

مجله علوم آماری، پاییز و زمستان ۱۳۹۵

جلد ۱۰، شماره ۲، ص ۲۳۳-۲۶۰

DOI: 10.18869/acadpub.jss.10.2.233

## تحلیل بیزی تقریبی داده‌های فضایی - زمانی با استفاده از یک میدان تصادفی مارکوفی گاوسی

فاطمه حسینی، الهام همایون فال

گروه آمار، دانشگاه سمنان

تاریخ دریافت: ۱۳۹۳/۱۱/۳ تاریخ آخرین بازنگری: ۱۳۹۴/۸/۱۸

**چکیده:** برای مدل‌بندی پاسخ‌های فضایی که در طول زمان مشاهده می‌شوند گاهی از مدل‌های سلسله‌مراتبی فضایی - زمانی استفاده می‌شود که در آن ساختار همبستگی فضایی - زمانی داده‌ها توسط یک میدان تصادفی پنهان گاوسی در نظر گرفته می‌شود. در این مقاله برای تحلیل این مدل‌ها، ابتدا رهیافت بیزی معمولی ارائه می‌شود. اما به دلیل پیچیدگی توزیع‌های پسین و توزیع‌های شرطی کامل و استفاده از نمونه‌های مونت کارلویی، زمان محاسبات بسیار طولانی است. برای رفع این مشکل میدان تصادفی پنهان گاوسی، به صورت یک میدان تصادفی مارکوفی در نظر گرفته می‌شود که برای تولید آن چار معادلات دیفرانسیل جزئی تصادفی می‌توان استفاده کرد. سپس از روش بیز تقریبی و تقریب لاپلاس آشیانی جمع‌بسته برای به‌دست آوردن یک تقریب دقیق از توزیع‌های پسین و استنباط‌ها پیرامون مدل استفاده می‌شود. در نهایت مجموعه داده‌های میزان بارندگی استان سمنان با مدل و روش‌های ارائه شده مورد تحلیل قرار می‌گیرد.

---

آدرس الکترونیک مسئول مقاله: فاطمه حسینی، [fatemeh.hoseini@semnan.ac.ir](mailto:fatemeh.hoseini@semnan.ac.ir)  
کد موضوع‌بندی ریاضی (۲۰۱۰): ۶۰H۱۵، ۶۲F۱۵

واژه‌های کلیدی : داده‌های فضایی-زمانی، میدان تصادفی مارکوفی گاوسی، تقریب لاپلاس آشیانی جمع‌بسته، معادلات دیفرانسیل جزئی تصادفی.

## ۱ مقدمه

گاهی در عمل تحلیل داده‌هایی مورد نظر است که در طول زمان و برحسب مکان قرارگیری در فضای مورد مطالعه وابسته هستند. برای تحلیل این نوع داده‌ها از مدل سلسله مراتبی فضایی-زمانی استفاده می‌شود. در این مدل متغیرهای پاسخ با استفاده از یک میدان تصادفی گاوسی<sup>۱</sup> مدل‌بندی می‌شوند، خطای اندازه‌گیری به صورت یک میدان تصادفی گاوسی نوفه سفید فرض می‌شود و وابستگی فضایی-زمانی با یک میدان پنهان تصادفی فضایی-زمانی به صورت یک میدان حالت که با زمان تغییر می‌کند و اتو رگرسیو مرتبه اول است، مدل‌بندی می‌شود این میدان تصادفی پنهان معمولاً گاوسی با تابع کوواریانس ماترن<sup>۲</sup> در نظر گرفته می‌شود.

مطالعات زیادی پیرامون این مدل‌ها انجام شده است. از جمله مطالعات انجام شده در یک چارچوب بیزی می‌توان به ویکل و همکاران (۱۹۹۸)، ویکل (۲۰۰۳)، بانرجی و همکاران (۲۰۰۴)، ریواز و همکاران (۱۳۸۶)، کوکچی و همکاران (۲۰۰۷)، کرسی و ویکل (۲۰۱۱)، کاملتی و همکاران (۲۰۱۱) و ساهو (۲۰۱۲) اشاره نمود.

یکی از موضوعات اصلی در مطالعه داده‌های فضایی-زمانی، برآورد پارامترهای مدل، برآورد متغیرهای پنهان و پیش‌گویی متغیر پاسخ نامعلوم در موقعیت‌ها و زمان‌های جدید براساس داده‌های مشاهده شده در موقعیت‌ها و زمان‌های معلوم می‌باشد. در تحلیل بیزی سلسله مراتبی مدل‌های گاوسی پنهان به دلیل پیچیدگی توزیع‌های پسین و عدم وجود شکل‌های بسته برای آن‌ها وجود تعداد زیادی متغیر پنهان و پارامتر، استفاده از رهیافت بیزی معمولی و روش‌های زنجیر مارکوف مونت کارلویی بسیار وقت‌گیر و دارای مشکلات همگرایی می‌باشد، به خصوص وقتی که مطالعه یک سری داده‌های بزرگ فضایی-زمانی مدنظر باشد، ماتریس کوواریانس

<sup>۱</sup> Gaussian Random field

<sup>۲</sup> Matern Covariance Function

مدل غیرتُنک<sup>۳</sup> می شود، یعنی بیشتر در ایه هایش غیر صفر می باشند. به دست آوردن تابع کوواریانس غیرتُنک  $\sigma^2 C(\cdot, \cdot)$  سخت شامل هزینه های محاسباتی زیادی است، به طوری که استفاده از روش های مونت کارلویی بسیار زمان بر و گاهی غیرممکن است، زیرا باید تابع کوواریانس غیرتُنک برای هر تکرار به دست آورده شود. بنابراین اگرچه الگوریتم های مونت کارلویی، روش هایی مفید برای به دست آوردن برآوردها و تحلیل بیزی پارامترهاست، اما به دلیل داشتن مشکلاتی مانند طولانی بودن زمان محاسبات و گاهی عدم همگرایی در مدل های فضایی پیچیده سلسله مراتبی، رو و مارتینو (۲۰۰۷) روش تقریب لاپلاس آشیانی جمع بسته<sup>۴</sup> (INLA) را برای تحلیل مدل های پیچیده فضایی معرفی کردند و نشان دادند این روش ضمن حفظ دقت برآورد پارامترها سرعت محاسبات را نیز افزایش می دهد. ایدسویک و همکاران (۲۰۰۹) از تقریب معرفی شده توسط رو و مارتینو (۲۰۰۷) و رو و همکاران (۲۰۰۹) استفاده کردند و به تحلیل بیزی تقریبی مدل های آمیخته خطی تعمیم یافته فضایی پرداختند. حسینی و همکاران (۲۰۱۱) و حسینی و محمدزاده (۲۰۱۲) با به کار بردن توزیع چوله نرمال بسته برای متغیرهای پنهان تحلیل بیزی و بیزی تقریبی این مدل ها را با الگوریتم های مونت کارلوی زنجیر مارکوفی و رهیافت INLA بیان و مقایسه نمودند. قلی زاده و همکاران (۲۰۱۳) و (۱۳۹۲) از تقریب INLA استفاده نمودند و به تحلیل بیزی تقریبی مدل های رگرسیون جمعی ساختاری پرداختند و داده های جرم شهر تهران را با این مدل و روش مورد تحلیل قرار دادند. کامالنتی و همکاران (۲۰۱۳) با در نظر گرفتن میدان تصادفی مارکوفی گاوسی<sup>۵</sup> (GMRF) و رهیافت بیز تقریبی معرفی شده توسط رو و همکاران (۲۰۰۹) به مطالعه داده های فضایی - زمانی پرداخته اند.

در این مقاله برای کاهش زمان محاسبات، میدان تصادفی پنهان گاوسی با تابع کوواریانس ماترن به یک میدان تصادفی مارکوفی گاوسی تبدیل می شود و با این تبدیل، در محاسبات به جای استفاده از یک ماتریس غیرتُنک از یک ماتریس تُنک<sup>۶</sup>

<sup>۳</sup> Dence matrix

<sup>۴</sup> Integrated Nested Laplace Approximation

<sup>۵</sup> Gaussian markov random field

<sup>۶</sup> Sparse matrix

استفاده می‌شود. ماتریس کُتک ایجاد شده بیشتر درایه‌هایش صفر است و درایه‌های غیرصفر آن روی قطر و حوالی آن پراکنده باشند. لیندگرن و همکاران (۲۰۱۱) برای تولید مناسب این میدان تصادفی مارکوفی گاوسی، رهیافت معادلات دیفرانسیل جزئی تصادفی<sup>۷</sup> (SPDE) به کار گرفتند. دلیل استفاده از میدان تصادفی مارکوفی گاوسی به جای میدان تصادفی گاوسی، خواص محاسباتی خوبی است که میدان تصادفی مارکوفی گاوسی برخوردار می‌باشد. در واقع میدان تصادفی مارکوفی گاوسی به وسیله یک ماتریس دقت<sup>۸</sup> کُتک، تعریف می‌شود و استفاده از روش‌های عددی محاسباتی موثر را امکان‌پذیر می‌سازد. به عبارت دیگر تابع کوواریانس فضایی-زمانی و ماتریس غیرکُتک کوواریانس مربوط به میدان تصادفی گاوسی، به ترتیب توسط یک ساختار همسایگی و یک ماتریس دقت کُتک جایگزین می‌شوند. در بخش ۲ مدل سلسله مراتبی فضایی-زمانی معرفی و تحلیل بیزی این مدل در بخش ۳ ارائه می‌شود. در بخش ۴ به معرفی GMRF پرداخته می‌شود و رهیافت SPDE و نحوه ایجاد یک GMRF با این رهیافت در بخش ۵ بیان می‌گردد. برآورد بیزی تقریبی پارامترها و پیش‌گویی‌ها با استفاده از روش INLA در بخش ۶ توضیح داده می‌شود. در بخش ۷ نحوه کاربست و پیاده‌سازی مدل و روش‌های ارائه شده بر روی داده‌های میزان بارندگی در سال ۱۳۹۱ استان سمنان اعمال می‌گردد. در نهایت پیشنهادات و نتیجه‌گیری در بخش ۸ آورده می‌شود.

## ۲ مدل سلسله مراتبی فضایی-زمانی

فرض کنید  $Y(s, t) = \{y(s, t), s \in D \subseteq R^2, t \in R\}$  یک میدان تصادفی فضایی-زمانی گاوسی باشد و  $y(s_i, t)$ ‌ها مشاهدات این میدان در موقعیت فضایی  $s_i = 1, \dots, d$  و در زمان  $t = 1, \dots, T$  باشند. این میدان تصادفی را می‌توان به صورت

$$y(s_i, t) = u(s_i, t) + \varepsilon(s_i, t) \quad (1)$$

<sup>۷</sup> Stochastic Partial Differential Equations

<sup>۸</sup> Precision matrix

تجزیه نمود، که در آن  $\varepsilon(s_i, t) \sim N(0, \sigma_\varepsilon^2)$  خطای اندازه گیری است و به صورت یک میدان نوفه سفید گاوسی<sup>۹</sup> مدل بندی می شود، به طوری که هر زوج خطا دوبه دو ناهمبسته هستند و  $\sigma_\varepsilon^2$  واریانس تصادفی یا اثر قطعه ای<sup>۱۰</sup> است. وابستگی فضایی-زمانی مشاهدات با اضافه کردن جمله  $u(s_i, t)$  به مدل در نظر گرفته می شود که یک میدان حالت است و می تواند به شکل های متفاوت در نظر گرفته شود.

