchical Bayesian model for image processing by introducing the locally adaptive hyper-prior variance τ, replacing a single homogeneous hyper-prior variance τ. The locally adaptive model adjusted the hyper-prior variances based on the different local spatial conditions, effectively allowing the hyper-prior variances to vary for each location estimation. This adaptation provided the model with greater flexibility in estimation, subsequently improving accuracy. In our exploration of hyper-prior parameter estimation, we found that weakly informative prior distributions, such as a relatively flat exponential distribution, performed more efficiently compared to non-informative priors. This was evidenced by higher convergence rates and lower estimation correlations.

The locally adaptive Bayesian model with inhomogeneous hyper-prior parameters can specifically describe the probabilistic distribution of each unknown pixel. Beyond pixel-wise posterior estimation, these inhomogeneous hyper-prior variances enabled a more detailed exploration of the outcomes. For instance, plotting the estimated hyper-prior variances directly revealed spatial information about the pixels. In conclusion, the locally adaptive Bayesian approach constructs a hierarchical model that encompasses multiple levels of parameters. This approach effectively integrates information from estimated parameters across different levels, leading to improved image estimation. Consequently, there is the potential for enhancements in quantification, diagnosis, and treatment monitoring in medical imaging applications.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

MZ: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. RGA: Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. CT: Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

CT declares collaboration with BIOEMTECH where he undertook a secondment of about 6 months in 2018 sponsored by the EU Horizon 2020 project: Vivoimag (https://vivoimag.eu/). The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnume.2025.1508816/full#supplementary-material

References

1. Marquis H, Willowson KP, Schmidtlein CR, Bailey DL. Investigation and optimization of PET-guided SPECT reconstructions for improved radionuclide therapy dosimetry estimates. *Front Nucl Med*. (2023) 3:1124283. doi: 10.3389/fnume.2023.1124283

2. Winkler G. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*. Berlin, Heidelberg: Springer (2003).

3. Hansen TM, Mosegaard K, Holm S, Andersen FL, Fischer BM, Hansen AE. Probabilistic deconvolution of pet images using informed priors. *Front Nucl Med*. (2023) 2:1028928. doi: 10.3389/fnume.2022.1028928

4. Schramm G, Thielemans K. Parallelproj—an open-source framework for fast calculation of projections in tomography. *Front Nucl Med*. (2024) 3:1324562. doi: 10.3389/fnume.2023.1324562

5. Noviyanti L. An empirical Bayesian and Buhlmann approach with non-homogeneous Poisson process. In: *AIP Conference Proceedings*. AIP Publishing (2015). Vol. 1692.

6. Olalude OA, Muse BO, Alaba OO. Informative prior on structural equation modelling with non-homogenous error structure. *F1000Res*. (2022) 11:494. doi: 10.12688/f1000research.108886.2

7. Zhang M, Aykroyd RG, Tsoumpas C. Mixture prior distributions and Bayesian models for robust radionuclide image processing. *Front Nucl Med*. (2024) 4:1380518. doi: 10.3389/fnume.2024.1380518

8. Aykroyd R. Bayesian estimation for homogeneous and inhomogeneous Gaussian random fields. *IEEE Trans Pattern Anal Mach Intell*. (1998) 20:533–9. doi: 10.1109/34.682182

9. Kukačka J, Metz S, Dehner C, Muckenhuber A, Paul-Yuan K, Karlas A, et al. Image processing improvements afford second-generation handheld optoacoustic imaging of breast cancer patients. *Photoacoustics*. (2022) 26:100343. doi: 10.1016/j.pacs.2022.100343

10. Marsi S, Bhattacharya J, Molina R, Ramponi G. A non-linear convolution network for image processing. *Electronics*. (2021) 10:201. doi: 10.3390/electronics10020201

11. Trull AK, Van Der Horst J, Palenstijn WJ, Van Vliet LJ, Van Leeuwen T, Kalkman J. Point spread function based image reconstruction in optical projection tomography. *Phys Med Biol*. (2017) 62:7784. doi: 10.1088/1361-6560/aa8945

12. Hutt A, Wahl T. Poisson-distributed noise induces cortical γ-activity: explanation of γ-enhancement by anaesthetics ketamine and propofol. *J Phys Complex*. (2022) 3:015002. doi: 10.1088/2632-072X/ac4004

13. Kirkpatrick JM, Young BM. Poisson statistical methods for the analysis of low-count gamma spectra. *IEEE Trans Nucl Sci*. (2009) 56:1278–82. doi: 10.1109/TNS.2009.2020516

14. Maestrini L, Aykroyd RG, Wand MP. A variational inference framework for inverse problems. *arXiv* [Preprint]. *arXiv:2103.05909* (2021).