در این مقاله میدان حالت  $u(s_i, t)$  به صورت  $u(s_i, t) = \mu(s_i, t) + \xi(s_i, t)$  تجزیه می شود، که در آن  $\mu(s_i, t)$  تعییرات بزرگ مقیاس<sup>۱۱</sup> یا روند<sup>۱۲</sup> گفته می شود که ممکن است ناشی از تغییر پذیری بین موقعیت های مشاهده شده باشند و تابعی از متغیرهای کمکی است و به صورت  $\mu(s_i, t) = z(s_i, t)\beta$  مشخص می شود، که در آن  $z(s_i, t) = (z_1(s_i, t), \dots, z_p(s_i, t))'$  برداری با بعد  $p$  از متغیرهای کمکی در موقعیت  $s_i$  و زمان  $t$  است و  $\beta = (\beta_1, \dots, \beta_p)'$  بردار ضرایب رگرسیونی می باشد.  $\xi(s_i, t)$  تحقق از یک میدان تصادفی فضایی-زمانی گاوسی با میانگین صفر است که با زمان تغییر می کند و مدل اتو رگرسیو<sup>۱۳</sup> (AR) مرتبه اول با ضرایب  $a$  و همبستگی فضایی-زمانی برای  $t = 1, \dots, T$  به صورت

$$\xi(s_i, t) = a\xi(s_i, t-1) + \omega(s_i, t), \quad (2)$$

در نظر گرفته می شود، که در آن  $|a| < 1$  و  $\xi(s_i, 1)$  دارای توزیع مانای  $N(0, \frac{\sigma_\omega^2}{1-a^2})$  است. همچنین فرض می شود  $\omega(s_i, t)$  فرآیندی مستقل از زمان و دارای توزیع گاوسی با میانگین صفر و تابع کوواریانس همسان گرد و مانای مرتبه دوم فضایی-زمانی صفر برای  $t = t'$  و برای هر  $i \neq j$  و  $t \neq t'$  به صورت  $Cov(\omega(s_i, t), \omega(s_j, t')) = \sigma_\omega^2 c(h)$  است، به طوری که برای هر  $s_i$  و  $t$ ،  $var(\omega(s_i, t)) = \sigma_\omega^2$  و تابع همبستگی فضایی  $c(h)$  به موقعیت های  $s_i$  و  $s_j$  تنها از طریق فاصله اقلیدسی  $h = \|s_i - s_j\| \in R$  وابسته می باشد و به صورت یک تابع

<sup>۹</sup> Gaussian white noise process

<sup>۱۰</sup> Nugget effect

<sup>۱۱</sup> Large scale variation

<sup>۱۲</sup> Trend

<sup>۱۳</sup> Auto Regressive

ماترن،

$$c(h) = \frac{1}{\Gamma(\nu) 2^{\nu-1}} (\kappa h)^\nu K_\nu(\kappa h)$$

در نظر گرفته می‌شود، که در آن  $\nu$  پارامتر هموارسازی است و درجه همواری فرآیند را اندازه‌گیری می‌کند و  $\kappa > 0$  پارامتر مقیاس نامیده می‌شود.  $K_\nu(\cdot)$  یک تابع بسل تصحیح شده نوع دو و مرتبه  $\nu$  است. پارامتر مقیاس با دامنه  $\rho$  به صورت  $\rho = \frac{\sqrt{\lambda \nu}}{k}$  در ارتباط است.  $\rho$  فاصله‌ای است که در آن همبستگی نزدیک به ۱/۰ برای هر  $\nu$  است. اگر  $\nu = 1/2$  تابع همبستگی ماترن تبدیل به تابع همبستگی نمایی و اگر  $\nu \rightarrow \infty$  تابع همبستگی ماترن به تابع همبستگی گاوسی تبدیل می‌شود. اکنون مدل (۱) را می‌توان به صورت

$$y(s_i, t) = z(s_i, t)\beta + \xi(s_i, t) + \varepsilon(s_i, t) \quad (3)$$

نوشت و در نهایت از (۲) و (۳) مدل کلی را می‌توان به شکل ماتریسی به صورت

$$y_t = z_t \beta + \xi_t + \varepsilon_t, \quad \varepsilon_t \sim N(\mathbf{0}, \sigma_\varepsilon^2 I_d) \quad (4)$$

$$\xi_t = a\xi_{t-1} + \omega_t, \quad \omega_t \sim N(\mathbf{0}, \Sigma = \sigma_\omega^2 \tilde{\Sigma}) \quad (5)$$

بیان کرد، که در آن  $y_t = (y(s_1, t), \dots, y(s_d, t))'$  بردار مشاهدات در هر زمان  $t, I_d$  ماتریس واحد از مرتبه  $d$  و

$$z_t = (z(s_1, t)', \dots, z(s_d, t)')', \quad \xi_t = (\xi(s_1, t), \dots, \xi(s_d, t))'$$

به طوری که  $\xi_1$  از یک فرآیند مانای  $AR(1)$  با توزیع  $N(\mathbf{0}, \frac{\Sigma}{1-a^2})$  استخراج می‌شود.  $\tilde{\Sigma}$  ماتریس همبستگی با بعد  $d$  و عناصر  $c(\|s_i - s_j\|)$  که در آن  $c(\cdot)$  تابع ماترن با پارامترهای  $\nu$  و  $\kappa$  است. بردار متغیرهای پنهان  $\xi = (\xi_1', \dots, \xi_T)'$  به صورت یک میدان تصادفی مارکوفی گاوسی در نظر گرفته می‌شود.

### ۳ تحلیل بیزی مدل فضایی-زمانی

در این بخش تحلیل بیزی مدل‌های فضایی-زمانی روی یک ناحیه پیوسته فضایی مورد بررسی قرار می‌گیرد. بردار متغیرهای پنهان  $\xi = (\xi_1', \dots, \xi_T)'$  و

بردار پارامترهای رگرسیونی به صورت  $\beta = (\beta_1, \dots, \beta_p)'$  در یک بردار به صورت  $x = \{\beta', \xi'\}'$  در نظر گرفته می‌شود، که میدان تصافی پنهان با مولفه‌های مستقل است. ابرپارامترها،  $\sigma_\varepsilon^2$  پارامتر واریانس تصادفی،  $a$  ضریب مدل اتو رگرسیو مرتبه اول،  $\sigma_\omega^2$  و  $\kappa$  پارامترهای تابع کوواریانس، در یک بردار به صورت  $\theta = (\sigma_\varepsilon^2, a, \sigma_\omega^2, \kappa)$  در نظر گرفته می‌شوند. پارامتر  $\nu$  ثابت و برابر یک در نظر گرفته می‌شود. اکنون با فرض استقلال پیشین‌های ابرپارامترها، توزیع پیشین توأم  $\pi(\theta) = \pi(\sigma_\varepsilon^2)\pi(a)\pi(\sigma_\omega^2)\pi(\kappa)$  می‌باشد. با توجه به این که مشاهدات  $y_t$  دنباله‌ای مستقل به شرط  $\xi_t$  هستند و  $\xi_t$  یک میدان تصادفی پنهان به صورت پویای مارکوفی زمانی است، توزیع پسین را می‌توان به شکل

$$\begin{aligned} \pi(x, \theta | y) &\propto \pi(x | \theta) \pi(\theta) \pi(y | x, \theta) \\ &= \pi(\xi | \theta) \pi(\beta) \pi(\theta) \prod_{t=1}^T \pi(y_t | \xi_t, \beta, \theta) \\ &= \pi(\xi_1 | \theta) \prod_{t=2}^T \pi(\xi_t | \xi_{t-1}, \theta) \pi(\beta) \pi(\theta) \prod_{t=1}^T \pi(y_t | \xi_t, \beta, \theta) \quad (6) \end{aligned}$$

نوشت. گاهی در عمل اطلاعاتی کافی در مورد پارامترها در دسترس نمی‌باشد و در تحلیل بیزی از توزیع‌های پیشین ناآگاهی بخش<sup>۱۴</sup> استفاده می‌شود، پیشین‌های جفریز، لاپلاس و مبهم<sup>۱۵</sup> در کلاس پیشین‌های ناآگاهی بخش قرار می‌گیرند. در این مقاله توزیع‌های پیشین سره با واریانس بزرگ که نوعی پیشین مبهم هستند، برای پارامترهای مدل به کار گرفته می‌شوند. معمولاً برای پارامترهای رگرسیونی یک توزیع پیشین مبهم نرمال  $\beta \sim N(\mathbf{0}, \sigma_\beta^2 I_p)$  با  $\sigma_\beta^2$  بزرگ در نظر می‌گیرند، که در این مقاله برای داشتن یک توزیع پیشین مبهم  $\sigma_\beta^2 = 10^3$  فرض شده است. برای واریانس‌ها پیشین مبهم گامای معکوس برای مثال  $IG(2, 1)$  (توزیعی با میانگین ۱ و واریانس نامتناهی)، برای پارامتر  $\kappa$  و  $a$  توزیع پیشین یکنواخت و میدان پنهان  $\xi$  به صورت یک میدان تصادفی مارکوفی گاوسی در نظر گرفته می‌شود. بنابراین شکل توزیع پسین بسیار پیچیده است و از روش‌های مونت کارلوی زنجیر مارکوفی برای

<sup>۱۴</sup> Non-informative Prior

<sup>۱۵</sup> Vague

شبیه‌سازی از توزیع پسین استفاده می‌شود. برای به کارگیری الگوریتم نمونه‌گیری گیبز توزیع‌های شرطی کامل پارامترها تشکیل می‌شوند.

فرض کنید  $\beta \sim N_p(\circ, \sigma_\beta^2 I_p)$ ، آن‌گاه توزیع شرطی کامل  $\beta$  به صورت

$$\begin{aligned} \pi(\beta | \mathbf{y}, \boldsymbol{\xi}, \boldsymbol{\theta}) &\propto \prod_{t=1}^T \pi(\mathbf{y}_t | \beta, \sigma_\varepsilon^2) \pi(\beta) \\ &\propto \phi_d(\mathbf{y}; \mathbf{z}_t \beta + \boldsymbol{\xi}_t, \sigma_\varepsilon^2 I_d) \phi_p(\circ, \sigma_\beta^2 I_p) \\ &\propto \exp\left\{-\frac{1}{\sigma_\beta^2} (\beta - \boldsymbol{\mu}_\beta)' \Sigma_\beta^{-1} (\beta - \boldsymbol{\mu}_\beta)\right\}, \end{aligned}$$

است. بنابراین چون توزیع شرطی کامل  $\beta$  شکل بسته نرمال دارد، تولید نمونه از توزیع شرطی کامل  $\beta$  به راحتی امکان‌پذیر است. توزیع شرطی کامل برای سایر پارامترها

$$\begin{aligned} \pi(\sigma_\varepsilon^2 | \mathbf{y}, \mathbf{x}, \boldsymbol{\theta}) &\propto \prod_{t=1}^T \pi(\mathbf{y}_t | \beta, \sigma_\varepsilon^2) \pi(\sigma_\varepsilon^2) \\ &= \prod_{t=1}^T \phi_d(\mathbf{y}_t; \mathbf{z}_t \beta + \boldsymbol{\xi}_t, \sigma_\varepsilon^2 I_d) IG(\alpha_\varepsilon, \gamma_\varepsilon), \\ \pi(\sigma_\omega^2 | \mathbf{y}, \mathbf{x}, \boldsymbol{\theta}) &\propto \pi(\sigma_\omega^2) \pi(\boldsymbol{\xi}_1 | \sigma_\varepsilon^2, a, \kappa) \\ &= \pi(\sigma_\omega^2) \pi(\boldsymbol{\xi}_1 | \sigma_\varepsilon^2, a, \kappa) \prod_{t=2}^T \pi(\boldsymbol{\xi}_t | \sigma_\varepsilon^2, a, \kappa) \\ &= IG(\alpha_\omega, \gamma_\omega) \phi_d(\boldsymbol{\xi}_1; \circ, \frac{\Sigma}{\sqrt{1-a^2}}) \prod_{t=2}^T \phi_d(\boldsymbol{\xi}_t; a \boldsymbol{\xi}_{t-1}, \Sigma) \\ \pi(a | \mathbf{y}, \mathbf{x}, \boldsymbol{\theta}) &\propto \pi(a) \pi(\boldsymbol{\xi}_1 | \sigma_\varepsilon^2, a, \kappa) \\ &= \pi(a) \pi(\boldsymbol{\xi}_1 | \sigma_\varepsilon^2, a, \kappa) \prod_{t=2}^T \pi(\boldsymbol{\xi}_t | \sigma_\varepsilon^2, a, \kappa) \\ &= \pi(a) \phi_d(\boldsymbol{\xi}_1; \circ, \frac{\Sigma}{\sqrt{1-a^2}}) \prod_{t=2}^T \phi_d(\boldsymbol{\xi}_t; a \boldsymbol{\xi}_{t-1}, \Sigma) \\ \pi(\kappa | \mathbf{y}, \mathbf{x}, \boldsymbol{\theta}) &\propto \pi(\kappa) \pi(\boldsymbol{\xi}_1 | \sigma_\varepsilon^2, a, \kappa) \\ &= \pi(\kappa) \pi(\boldsymbol{\xi}_1 | \sigma_\varepsilon^2, a, \kappa) \prod_{t=2}^T \pi(\boldsymbol{\xi}_t | \sigma_\varepsilon^2, a, \kappa) \\ &= \pi(\kappa) \phi_d(\boldsymbol{\xi}_1; \circ, \frac{\Sigma}{\sqrt{1-a^2}}) \prod_{t=2}^T \phi_d(\boldsymbol{\xi}_t; a \boldsymbol{\xi}_{t-1}, \Sigma) \end{aligned}$$



حاصل می‌شوند، که دارای شکل خاصی از توزیع‌های شناخته شده نیستند و برای تولید نمونه از الگوریتم متروپلیس-هستینگس استفاده می‌شود. پیش‌گویی بیزی در موقعیت جدید  $s_0$  از توزیع پیش‌گو

$$\pi(y(s_0, t)|\mathbf{y}) = \int \pi(y(s_0, t)|\xi(s_0, [t], \theta)\pi(\xi(s_0, [t])|\mathbf{y})d\xi(s_0, [t]) \quad (V)$$

به دست می‌آید، که در آن  $\pi(\xi(s_0, [t])|\mathbf{y}) = \int \pi(\xi(s_0, [t])|\xi, \theta)\pi(\xi, \theta|\mathbf{y})d\xi d\theta$  منظور از  $\xi(s_0, [t]) = (\xi(s_0, 1), \dots, \xi(s_0, t))'$  است.

به دلیل پیچیدگی توزیع پسین و عدم وجود شکل بسته و وجود تعداد زیادی متغیر پنهان و پارامتر معمولاً استفاده از رهیافت بیزی معمولی و نمونه‌های مونت کارلویی بسیار وقت‌گیر و گاهی اوقات همگرایی‌ها به راحتی امکان‌پذیر نیست. استفاده از روش بیزی تقریبی برای مدل‌های فضایی-زمانی، به نمونه‌گیری‌های وقت‌گیر مونت کارلویی نیاز نداشته و به‌طور قابل ملاحظه‌ای سرعت محاسبات را افزایش می‌دهد. در این رهیافت از تقریب INLA که توسط رو و همکاران (۲۰۰۹) معرفی شده است استفاده می‌شود، که محاسباتی سریع و تقریبی دقیق از توزیع‌های پسین را جایگزین شبیه‌سازی‌های سنگین الگوریتم‌های مونت کارلویی می‌کند. همچنین در راستای راحتی و کوتاه کردن محاسبات، بردار متغیرهای پنهان  $\xi$  در این مدل‌ها به صورت یک میدان تصادفی مارکوفی گاوسی در نظر گرفته می‌شود. در قسمت بعد به تعریف میدان تصادفی مارکوفی گاوسی و نحوه تولید این میدان با رهیافت SPDE پرداخته می‌شود.

#### ۴ میدان تصادفی مارکوفی گاوسی

قبل از تعریف میدان تصادفی مارکوفی گاوسی نیاز به تعریف گراف است. مجموعه‌ای از رأس‌ها که توسط تعدادی یال به هم وصل شده‌اند گراف نامیده می‌شود. به عبارت دیگر یک گراف به صورت زوج  $\mathfrak{G} = (\nu, \varepsilon)$  تعریف می‌شود، که در آن  $\nu$  یک مجموعه متناهی ناتهی از عناصری به نام رئوس است و  $\varepsilon$  مجموعه‌ای متناهی از زوج‌های نامرتب و نامساوی از عناصر متمایز  $\nu$  به صورت  $\varepsilon \subseteq \{\{i, j\} : i, j \in \nu, i \neq j\}$  موسوم به یال‌ها است، به طوری که رأس‌ها را به هم

وصل می‌کنند. از آنجا که یال‌های گراف به صورت زوج‌های نامرتب و نامساوی تعریف شده است، این بدان معناست که میان هر دو رأس حداکثر یک یال وجود خواهد داشت. در واقع یک گراف به صورت مجموعه‌ای از نقاط و نحوه‌ی ارتباط بین آن‌ها در نظر گرفته می‌شود. در صورتی که رأس‌های یک گراف با مجموعه‌ای از نمادها مانند  $\nu = \{1, \dots, n\}$  مشخص شود، گراف نشان‌دار نامیده می‌شود.

فرض کنید بردار تصادفی  $x = (x_1, \dots, x_n)'$  دارای توزیع نرمال چندمتغیره با بردار میانگین  $\mu$  و ماتریس کوواریانس معین مثبت  $Q^{-1}$  باشد، یعنی

$$\pi(x) = (\pi)^{-\frac{n}{2}} |Q|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x - \mu)'Q(x - \mu)\right),$$

و  $\mathfrak{G} = (\nu, \varepsilon)$  یک گراف باشد، که در آن  $\nu$  مجموعه‌ای متناهی و غیرتهی از رئوس است به طوری که  $\nu = \{1, \dots, n\}$  یعنی  $\mathfrak{G}$  یک گراف نشان‌دار است و  $\varepsilon$  شامل همه زوج‌های  $\{i, j\}$  است که رأس‌های  $i$  و  $j$  هیچ یال مشترکی نداشته باشند، اگر و تنها اگر  $x_i \perp x_j | x_{-ij}$ ، یعنی دو مولفه  $x_i$  و  $x_j$  به شرط بردار  $x$  به جز مولفه  $i$  و  $j$  آن مستقل باشند، در این صورت  $x$  یک میدان تصادفی مارکوفی گاوسی (GMRF) نسبت به گراف  $\mathfrak{G}$  نامیده می‌شود، (رو و هلد، ۲۰۰۵).

در نظر گرفتن  $x$  به صورت یک GMRF منجر به این می‌شود که توزیع شرطی کامل هر مولفه از  $x$  به صورت توزیع شرطی آن مولفه به شرط سایر مولفه‌ها شود، یعنی  $\pi(x_i | x_{-i})$  که در آن بردار  $x_{-i}$  به جز مولفه  $x_i$  است و از خاصیت مارکوفی توزیع شرطی کامل  $x_i$ ،  $i = 1, \dots, n$ ، تنها به تعداد کمی از مولفه‌های  $x$  بستگی دارد. اگر این مجموعه از مولفه‌ها را با  $\delta_i$  مشخص شود که تشکیل دهنده مجموعه‌ای از همسایگی‌های مولفه  $i$  است، آن‌گاه  $\pi(x_i | x_{\delta_i}) = \pi(x_i | x_{-i})$  به طوری که فرض می‌شود جملات  $x_i$  و  $x_{-\{i, \delta_i\}}$  مستقل هستند. برای هر  $i = 1, \dots, n$  این رابطه استقلال شرطی می‌تواند به صورت  $x_i \perp x_{-\{i, \delta_i\}} | x_{\delta_i}$  نوشته شود. رو و هلد (۲۰۰۵) نشان دادند که درایه  $Q$  ماتریس دقت معین مثبت<sup>۱۶</sup>  $Q$  صفر است اگر و تنها اگر  $x_i \perp x_j | x_{-\{i, j\}}$  و درایه‌های غیرصفر ماتریس  $Q$  توسط ساختار همسایگی میدان مشخص می‌شوند،  $Q_{ij} \neq 0$ ،  $j \in \{i, \delta_i\}$  و بنابراین ماتریس دقت  $Q$  به صورت

<sup>۱۶</sup> Positive definite precision matrix

یک ماتریس کوانتیک می‌باشد و در راستای راحتی و کوتاه کردن محاسبات بسیار مفید می‌باشد. در واقع میدان تصادفی مارکوفی گاوسی به وسیله‌ی یک ماتریس دقت که ساختار کوانتیک دارد، تعریف می‌شود و استفاده از روش‌های عددی محاسباتی موثر را امکان‌پذیر می‌سازد و دلیل اصلی برای استفاده از یک GMRF در مدل‌های پیچیده سلسله مراتبی فضایی این است که تابع کوواریانس فضایی- زمانی و ماتریس کوواریانس غیرکوانتیک مربوط به میدان تصادفی گاوسی، به ترتیب توسط یک ساختار همسایگی و یک ماتریس دقت پراکنده جایگزین می‌شوند.