15. Georgiou M, Fysikopoulos E, Mikropoulos K, Fragogeorgi E, Loudos G. Characterization of "γ-eye": a low-cost benchtop mouse-sized gamma camera for dynamic and static imaging studies. *Mol Imaging Biol*. (2017) 19:398–407. doi: 10.1007/s11307-016-1011-4

16. Al-Gezeri SM, Aykroyd RG. Spatially adaptive Bayesian image reconstruction through locally-modulated Markov random field models. *Braz J Probab Stat*. (2019) 33:498–519. doi: 10.1214/18-BJPS399

17. Beyerer J, Heizmann M, Sander J, Gheta I. Bayesian methods for image fusion. *Image Fusion Algorithms Appl*. (2011) 157.

18. Kozlov O. Gibbsian description of point random fields. *Theory Probab Appl*. (1977) 21:339–56. doi: 10.1137/1121038

19. Gibbs JW. On the equilibrium of heterogeneous substances. (1879).

20. Willard GJ. *Scientific Papers of J. Willard Gibbs*. London & New York: Longmans, Green and Company (1901). Vol. 1.

21. Cross GR, Jain AK. Markov random field texture models. *IEEE Trans Pattern Anal Mach Intell*. (1983) PAMI-5:25–39. doi: 10.1109/TPAMI.1983.4767341

22. Kato Z. Markov random fields in image segmentation. *Found Trends Signal Process*. (2011) 5:1–155. doi: 10.1561/2000000035

23. Gelman A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal*. (2006) 1:515–34. doi: 10.1214/06-BA117A

24. Link WA. A cautionary note on the discrete uniform prior for the binomial n. *Ecology*. (2013) 94:2173–9. doi: 10.1890/13-0176.1

25. Adkison MD. Drawbacks of complex models in frequentist and Bayesian approaches to natural-resource management. *Ecol Appl*. (2009) 19:198–205. doi: 10.1890/07-1641.1

26. Chung J, Español MI. Learning regularization parameters for general-form Tikhonov. *Inverse Probl*. (2017) 33:074004. doi: 10.1088/1361-6420/33/7/074004

27. Golub GH, Von Matt U. *Tikhonov Regularization for Large Scale Problems*. Citeseer: Computer Science Department, Stanford University (1997).

28. Wen Y-W, Chan RH. Using generalized cross validation to select regularization parameter for total variation regularization problems. *Inverse Probl Imaging*. (2018) 12:1103–20. doi: 10.3934/ipi.2018046

Appendix

Estimation algorithm of Markov chain Monte Carlo

The hyper-parameter τ is a global variable, and we assume the distribution for a global hyper-parameter τ is a uniform distribution. However, $\tau = \{\tau_j, j = 1, 2, , m\}$ becomes a vector of unknown parameters. Therefore, there are corresponding changes in terms of distribution $\pi_{\tau|\gamma}(\tau \mid \gamma)$ and estimation process: MCMC. We initially define the Gamma hyper-prior distribution with artificial values $\xi = 1$ and $\theta = 2$: $\gamma \sim \text{Gamma}(\xi = 1, \theta = 2)$. If we want to estimate τ for the kth iteration in the Markov chain Monte Carlo and update unknown parameters, follow the steps below:

Once the chain has burned in, a process to drop a bunch of initial iterations whose value probabilities are low, the estimation

for each parameter is supposed to be the sampling mean, we seek, i.e., $E(\gamma) \approx (\sum_k \gamma^{(k)})/k$.

Estimation measurement

The estimation X_j can be assessed by averaging the application outcomes for the last K stationary MCMC iterations. The most popular and common indexes for accurate measurement are the MSE, bias, and standard deviation (SD). Furthermore, residual sum squares (RSS), the modeling fitness measurement, is widely applied among model comparisons. For the subset of the whole parameters, the forms of measurement are seen in Table A2.

The entire image is of size m, while any image sub-region is of size R, i.e., $R \leq m$. The average MCMC estimation is $E(x_j)$ and $\hat{x}_j^K$ refers to the estimation of corresponding pixels in the Kth iteration. The denominator in each of the expressions varies based on the number of pixels considered.

TABLE A1 MCMC for modeling with locally adaptive hyper-parameter γ.