## ۵ رهیافت SPDE برای پیدا کردن یک GMRF

فرض کنید  $X(s) = \{x(s), s \in D \subseteq R^2\}$  یک میدان تصادفی گاوسی مانای مرتبه دوم و همسان‌گرد با تابع کوواریانس ماترن با پارامتر همواری  $\nu$  و پارامتر مقیاس  $\kappa$  باشد. همچنین فرض کنید تحقیقی از فرایند  $X(s_i)$  در  $d$  موقعیت  $s_1, \dots, s_d$  مشاهده شده باشد. روش SPDE یک GMRF با همسایگی فضایی و ماتریس دقت کوانتیک  $Q$  پیدا می‌کند که به بهترین وجه نشان‌دهنده میدان ماترن باشد. در زیربخش ۱.۵ نحوه تبدیل یک میدان ماترن به یک میدان تصادفی مارکوفی گاوسی بیان می‌شود. در این روش میدان تصادفی ماترن به صورت یک ترکیب خطی از توابع تکه‌ای خطی که بر یک مثلث‌بندی از دامنه  $D$  تعریف شده است، بیان می‌شود، دامنه  $D$  به مجموعه‌ای از مثلث‌ها تقسیم‌بندی می‌شود که در بیشتر از یک یال مشترک یا زاویه مشترک متقاطع نیستند. رئوس مثلث‌بندی اولیه در موقعیت‌های  $s_1, \dots, s_d$  قرار می‌گیرد و سپس رئوس دیگر به منظور رسیدن به یک مثلث‌بندی مناسب برای پیش‌بینی فضایی اهداف اضافه می‌شوند.

### ۱.۵ معادلات دیفرانسیل جزئی

ویتل (۱۹۶۳ و ۱۹۵۴) به مطالعه معادلات دیفرانسیل جزئی پرداخت و ثابت کرد که میدان‌های تصادفی ماترن تنها جواب‌های پایدار برای یک معادله دیفرانسیل جزئی هستند، به طوری که یک میدان تصادفی گاوسی  $x(s)$  با تابع کوواریانس ماترن و

پارامتر هموارسازی  $\nu$  و پارامتر مقیاس  $\kappa$  به صورت

$$C(h) = \left(\frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}\right)(\kappa h)^\nu K_\nu(\kappa h) \quad (8)$$

است، که در آن  $h = \|s_i - s_j\|$  و  $\sigma^2 = \frac{\Gamma(\nu)}{\Gamma(\nu+\frac{d}{2})(\frac{\pi}{2})^{\frac{d}{2}}\kappa^{2\nu}}$  می‌باشند. می‌فد اصاصتی نژ ج ل یسنار فید ه لداعه می ارب<sup>۱۷</sup> رادیا پابل ح ک ی

$$(\kappa^2 - \nabla^2)^{\frac{\alpha}{2}} x(s) = W(s), \quad s \in \mathbf{R}^d, \quad \alpha = \nu + \frac{d}{2}, \quad \kappa > 0, \quad \nu > 0 \quad (9)$$

است، که در آن  $W$  میدان تصادفی فضایی نوفه سفید و  $\nabla^2$  عمل گر دیفرانسیلی مرتبه دوم یا همان عمل گر لاپلاس است.  $(\kappa^2 - \nabla^2)^{\frac{\alpha}{2}}$  یک عمل گر شبه-دیفرانسیل<sup>۱۸</sup> می‌باشد، به طوری که با به کار بردن تعریف تبدیل فوریه یک عملگر لاپلاس کسری<sup>۱۹</sup> به صورت

$$\{F(\kappa^2 - \nabla^2)^{\frac{\alpha}{2}}\phi\}(\mathbf{K}) = (\kappa^2 + \|\mathbf{K}\|^2)^{\frac{\alpha}{2}}(F\phi)(\mathbf{K})$$

در نظر گرفته می‌شود، که در آن  $\phi$  تابعی در  $\mathbf{R}^d$  است.

لیندگرن و همکاران (۲۰۱۱) برای ساختن یک میدان تصادفی مارکوفی گاوسی که بیان گر یک میدان تصادفی ماترن روی یک شبکه مثلث بندی باشد، در مرحله اول یک فرم ضعیف تصادفی<sup>۲۰</sup> معادله دیفرانسیل جزئی (۹) را در نظر گرفتند. با تعریف ضرب داخلی به صورت  $\langle f, g \rangle = \int f(\mathbf{u})g(\mathbf{u})d\mathbf{u}$  برای هر مجموعه متناهی مناسب از توابع تست<sup>۲۱</sup>  $\{\phi_j(s), j = 1, \dots, m\}$ ، با داشتن شرط

$$\{\langle \phi_j, (\kappa^2 - \nabla^2)^{\frac{\alpha}{2}} x \rangle, j = 1, \dots, n\} \stackrel{d}{=} \{\langle \phi_j, W \rangle, j = 1, \dots, n\}$$

حل ضعیف تصادفی به دست آوردند، که در آن  $d$  نشان دهنده هم توزیعی است. در مرحله بعد برای نمایش المان محدود<sup>۲۲</sup> حل معادله دیفرانسیل جزئی، یک ترکیب

<sup>۱۷</sup> Stationary solution

<sup>۱۸</sup> Pseudo-differential Operator

<sup>۱۹</sup> Fractional Laplacian

<sup>۲۰</sup> Stochastic weak formulation

<sup>۲۱</sup> Test functions

<sup>۲۲</sup> Finite Element Method

خطی از توابع پایه ای  $\psi_l(s)$  به شکل

$$x(s) = \sum_{l=1}^n \psi_l(s) \omega_l \quad (10)$$

ایجاد نمودند. روش المان محدود روشی عددی برای حل تقریبی معادلات دیفرانسیل جزئی است و مثلث بندی یک راه حل معمول برای حل معادلات دیفرانسیل جزئی با استفاده از روش المان محدود می باشد، (برنر و اسکات، ۲۰۰۷)، به طوری که دقت جواب ها بستگی به نحوه مثلث بندی دارد. در  $(10)$ ،  $n$  تعداد کل رئوس،  $\omega_l$  وزن هایی هستند که دارای توزیع گاوسی هستند و توابع  $\psi_l(s)$  خطی تکه ای روی هر مثلث می باشند، به طوری که  $\psi_l(s)$  در رأس  $l$  برابر یک و در سایر رئوس صفر است.

معمولاً در مطالعات فضایی یک ناحیه فضایی مثل  $D \subseteq \mathbf{R}^d$  یعنی  $d = 2$  مورد نظر است. دامنه  $D$  به مجموعه ای از مثلث ها تقسیم بندی می شود که در بیشتر از یک یال مشترک یا زاویه مشترک متقاطع نیستند. در عمل رئوس مثلث بندی اولیه در موقعیت های مشاهده شده قرار می گیرد و سپس رئوس دیگر به منظور رسیدن به یک مثلث بندی مناسب برای پیش گویی فضایی اضافه می شوند. یک مثلث بندی مناسب باید بتواند ناحیه مورد مطالعه را به بهترین وجه مدل کند و نتایج باید با آسانی قابل تکرار و قابل پیش بینی باشد. مثلث بندی دلونی<sup>۲۳</sup> نوعی مثلث بندی است که در آن دایره محیطی هر مثلث شامل هیچ رأسی از رئوس مثلث های دیگر نباشد (شرط دایره محیطی) و واضح است که این نوع مثلث بندی یکتا خواهد بود. تاکنون الگوریتم های مختلفی جهت بهبود زمانی ایجاد مثلث بندی و همچنین مثلث بندی دلونی مطرح شده است. این الگوریتم ها دارای پیچیدگی زمانی متفاوتی هستند. در این مقاله برای مثلث بندی از بسته INLA استفاده شده است، که در این بسته مثلث بندی دلونی و راه حل هیلری و دایلن (۲۰۰۶) به کار گرفته شده است. ثابت می شود که مثلث بندی دلونی کوچک ترین زاویه را ماکزیمم می کند، یک مثلث بندی مجاز و همچنین زاویه - بهینه است. همچنین تعداد مورد انتظار مثلث های ایجاد شده توسط الگوریتم دلونی، حداکثر  $n + 1$  است و زمان هزینه شده توسط این

<sup>۲۳</sup> Delaunay triangulation

الگوریتم با تعداد مثلث‌های ایجاد شده متناسب است، یعنی  $O(n)$ . برای مطالعه پیش‌تر به هیلی و دایلن (۲۰۰۶) دی برگ و همکاران (۲۰۰۸) و لیندگرن و همکاران (۲۰۱۱) مراجعه شود.

## ۲.۵ ماتریس دقت میدان تصادفی مارکوفی گاوسی تولید شده

در بسته INLA برای  $d = 2$  معادله دیفرانسیل جزئی (۹) به صورت عمومی

$$(\kappa^2(s) - \nabla^2)^{\frac{\alpha}{2}}(\tau(s)x(s)) = W(s), \quad s \in \mathbf{R}^2, \quad \alpha = \nu + 1, \quad \nu > 0$$

در نظر گرفته می‌شود که میدان‌های نایستا را شامل شود، یعنی واریانس  $W(s)$  و پارامتر  $\kappa$  ثابت نباشند. برای سادگی واریانس  $W(s)$  را ثابت فرض می‌کنند و یک پارامتر مقیاس دیگر  $\tau(s)$  برای  $x(s)$  در نظر می‌گیرند، به طوری که اگر یکی یا هر دو پارامتر مقیاس  $\tau$  و  $\kappa$  ثابت نباشند، میدان نایستا خواهد بود و در نهایت از رابطه (۸)  $\sigma_w^2 = \frac{1}{4\pi\kappa^2\tau}$  می‌شود.

لیندگرن و همکاران (۲۰۱۱) نشان دادند وزن‌های گاوسی  $\{\omega_i\}$  در (۱۰) یک ساختار مارکوفی دارند و این وزن‌ها میدان تصادفی گاوسی ماترن را به یک میدان تصادفی مارکوفی گاوسی تبدیل می‌کنند. همچنین لیندگرن و همکاران (۲۰۱۱) سه ماتریس  $n \times n$  به صورت

$$C_{ij} = \langle \psi_i, \psi_j \rangle, \quad G_{ij} = \langle \nabla \psi_i, \nabla \psi_j \rangle, \quad (K_{\kappa^2})_{ij} = \kappa^2 C_{ij} + G_{ij}$$

در نظر گرفتند و نشان دادند که ماتریس دقت  $Q$  میدان تصادفی مارکوفی گاوسی  $\omega \sim N(0, Q^{-1})$  از تابعی از  $\kappa^2$  برای هر  $\alpha = \nu + 1$  و  $\nu = 0, 1, \dots$  می‌باشد. این تبدیل یک نگاهت صریح و روشن از پارامترهای تابع کوواریانس میدان تصادفی گاوسی  $(\kappa, \nu)$  به مولفه‌های ماتریس دقت  $Q$  میدان تصادفی مارکوفی گاوسی  $\omega$  با یک هزینه محاسباتی از مرتبه  $O(n)$  برای هر مثلث‌بندی را ایجاد می‌کند. لیندگرن و همکاران (۲۰۱۱) برای  $\alpha = 2$  و  $\phi_{\kappa} = \psi_{\kappa}$  نشان داد ماتریس دقت میدان تصادفی مارکوفی گاوسی  $\omega_i$  تابعی از  $\kappa^2$  به صورت  $K_{\kappa^2} C^{-1} K_{\kappa^2}$  می‌شود. ماتریس‌های  $C$

و  $G$  به راحتی قابل محاسبه هستند و ماتریس  $C^{-1}$  یک ماتریس  $n \times n$  است ( لیندگرن و همکاران، ۲۰۱۱).

### ۳.۵ بازنویسی مدل

اکنون در هر زمان  $t = 1, \dots, T$  میدان ماترن  $\omega_t$  که در رابطه (۵) معرفی شده از طریق رهیافت SPDE به یک GMRF تبدیل و به صورت  $\omega_t \sim N(\mathbf{0}, Q_s^{-1})$  نشان داده می شود، که در آن ماتریس دقت  $Q_s$  با توجه به فرض استقلال

$$Cov(\omega(s_i, t), \omega(s_j, t')) = \begin{cases} \circ & t = t' \\ \sigma_w^2 c(h) & t \neq t', i \neq j \end{cases}$$

نسبت به زمان تغییر نمی کند و بعد آن برابر  $n$  یعنی تعداد رئوس مثلث بندی از روش SPDE می باشد. بنابراین رابطه (۵) برای هر  $t = 1, \dots, T$  به صورت

$$\xi_t = a\xi_{t-1} + \omega_t \quad \omega_t \sim N(\mathbf{0}, Q_s^{-1}), \xi_1 \sim N(\mathbf{0}, \frac{Q_s^{-1}}{(1-a^2)}) \quad (11)$$

می شود و میدان تصادفی مارکوفی گاوسی  $\xi = (\xi_1', \dots, \xi_T')$  دارای توزیع توام  $Tn$  متغیره به شکل  $\xi \sim N(\mathbf{0}, Q^{-1})$  است، که در آن  $Q$  از حاصلضرب کرونگر  $Q = Q_T \otimes Q_s$  به دست می آیند. در بخش ۲.۵ تعریف شد و  $Q_T$  ماتریس دقت فرآیند اتو رگرسیو مرتبه اول رابطه (۱۱) و به صورت

$$Q_T = \begin{pmatrix} \frac{1}{\sigma_w^2} & -a/\sigma_w^2 & & & \\ -a/\sigma_w^2 & (1+a^2)/\sigma_w^2 & & & \\ & & \dots & & \\ & & & (1+a^2)/\sigma_w^2 & -a/\sigma_w^2 \\ & & & -a/\sigma_w^2 & \frac{1}{\sigma_w^2} \end{pmatrix}$$

است. با توجه به میدان تصادفی مارکوفی گاوسی تعریف شده، اکنون رابطه (۴) به صورت

$$y_t = z_t \beta + B \xi_t + \varepsilon_t \quad \varepsilon_t \sim N(\mathbf{0}, \sigma_\varepsilon^2 I_d) \quad (12)$$

بازنویسی می شود، که در آن  $B$  ماتریسی  $n \times d$  با بعد  $d \times n$  می باشد که اکثر عناصر آن صفر است و در هر سطر فقط یک عنصر یک است و توسط این ماتریس

مقادیری از میدان تصادفی مارکوفی گاوسی  $\xi_t$  برای هر مشاهده‌ی بردار  $y_t$  انتخاب می‌شود. رابطه (۱۲) برای هر عضو  $y_t$  به صورت

$$y(s_i, t) = z(s_i, t)\beta + \sum_{j=1}^n B_{ij}\xi_t + \varepsilon(s_i, t)$$

است که اگر رأس مثلث  $\lambda$  در موقعیت  $s_i$  قرار داشته باشد  $B_{ij} = 1$  و در غیر این صورت برابر صفر است.

## ۶ تحلیل بیزی تقریبی مدل فضایی زمانی

مشابه بخش ۳ بردار  $\xi$  و  $\beta$  به صورت یک میدان تصادفی پنهان  $x = (\xi, \beta)$  و  $\theta = (\sigma_\varepsilon^2, a, \sigma_w^2, \kappa)$  در نظر گرفته می‌شود. با فرض یک توزیع پیشین مبهم گاوسی با دقت معلوم  $\sigma_\beta^2$  برای بردار  $\beta$  و توزیع GMRF مشخص شده در رابطه (۱۱) برای  $\xi$ ، چگالی  $\pi(x | \theta)$  توزیع گاوسی با میانگین صفر و ماتریس دقت  $Q(\theta_1)$  با بردار  $\theta_1 = (\sigma_w^2, a, \kappa)$  دارد. همچنین مشاهدات  $y = \{y_t\}$  به طور نرمال توزیع شده‌اند و به شرط  $\theta_2 = (\sigma_\varepsilon^2)$  مستقل شرطی هستند. توزیع پسین توأم به صورت

$$\pi(x, \theta | y) \propto \pi(\theta) \pi(x | \theta) \prod_{t=1}^T \pi(y_t | x, \theta)$$

می‌شود و از رابطه (۱۲) توزیع شرطی مشاهدات در زمان  $t$  به صورت

$$(y_t | x, \theta) \sim N(z_t \beta + B \xi_t, \sigma_\varepsilon^2 I_d)$$

است. توزیع‌های پسین حاشیه‌ای برای متغیرهای پنهان و ابرپارامترها به صورت

$$\pi(x_i | y) = \int \pi(x_i | \theta, y) \pi(\theta | y) d\theta, \quad i = 1, 2, \dots, T + P \quad (13)$$

$$\pi(\theta_j | y) = \int \pi(\theta | y) d\theta_{-j}, \quad j = 1, 2, 3, 4 \quad (14)$$

به دست می‌آیند، بنابراین مسأله اصلی برای برآورد پارامترها و پیش‌گویی‌ها، محاسبه دو توزیع پسین  $\pi(x | y, \theta)$  و  $\pi(\theta | y)$  است. رو و همکاران (۲۰۰۹) توزیع  $\pi(x | y, \theta)$  را با به کار بردن بسط تیلور و خطی‌سازی قسمت درست‌نمایی حول یک



مقدار ثابت با توزیع نرمال تقریب زدند و پس از مشخص شدن توزیع تقریبی نرمال به صورت

$$\hat{\pi}(x|y, \theta) \approx N_n(\hat{\mu}_{x|y, \theta}(x^\circ), \hat{\Sigma}_{x|y, \theta}(x^\circ)),$$

طبق خاصیت بسته بودن توزیع نرمال نسبت به حاشیه‌سازی، می‌توان توزیع تقریبی نرمال  $\pi(x_i|y, \theta)$  را به دست آورد. یک تقریب لاپلاس اصلاح شده نیز برای توزیع‌های حاشیه‌ای پسین متعیرهای پنهان توسط رو و همکاران (۲۰۰۹) به صورت

$$\tilde{\pi}_{LA}(x_i | y, \theta) \propto \frac{\pi(y | x) \pi(x | \theta)}{\tilde{\pi}(x_{-i} | x_i, y, \theta)} \Big|_{x=\hat{E}(x_{-i} | x_i, y, \theta)},$$

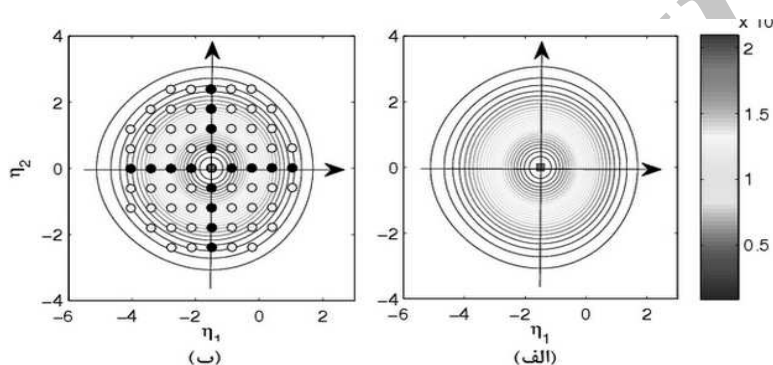
تعریف شده است، بنا به خاصیت حاشیه‌سازی توزیع نرمال  $\hat{\pi}(x_{-i} | x_i, y, \theta)$  یک توزیع تقریبی نرمال دارد. پس از تعیین این توزیع پسین تقریبی می‌توان تقریب لاپلاس توزیع پسین پارامترها را از رابطه

$$\tilde{\pi}_{LA}(\theta|y) \propto \frac{\pi(y|x)\pi(x|\theta)\pi(\theta)}{\tilde{\pi}(x|y, \theta)} \Big|_{x=\hat{E}(x|y, \theta)}, \quad (15)$$

تعیین نمود. برای تعیین تقریب (۱۵) مد تابع  $\log \tilde{\pi}(\theta|y)$  با روش‌های بهینه‌سازی به صورت  $\theta^*$  تعیین می‌شود. ماتریس هسیان  $H = \frac{\partial^2 \log \tilde{\pi}(\theta|y)}{\partial \theta^i \partial \theta^j}$  محاسبه و عکس آن به صورت  $H^{-1} = V \Lambda V'$  تجزیه می‌شود، که در آن  $V$  ماتریس بردارهای ویژه و  $\Lambda$  ماتریس قطری مقادیر ویژه آن است. مبدأ مختصات را به مد  $\theta^*$  انتقال داده و  $\theta$  توسط متغیر استاندارد  $t$  به صورت  $\theta(t) = \theta^* + V \Lambda^{-1/2} t$  تعریف می‌شود که فرمول مختصات جدید در مبدأ  $\theta^*$  است. با شروع از مبدأ مختصات جدید روی هر یک از محورهای نقطه‌ای به فاصله مقادیر صحیح  $\delta_t$  به گونه‌ای اختیار می‌شوند که شرط  $\log \tilde{\pi}(\theta(\circ)|y) - \log \tilde{\pi}(\theta(t)|y) < \delta_\pi$  برقرار باشد. سپس به طور مشابه نقاط درون صفحات نیز تعیین می‌شوند. با به کار بردن این الگوریتم نقطه‌ای از توزیع  $\tilde{\pi}(\theta|y)$  تولید می‌شوند و می‌توان تقریبی از توزیع به دست آورد. معمولاً  $\delta_t = 2/5$  فرض می‌شود و در بیشتر مسائل با این فرض تعداد نقاط مناسب از توزیع پسین تقریبی تولید می‌گردد (رو و همکاران، ۲۰۰۹).