Algorithm MCMC for modelling with locally adaptive hyper-prior parameters (iteration k)
Input: A list of initial values $\{X_1 = x_1^{(0)}, X_2 = x_2^{(0)}, , X_m = x_m^{(0)}\}$;
A list of initial values $\{\tau_1 = \tau_1^{(0)}, \tau_2 = \tau_2^{(0)}, , \tau_m = \tau_m^{(0)}\}$;
An updated positive constant $\sigma^{(k)}, \sigma_\tau^{(k)}, \sigma_\gamma^{(k)}$.
For $j = \{1, 2, , m\}$
1. Propose a new value $x_j^{(k)} \sim N(x_j^{(k-1)}, (\sigma^{(k)})^2)$; **if and only if** $x_j^k \geq 0$
2. Generate $\mu \sim \text{unif}(0, 1)$
3. Accept x_j^k with probability
$\alpha = \min\left(1, \frac{f_{X|Y,\tau}(x_1^{(k-1)}, x_2^{(k-1)}, , x_j^{(k)}, x_m^{(k-1)}|Y, \tau^{(k-1)})}{f_{X|Y,\tau}(x_1^{(k-1)}, x_2^{(k-1)}, , x_j^{(k-1)}, x_m^{(k-1)}|Y, \tau^{(k-1)})}\right)$
4. Compare the μ with the calculated α,
5. **if**: $\mu \leq \alpha$ **then**
6. Accept the proposal value $x_j = x_j^{(k)}$
7. **else** $x_j = x_j^{(k-1)}$
end update x
8. Propose a new candidate value $\tau_j^k \sim N(\tau_j^{k-1}, (\sigma_\tau^{(k)})^2)$; **if and only if** $\tau_j^k \geq 0$
9. Generate $\mu_1 \sim \text{unif}(0, 1)$.
10. Accept $\tau^{(k)}$ with probability
$\alpha_1 = \min\left(1, \frac{f_{\tau|X}(\tau_1^{(k-1)}, \tau_2^{(k-1)}, , \tau_j^{(k)}, , \tau_m^{(k-1)}|x^k, \gamma^{(k-1)})}{f_{\tau|X}(\tau_1^{(k-1)}, \tau_2^{(k-1)}, , \tau_j^{(k-1)}, , \tau_m^{(k-1)}|x^k, \gamma^{(k-1)})}\right),$
11. **if**: $\mu_1 \leq \alpha_1$ **then**
12. Accept the proposal value $\tau_j = \tau_j^{(k)}$,
13. **else** $\tau_j = \tau_j^{(k-1)}$
end update τ
14. Propose a new candidate value $\gamma^k \sim N(\gamma^{k-1}, (\sigma_\tau^{(k)})^2)$; **if and only if** $\gamma^k \geq 0$
15. Generate $\mu_2 \sim \text{unif}(0, 1)$.
16. Accept $\gamma^{(k)}$ with probability
$\alpha_2 = \min\left(1, \frac{f_{\gamma|X,\tau}(\gamma^k|x^k, \tau^{(k)})}{f_{\gamma|X,\tau}(\gamma^{k-1}|x^k, \tau^{(k)})}\right),$
17. **if**: $\mu_2 \leq \alpha_2$ **then**
18. Accept the proposal value $\gamma = \gamma^{(k)}$,
19. **else** $\gamma = \gamma^{(k-1)}$
end update γ
20. **end if**
Repeat the above steps until it receives enough sampling size.

TABLE A2 List of statistical measurements.

Measurement	Equation
Absolute bias	$(\sum_{j=1}^m \mid x_j - E(x_j) \mid)/m$
Regional absolute bias	$\left(\sum_{j \in R} \mid x_j - E(x_j) \mid\right)/R$
Standard deviation	$\sqrt{\sum_{k=1}^K \sum_{j=1}^m \mid x_j^{(k)} - E(x_j) \mid^2 / (K \cdot m)}$
Regional standard deviation	$\sqrt{\sum_{k=1}^K \sum_{j \in R} \mid x_j^{(k)} - E(x_j) \mid^2 / (K \cdot R)}$
Mean square error	$\sqrt{\sum_{k=1}^K \sum_{j=1}^m \mid x_j^{(k)} - E(x_j) \mid^2 / (K \cdot m)}$
Regional mean square error	$\sqrt{\sum_{k=1}^K \sum_{j \in R} \mid x_j^{(k)} - E(x_j) \mid^2 / (K \cdot R)}$
Residual sum of squares	$\sum_{k=1}^K \sum_{j=1}^m \left(y_j - y_j^{(k)}\right)^2 / (K \cdot m)$
Regional residual sum of squares	$\sum_{k=1}^K \sum_{j \in R} \left(y_j - y_j^{(k)}\right)^2 / (K \cdot R)$