به عنوان مثال فرض کنید  $x = (x_1, x_2)'$  دارای توزیع نرمال با پارامتر مکان  $\theta = (\theta_1, \theta_2)'$  و ماتریس کوواریانس همانی باشد. همچنین توزیع پیشین

۱  $\pi(\theta) \propto$  و مشاهدات  $y = (2, 5)'$  را برای متغیر پاسخ دو جمله‌ای با تعداد آزمایش‌ها  $u = (10, 10)'$  در نظر بگیرید. با در نظر گرفتن  $\delta_t = 0/5$  نقاط روی محورهای  $t_1$  و  $t_2$  به گونه‌ای تعیین شوند تا شرط به ازای  $\delta_\pi = 2/5$  برقرار باشد. شکل ۱ (راست) خطوط تراز  $\hat{\pi}(\theta|y)$  را نشان می‌دهد، که در آن مبدأ مختصات به مد  $\theta^* = (\theta_1^*, \theta_2^*)$  انتقال یافته است. بردارهای ویژه براساس ماتریس همبستگی در مد محاسبه شده‌اند. شکل ۱ نمودار خطوط تراز و ۶۸ نقطه تولید شده می‌باشد. اکنون با



شکل ۱: اکتشاف پسین پارامتر، الف-مد معلوم شده و مبدأ مختصات به مد انتقال داده شده است. ب- روی محورهای مختصات و درون صفحات نقاط تولید شده‌اند.

داشتن نمونه‌های تولید شده برای توزیع پسین تقریبی  $\tilde{\pi}_{LA}(\theta|y)$  و مشخص شدن توزیع پسین تقریبی  $\tilde{\pi}_{LA}(x_i|y, \theta)$ ، تقریبی از دو انتگرال (۱۳) و (۱۴) به صورت

$$\begin{aligned} \tilde{\pi}(x_i|y) &= \int \tilde{\pi}_{LA}(x_i|y, \theta_l) \tilde{\pi}_{LA}(\theta|y) \\ &= \sum_l \tilde{\pi}_{LA}(x_i|y, \theta_l) \tilde{\pi}_{LA}(\theta_l|y), \quad i = 1, \dots, T + p \end{aligned} \quad (16)$$

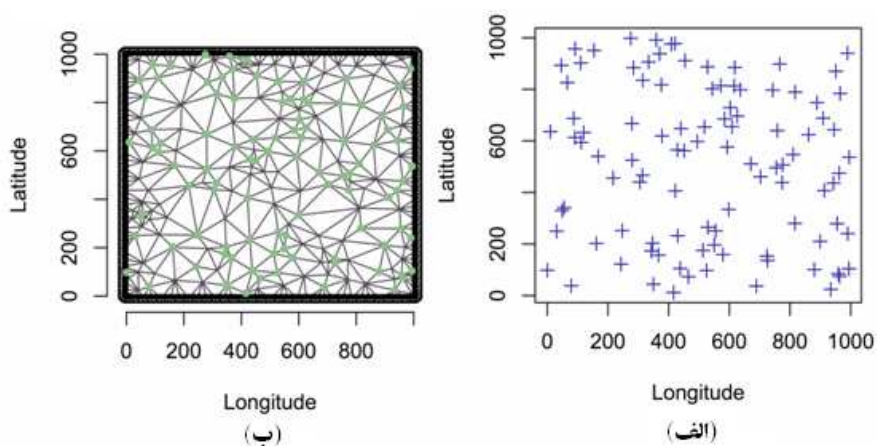
$$\tilde{\pi}(\theta_j|y) = \int \tilde{\pi}(\theta|y) d\theta_{-j}, \quad j = 1, 2, 3, 4 \quad (17)$$

حاصل می‌شود، که در آن  $l$  تعداد نقاط تولید شده از الگوریتم تعیین توزیع پسین لاپلاس تقریبی پارامترها است. پس از برآورد متغیرهای پنهان و ضرایب رگرسیون از (۱۶) و برآورد پارامترها از (۱۷) و استفاده از توزیع‌های پسین تقریبی و به کار

بردن رابطه (۷)، پیش‌گویی در موقعیت‌های معلوم و زمان معلوم مورد علاقه فایده مشاهده پاسخ، را به دست آورد.

## ۱.۶ مطالعه شبیه‌سازی

با کار و ساهو (۲۰۱۵) بسته جدید R-spTimer را برای تحلیل بیزی داده‌های فضایی-زمانی معرفی کردند و نشان دادند زمان محاسبات R-spTimer از بسته‌های موجود در نرم‌افزار R مثل spBayes برای بررسی این نوع از داده‌ها سریع‌تر می‌باشد. در این بخش یک مطالعه شبیه‌سازی برای بررسی اعتبار روش ارائه شده و مقایسه بسته R-INLA با R-spTimer ارائه می‌شود. برای تولید داده‌های فضایی-زمانی  $y(s_i, t)$  در  $100$  موقعیت،  $S = \{1, \dots, 100\}$  و  $12$  زمان،  $T = 1, \dots, 12$  در یک شبکه نامنظم، موقعیت‌ها به صورت شکل ۲ (الف) در نظر گرفته شده است. تابع کوواریانس متغیرهای پنهان فضایی  $w$ ، ماترن با مقدار معلوم  $\nu = \frac{1}{4}$ ، مقادیر پارامترها  $\sigma_w^2 = 0/3$  و  $\kappa = 0/05$ ، واریانس خطا  $\sigma_\epsilon^2 = 0/01$  و ضریب فرآیند اتو رگرسیو مرتبه اول،  $a = 0/1$  در نظر گرفته شده است. سه متغیر تبیینی از توزیع نرمال استاندارد تولید و بردار پارامترهای رگرسیونی به صورت  $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)' = (5, 2, 1, 0/5)'$  فرض گردید. با استفاده از رهیافت SPDE مثلث‌بندی بر روی ناحیه ایجاد و در شکل ۲ (ب) مشخص شده است. از رهیافت بیزی با الگوریتم مونت کارلوی زنجیر مارکوفی با  $15000$  تکرار و رهیافت بیزی تقریبی، برآورد پارامترها محاسبه گردید. نتایج تحلیل با رهیافت بیزی و بیز تقریبی در جدول ۱ ارائه شده است. الگوریتم مونت کارلوی زنجیر مارکوفی با تعداد تکرار متفاوت  $5000$ ،  $10000$  و  $20000$  انجام گردید و نتایج برای تکرار کمتر از  $15000$  از دقت کافی برخوردار نبود و با تکرار بیش از  $15000$  تغییر محسوسی در نتایج ایجاد نمی‌گردید. زمان انجام محاسبات با رهیافت بیزی معمولی با یک کامپیوتر با مشخصات  $IntelCorei7$  ( $4GB$  و  $2/7GHz$ )، حدود  $480$  ثانیه زمان و با رهیافت بیزی تقریبی حدود  $140$  ثانیه به طول انجامید.



شکل ۲: الف- نمایشی از ۱۰۰ موقعیت‌های در نظر گرفته شده بر روی یک شبکه نامنظم برای تولید داده‌ها ب- مثلث‌بندی شبکه با رهیافت SPDE.

به منظور مقایسه دقت دو رهیافت، با شرایط یکسان از فرآیند شبیه‌سازی توضیح داده شده، تعداد ۱۰۰ مجموعه داده دیگر تولید گردید و مقادیر واقعی داده‌ها و مقادیر برازش شده دو مدل با معیار  $MSE$  مورد مقایسه قرار گرفتند، این مقادیر برای روش بیزی معمولی ۰/۲۴۳۷ و برای روش بیزی تقریبی ۰/۲۳۲۱ به دست آمد. همان‌طور که ملاحظه می‌شود نتایج هر دو روش تقریباً معادل است و حتی در بعضی موارد روش بیزی تقریبی بهتر عمل کرده است.

## ۲.۶ تحلیل داده‌های بارندگی استان سمنان

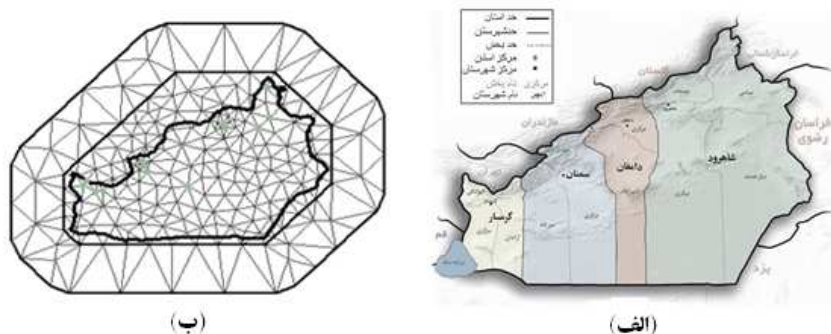
در این بخش چگونگی استفاده از روش‌های توضیح داده شده به داده‌های میزان بارندگی در استان سمنان برازش داده می‌شود. داده‌های مورد بررسی میزان بارندگی طی ۱۲ ماه سال ۱۳۹۱ اندازه‌گیری شده توسط ۱۴ ایستگاه هواشناسی در استان سمنان است.  $y(s_i, t)$  میزان بارندگی اندازه‌گیری شده در ایستگاه هواشناسی  $s_i$  در ماه  $t = 1, \dots, 12$  و  $d = 1, \dots, 14$  می‌باشد. به دلیل عدم دسترسی به

جدول ۱: مقادیر واقعی، برآورد، انحراف معیار و چندک‌های پارامترها با رهیافت

چندک‌ها			انحراف معیار	میانگین	مقدار واقعی	پارامتر	رهیافت	
۰/۹۷۵	۰/۵	۰/۰۲۵						
۵/۱۴۴۰	۵/۰۴۱۲	۴/۹۳۸۳	۰/۰۵۲۶	۵/۰۴۰۸	۵	$\beta_0$	بیزی	
۱/۹۷۰۹	۱/۹۳۱۱	۱/۸۹۱۰	۰/۰۲۰۳	۱/۹۳۱۲	۲	$\beta_1$		
۱/۰۲۶۶	۰/۹۸۶۰	۰/۹۴۶۲	۰/۰۲۰۷	۰/۹۸۶۱	۱	$\beta_2$		
۰/۵۳۳۲	۰/۴۹۵۶	۰/۴۵۷۲	۰/۰۱۹۳	۰/۴۹۵۴	۰/۵	$\beta_3$		
۰/۲۳۵۴	۰/۱۷۲۸	۰/۱۲۲۵	۰/۰۲۸۶	۰/۱۷۴۷	۰/۰۱	$\sigma_{\epsilon}^2$		
۰/۵۲۵۱	۰/۴۸۳۷	۰/۴۴۶۹	۰/۰۲۰۱	۰/۴۸۴۳	۰/۳	$\sigma_{\omega}^2$		
۰/۱۰۴۳	۰/۰۸۶۹۲	۰/۰۶۹۴	۰/۰۰۸۸	۰/۰۸۶۸	۰/۱	$a$		
۰/۱۰۷۲	۰/۰۶۱۹	۰/۰۴۵۳	۰/۰۰۵۹۲	۰/۰۷۰۰	۰/۰۵	$\kappa$		
۵/۴۴۹۲	۵/۴۰۱۳	۵/۳۶۳۱	۰/۰۱۹۴	۵/۴۰۱۳	۵	$\beta_0$		بیزی تقریبی
۲/۰۲۴۴	۱/۹۸۹۷	۱/۹۵۵۰	۰/۰۱۷۷	۱/۹۸۹۷	۲	$\beta_1$		
۱/۰۲۹۰	۰/۹۹۴۸	۰/۹۶۰۶	۰/۰۱۷۴	۰/۹۹۴۸	۱	$\beta_2$		
۰/۵۴۸۶	۰/۵۱۴۷	۰/۴۸۰۷	۰/۰۱۷۳	۰/۵۱۴۷	۰/۵	$\beta_3$		
۰/۰۲۹۶۳	۰/۲۳۷۳	۰/۱۹۳۸	۰/۰۲۶۱۳	۰/۲۳۹۵	۰/۰۱	$\sigma_{\epsilon}^2$		
۰/۱۸۴۸	۰/۱۰۸۱	۰/۰۵۸۹	۰/۰۳۲۲۷	۰/۱۱۱۶	۰/۳	$\sigma_{\omega}^2$		
۰/۳۳۱۶	۰/۱۰۲۶	-۰/۱۱۲۰	۰/۰۰۸۱۲	۰/۱۰۷۳	۰/۱	$a$		
۰/۰۵۵۰	۰/۰۳۵۴	۰/۰۱۸۴	۰/۰۰۹۶	۰/۰۳۵۶	۰/۰۵	$\kappa$		

متغیرهای کمکی این مطالعه بدون در نظر گرفتن متغیر کمکی انجام شده است. مثلث‌بندی این ناحیه بر اساس رئوس اولیه در  $d = 14$  موقعیت ایستگاه‌ها دنبال می‌شود و برای به دست آوردن مثلث‌بندی مناسب رئوس دیگر اضافه می‌شود. برای رسم شکل ۳ و مثلث‌بندی ناحیه فضایی مورد بررسی، بسته *INLA* از نرم‌افزار *R* به کار گرفته می‌شود.

ابتدا بسته *INLA* فراخوانی می‌گردد و از تابع *inla.mesh.create.helper* استفاده می‌شود. این تابع مثلث‌بندی را بر روی دامنه مشخص شده با مرزهایش، با استفاده از مجموعه محدب با ۷ ضلع در فاصله ۱۰ کیلومتری یک مثلث‌بندی ایجاد می‌کند و مثلث‌ها تمام دامنه را برای پیش‌گویی فضایی استان سمنان پوشش می‌دهند. به منظور کاهش اثرات مرزی یک گسترش با ۸ ضلع در یک فاصله ۱۰۰ کیلومتری از مرز منطقه، مثلث‌ها اضافه گردید، لذا گسترش کل منطقه، در ۱۱۰ کیلومتری ارائه شده است. حداکثر طول ضلع مثلث‌ها در داخل منطقه ۵۰ کیلومتر در نظر گرفته می‌شود و در گسترش‌های خارجی به ۱۰۰ کیلومتر می‌رسد، از آنجایی که تعداد مثلث‌های زیاد، باعث افزایش زمان محاسبات می‌شود، پس در خارج از منطقه مورد بررسی، طول ضلع مثلث‌ها بزرگ در نظر گرفته می‌شود تا از صرف زمان بیهوده



شکل ۳: (الف) نقشه استان سمنان (ب) مثلث‌بندی استان سمنان با رهیافت SPDE.

جلوگیری شود. همچنین از آنجایی که گسترش‌های خارج از منطقه بهره عملی ندارد، برای در نظر گرفتن اثرات مرزی به جای گسترش خارج از منطقه‌ای می‌توان از تعیین حداقل زاویه‌ی داخلی مجاز برای مثلث‌ها استفاده کرد. در واقع اثرات مرزی، تنها با تعیین حداقل زاویه‌ی داخلی مجاز برای مثلث‌ها به طرز خوبی مورد عمل قرار می‌گیرد. برای این منظور زاویه‌ی ۲۱ درجه به عنوان کمترین زاویه‌ی مجاز برای مثلث‌های داخل منطقه‌ی سمنان و زاویه‌ی ۲۱ درجه برای گسترش خارج از منطقه‌ی سمنان در نظر گرفته می‌شود. با تغییر مقدار زاویه در تابع مذکور، مثلث‌بندی‌های متفاوت با رئوس بیشتر یا کمتر می‌توان ایجاد کرد. برای رسم شکل ۳، از مقدار ۲۶ درجه برای منطقه داخلی و ۲۱ درجه برای گسترش خارجی استفاده شده است. در واقع هیچ تفاوت عملی در نتایج با استفاده از مثلث‌بندی ریزتر پیدا نمی‌شود، تنها هزینه و زمان محاسبات بالاتر می‌رود.

پس از مثلث‌بندی و مدل‌بندی متغیرهای پنهان با یک میدان تصادفی مارکوفی گاوسی و در نظر گرفتن پیشین‌های معمول برای پارامترها و به کار بردن رهیافت INLA برآورد پارامترها به دست آورده شده است، که نتایج در جدول ۳ ارائه شده است. برای به دست آوردن نقشه پیش‌گویی میزان بارندگی در استان سمنان، یک شبکه  $۸۰ \times ۶۰$  بر روی نقشه استان از ۱۰ تا ۵۵۰ کیلومتر در جهت شرق و از ۳۷۵۰ تا ۴۱۵۰ کیلومتر به سمت شمال در نظر گرفته شده است. برای نمونه نقشه

جدول ۲: برآورد بیز تقریبی پارامترها، (میانگین پسین، انحراف معیار و چندکها)

پارامتر	میانگین			انحراف معیار			چندکها			
	$\beta_0$	$\sigma_\tau^2$	$\sigma_\omega^2$	$a$	$\rho$	$\log(\tau)$	$\log(\kappa)$	$\beta_0$	$\sigma_\tau^2$	$\sigma_\omega^2$
	۱۸/۲۵۸	$1/27 \times 10^{-4}$	۳/۹۵۷	۰/۴۶۹	۹۰/۲۱۹	-۰/۷۹۸	-۳/۴۵۳	۰/۹۷۵	۰/۵	۰/۰۲۵
	۲۶/۱۴۷	$1/48 \times 10^{-4}$	۱۰/۴۷۶	۰/۷۰	۱۲/۵۳۳	-۰/۱۱۸	۰/۱۴۰	۱۸/۲۳۷	۱۰/۴۷۶	۳/۹۵۷
	$5/44 \times 10^{-4}$	$7/08 \times 10^{-5}$	$1/34 \times 10^{-5}$	۶۲/۶۰۱	۹۰/۲۱۹	-۰/۷۹۸	-۳/۴۵۳	$5/44 \times 10^{-4}$	$7/08 \times 10^{-5}$	$1/34 \times 10^{-5}$
	۵۳۳/۳۷۸	۳۸۷/۴۹۷	۲۸۷/۸۶۰	۰/۷۰	۹۰/۲۱۹	-۰/۷۹۸	-۳/۴۵۳	۵۳۳/۳۷۸	۳۸۷/۴۹۷	۲۸۷/۸۶۰
	۰/۶۰۵	۰/۴۶۹	۰/۳۲۹	۰/۷۰	۹۰/۲۱۹	-۰/۷۹۸	-۳/۴۵۳	۰/۶۰۵	۰/۴۶۹	۰/۳۲۹
	۱۱۶/۸۰۷	۸۹/۴۹۷	۶۷/۶۴۳	۱۲/۵۳۳	۹۰/۲۱۹	-۰/۷۹۸	-۳/۴۵۳	۱۱۶/۸۰۷	۸۹/۴۹۷	۶۷/۶۴۳
	-۰/۵۷۸	-۰/۷۹۴	-۱/۰۳۹	۰/۱۱۸	۹۰/۲۱۹	-۰/۷۹۸	-۳/۴۵۳	-۰/۵۷۸	-۰/۷۹۴	-۱/۰۳۹
	-۳/۱۷۲	-۳/۴۵۵	-۳/۷۲۴	۰/۱۴۰	۹۰/۲۱۹	-۰/۷۹۸	-۳/۴۵۳	-۳/۱۷۲	-۳/۴۵۵	-۳/۷۲۴

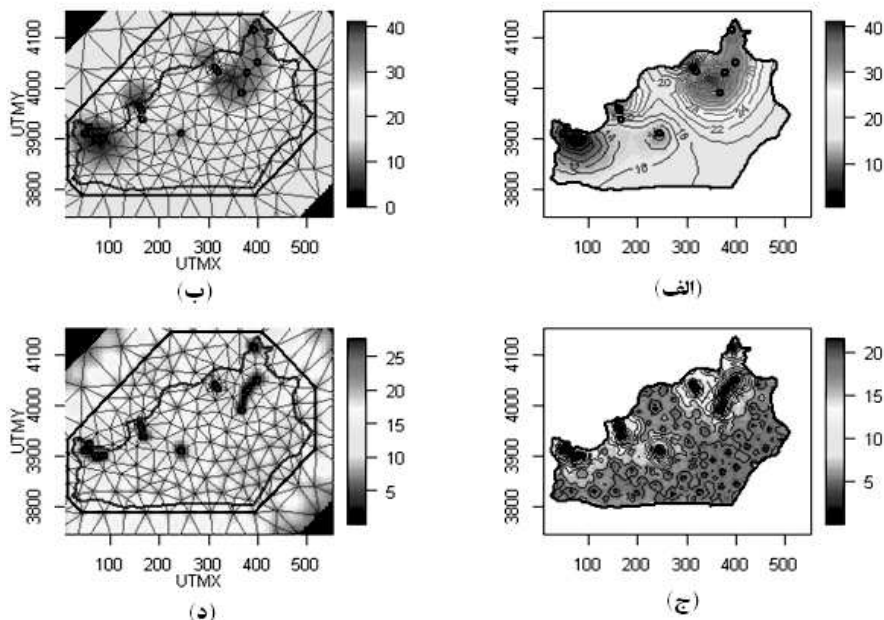
پیش‌گویی میزان بارندگی در استان سمنان برای ماه اردیبهشت در شکل ۴ رسم شده‌است. نقشه پیش‌گویی به‌دست آمده بر روی نقشه سمنان و *SPDE* ناحیه به ترتیب در نمودارهای (الف) و (ب) شکل ۴ ارائه شده است.

نمودارهای (ج) و (د) شکل ۴، به ترتیب نقشه خطای پیش‌گویی بر روی ناحیه مورد بررسی و *SPDE* تشکیل شده، می‌باشند. همان‌طور که در این نمودارها ملاحظه می‌شود مناطق هم‌مرز با استان‌های شمالی و استان خراسان شمالی و منطقه شمال شهر گرمسار از میزان بارندگی بیشتری برخوردار هستند و این میزان در غرب استان به سمت استان خراسان رضوی کاهش پیدا می‌کند. پیش‌گویی‌ها در نوار جنوبی استان از دقت بالایی برخوردار نیست و به دلیل خالی از سکنه بودن این مناطق و عدم وجود ایستگاه هواشناسی در این مناطق نمی‌توان پیش‌گویی مناسبی از میزان بارندگی در این منطقه داشته باشیم.

## ۷ بحث و نتیجه‌گیری

با توجه به پیچیدگی مدل‌های فضایی-زمانی و عدم وجود شکل‌های بسته برای پسین و توزیع‌های شرطی کامل، معمولاً تحلیل این مدل‌ها به راحتی امکان‌پذیر نیست.

در این مقاله وابستگی فضایی-زمانی توسط یک میدان تصادفی گاوسی ماترن مدل‌بندی و سپس از رهیافت *SPDE* استفاده شد و این میدان به یک میدان تصادفی مارکوفی گاوسی تبدیل شد. از خصوصیات خوب این میدان در راستای تسهیل و



شکل ۴: نقشه پیش‌گویی میزان بارندگی در ماه اردیبهشت بر روی (الف) نقشه استان سمنان و (ب) ناحیه مثلث‌بندی ایجاد شده با رهیافت SPDE، نقشه دقت پیش‌گویی بر روی (ج) نقشه استان و (د) ناحیه مثلث‌بندی استان سمنان.

کاهش زمان محاسبات استفاده گردید و رهیافت بیزی معمولی و بیزی تقریبی لاپلاس آشیانی جمع‌بسته برای برآورد پارامترها و پیش‌گویی‌ها ارائه شد. در یک مثال شبیه‌سازی دو روش بیزی معمولی و بیزی تقریبی از نظر زمان مورد مقایسه قرار گرفتند. برای روش بیزی معمولی و استفاده از الگوریتم مونت کارلوی زنجیر مارکوفی، از بسته جدید  $R - spTimer$  که باکار و ساهو (۲۰۱۵) برای تحلیل بیزی داده‌های فضایی- زمانی معرفی کردند، استفاده گردید و روش بیزی تقریبی با تقریب لاپلاس آشیانی جمع‌بسته با بسته  $R - INLA$  انجام شد.

زمان محاسبات با به‌کار بردن رهیافت SPDE و روش بیزی تقریبی لاپلاس آشیانی جمع‌بسته حدود ۱۴۰ ثانیه به طول انجامید. به‌طوری‌که برای رهیافت بیزی معمولی و الگوریتم مونت کارلوی زنجیر مارکوفی با ۱۵۰۰۰ تکرار، توسط یک



کامپیوتر با مشخصات  $IntelCorei7(4GB$  و  $2/7GHz)$  حدود ۴۸۰ ثانیه زمان صرف شد در صورتی که دقت نتایج برای هر دو رهیافت تقریباً معادل است.

نحوه کاربست مدل و روش ارائه شده بر روی یک مجموعه داده واقعی مربوط به داده‌های بارندگی استان سمنان نشان داده شد. حجم نمونه داده‌های مورد بررسی در این مقاله کوچک است در صورتی که در بسیاری از مسائل با مجموعه داده‌ها با حجم بالا سروکار داریم که استفاده از استنباط‌های معمول کلاسیک و بیزی بسیار پیچیده و زمان‌بر است و استفاده از میدان تصادفی مارکوفی گاوسی به دلیل دارا بودن خواص خوب و روش بیزی تقریبی لاپلاس آشیانی جمع‌بسته به دلیل استفاده از تقریب‌های گاوسی و لاپلاس بسیار مفید به نظر می‌رسند.

در این مقاله فقط یک مدل فضایی - زمانی به داده‌ها برازش شد اما می‌توان چند مدل را برای مجموعه داده مورد بررسی در نظر گرفت و پس از مقایسه بین آن‌ها مدل مناسب را انتخاب نمود. به عنوان مثال به نظر می‌رسد با توجه به گستردگی استان سمنان و تنوع در آب و هوای مناطق مختلف این استان، مدل‌هایی با روند شمالی - جنوبی و یا شرقی - غربی مفید باشد. برای پارامترها، توزیع‌های پیشین معمول در نظر گرفته شد و یکی از کارهای مورد نظر نویسندگان بررسی بیشتر روی توزیع‌های پیشین و مطالعه بر روی تاثیر آن‌ها بر روی نتایج است. از معایب اصلی روش‌هایی که تا کنون برای داده‌های فضایی - زمانی ارائه شده است، در نظر گرفتن یک فرض محدود کننده یعنی گاوسی بودن میدان تصادفی است که گاهی در عمل ممکن است فرضی نادرست و گمراه کننده باشد. لذا یکی دیگر از اهداف نویسندگان تعمیم مدل و روش‌های توضیح داده شده برای میدان‌های تصادفی چوله می‌باشد.

## تقدیر و تشکر

نویسندگان مقاله از داوران و سردبیر محترم مجله به خاطر ارائه پیشنهادات ارزنده که سبب بهبود مقاله گردید، کمال تشکر و قدردانی را دارند.

## مراجع

ریواز، ف.، محمدزاده، م. و جعفری خالدی، م. (۱۳۸۶)، پیش‌گویی بیز تجربی برای داده‌های فضایی-زمانی تحت مدل تفکیک‌پذیر، مجله علوم آماری، جلد ۱، ۴۵-۵۹.

قلی‌زاده گزور، ک.، محمدزاده، م. و قیومی، ز. (۱۳۹۲)، تحلیل فضایی رگرسیون جمعی ساختاری و مدل‌بندی داده‌های جرم شهر تهران با تقریب لاپلاس آشیانی جمع‌بسته، مجله علوم آماری، ۷، ۱۲۴-۱۰۳.

Bakar, K. S. and Sahu, S. K. (2015), Spatio-Temporal Bayesian Modeling Using R, *Journal of Statistical Software*, **63**, 1-32.

Banerjee, S., Carlin, B. and Gelfand, A. (2004), *Hierarchical Modeling and Analysis for Spatial Data. Monographs on statistics and Applied Probability*, Chapman and Hall, New York.

Brenner, S. C. and Scott, R. (2007), *The Mathematical Theory of Finite Element Methods*, Springer, New York.

Cameletti, M., Ignaccolo, R. and Bande, S. (2011), Comparing Spatio-Temporal Models for Particulate Matter in Piemonte, *Environmetrics*, **22**, 985-996.

Cameletti, M., Lindgren, F., Simpson, D. and Rue, H. (2013). Spatio-Temporal Modeling of Particulate Matter Concentration through the SPDE Approach, *Advances in Statistical Analysis*, **97**, 109-131.

Cocchi, D., Greco, F. and Trivisano, C. (2007), Hierarchical Space-Time Modelling of PM10 Pollution, *Atmospheric Environment*, **41**, 532-542.

- Cressie, N. and Wikle, C. (2011), *Statistics for Spatio- Temporal Data* , John Wiley, New York.
- De Berg, M., Cheong, O., van Kreveld, M. and Overmars, M. (2008), *Computational Geometry: Algorithms and Applications*, Springer, Berlin.
- Eidsvik, J., Martino, S. and Rue, H. (2009), Approximate Bayesian Inference in Spatial Generalized Linear Mixed Models, *Scandinavian Journal of Statistics*, **36**, 1-22.
- Hjelle, Q. and Daehlen, M. (2006), *Triangulations and Applications*, Springer, Heidelberg.
- Hosseini, F., Eidsvik, J. and Mohammadzadeh, M. (2011), Approximate Bayesian Inference in Spatial GLMM with Skew Normal Latent Variables, *Computational Statistics and Data Analysis*, **55**, 1791-1806.
- Hosseini, F. and Mohammadzadeh, M. (2012), Bayesian Prediction for Spatial GLMM's with Closed Skew Normal Latent Variables , *Australian & New Zealand Journal of Statistics*, **54**, 43-62.
- Gholizadeh, K., Mohammadzadeh, M. and Ghayyomi, Z. (2013), Spatial Analysis of Structured Additive Regression and Modeling of Crime Data in Tehran City using Integrated Nested Laplace Approximation, *Journal of Statistical Science*, **7**, 103-124.
- Guttorp, P. and Gneiting T. (2006), Studies in the History of Probability and Statistics XLIX, on the Matern Correlation Family, *Biometrika*, **93**, 989-995.
- Lindgren, F., Rue, H. and Lindstrom, J. (2011), An explicit link between Gaussian Fields and Gaussian Markov Random Fields: the Stochastic

Partial Differential Equation Approach (with discussion) *Journal of the Royal Statistical Society: Series B* , **73**, 423-498.

Rue, H. and Held, L. (2005), *Gaussian Markov Random Fields, Theory and Applications* , Chapman and Hall, London.

Rue, H. and Martino, S. (2007), Approximate Bayesian Inference for Hierarchical Gaussian Markov Random fields models, *Journal of Statistical Planning and Inference*, **137**, 3177-3199.

Rue, H. and Martino, S., and Chopin, N. (2009), Approximate Bayesian Inference for Latent Gaussian Model by Using Integrated Nested Laplace Approximations (with discussion), *Journal of the Royal Statistical Society: Series B* , **71**, 319-392.

Sahu, S. K. (2012), Hierarchical Bayesian Models for Space-Time Air Pollution Data, *In Handbook of Statistics-Vol 30. Time Series Analysis, Methods and Applications*, Editors: T Subba Rao and C R Rao. Elsevier Publishers, Holland. Elsevier Publishers, Amsterdam, 477-495.

Whittle, P. (1954), On Stationary Processes in the Plane. *Biometrika*, **41**, 434-449.

Whittle, P. (1963), Stochastic Processes in Several Dimensions, *Bulletin of the International Statistical Institute*, **40**, 974-994.

Wikle, CK. (2003), Hierarchical Models in Environmental Science. *International Statistical Review*, **71**, 181-199.

Wikle, CK., Berliner, L. and Cressie, N. (1998), Hierarchical Bayesian Space Time Models, *Environmental and Ecological Statistics*, **5**, 117-154.