



Bayesian Analysis of Skewed Spatio-Temporal Data Using Bayesian Neural Field and Variational Inference

Hosseini, F. , Karimi, O. 

Department of Statistics, Semnan University, Semnan, Iran.

Corresponding author: Fatemeh Hosseini, fatemeh.hosseini@semnan.ac.ir

Received: 20/9/2025 Revised: 23/2/2026 Accepted and Published Online: 25/2/2026.

Introduction

Spatio-temporal data arise in fields such as climatology, air pollution, epidemiology, and environmental sciences. These data typically exhibit complex dependence structures, nonstationarity, and skewness, which challenge conventional statistical models. While Gaussian random fields are widely used for spatio-temporal dependence, their symmetry and light-tailed assumptions make them inadequate for asymmetric data. In addition, fully Bayesian inference based on sampling methods can be computationally intensive in high-dimensional settings. To overcome these limitations, we propose a novel framework that combines Bayesian Neural Fields (BNF) with the Flexible Closed Skew-Normal (FCSN) distribution to model asymmetric spatio-temporal data, improving predictive performance, flexibility, and computational scalability within a coherent Bayesian framework.

Material and Methods

The proposed model follows an additive structure, decomposing the spatio-temporal response into a nonlinear mean component and a stochastic residual process. The mean is modeled via a BNF, capturing complex nonlinear relationships between spatial-temporal coordinates and covariates using a Bayesian neural network with Gaussian priors on the weights. The residuals are assumed to follow an FCSN distribution to accommodate skewness and asymmetric variability. Spatial-temporal dependence is modeled through a separable Matérn covariance function. Bayesian inference is performed using variational inference by maximizing the Evidence Lower Bound (ELBO), with MCMC estimation implemented in simulation studies for benchmark-

ing. The neural network includes two hidden layers with hyperbolic tangent activation functions, balancing flexibility and computational stability.

Results and Discussion

Simulation studies under moderate and high skewness show that the proposed BNF-FCSN model, estimated via variational inference, achieves lower prediction error, near-nominal 95% credible interval coverage, and substantially reduced computation time compared to the classical FCSN model estimated by MCMC. These advantages become more pronounced as skewness increases. In the real-data application using ozone measurements, exploratory analysis reveals positive skewness and departures from Gaussianity. Comparing the Gaussian BNF model with the FCSN-based BNF model indicates that incorporating skewness improves predictive accuracy and enhances predictive interval coverage. Overall, the results underscore the importance of explicitly modeling asymmetry in environmental spatio-temporal data.

Conclusion

This study proposes a scalable Bayesian framework for asymmetric spatio-temporal data by integrating Bayesian Neural Fields with the FCSN distribution. By decomposing the response into a nonlinear mean component and a skewed stochastic process, the model maintains interpretability while increasing flexibility. Variational inference enables efficient and stable parameter estimation, reducing the computational cost associated with sampling-based methods. Results from simulations and real data demonstrate improved predictive accuracy and uncertainty quantification compared with Gaussian and MCMC-based models. The framework applies to environmental, climatological, and epidemiological studies and can be extended to nonseparable covariance structures and time-varying skewness mechanisms.

Keywords: Bayesian Neural Field, Spatio-Temporal Data, Variational Inference, Skewed Distributions.

Mathematics Subject Classification (2020): 62F15, 62M10.



©The Author(s). The Publisher is Iranian Statistical Society.

This is an open access article distributed under the terms and conditions of [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/)

تحلیل بیزی داده‌های فضایی-زمانی چوله با استفاده از میدان عصبی بیزی و استنباط واریاسیونی

فاطمه حسینی، امید کریمی

گروه آمار، دانشگاه سمنان

نویسنده مسئول: فاطمه حسینی، fatemeh.hoseini@semnan.ac.ir

تاریخ دریافت: ۱۴۰۴/۶/۲۹ تاریخ بازنگری: ۱۴۰۴/۱۲/۴ تاریخ پذیرش و انتشار: ۱۴۰۴/۱۲/۶

چکیده: داده‌های فضایی-زمانی معمولاً دارای ساختارهای وابستگی پیچیده و چولگی هستند که مدل‌سازی آن‌ها با چارچوب‌های کلاسیک، مانند میدان‌های تصادفی گاوسی، یا بسیار پرهزینه و یا بیش از حد محدودکننده است. در این مقاله، چارچوبی نوین با عنوان میدان عصبی بیزی برای مدل‌سازی فرایندهای فضایی-زمانی چوله معرفی می‌شود. این چارچوب با استفاده از مختصات مکانی و زمانی در کنار متغیرهای توضیحی و توزیع‌های پیشین، امکان بازنمایی انعطاف‌پذیر همبستگی‌ها و چولگی را فراهم می‌سازد و قابلیت پیش‌بینی در مکان‌ها و زمان‌های جدید را دارد. استنباط پارامترها از طریق روش استنباط واریاسیونی انجام شد که علاوه بر بهره‌وری محاسباتی، برآورد عدم قطعیت را نیز امکان‌پذیر می‌سازد. نتایج مطالعه شبیه‌سازی نشان داد که چارچوب پیشنهادی از نظر دقت و سرعت نسبت به روش‌های مونت‌کارلویی استاندارد برتری دارد.

واژه‌های کلیدی: میدان عصبی بیزی، داده‌های فضایی-زمانی، استنباط واریاسیونی، توزیع‌های چوله

کد موضوع‌بندی ریاضی (۲۰۲۰): 62M10 ، 62F15



۱ مقدمه

داده‌های فضایی-زمانی، که در موقعیت‌های مکانی و زمانی مشخص گردآوری می‌شوند، در حوزه‌هایی چون اقلیم‌شناسی، اپیدمیولوژی، کشاورزی و علوم محیطی نقش کلیدی ایفا می‌کنند (کُرسی و ویکل (۲۰۱۱)). این داده‌ها اغلب دارای ویژگی‌هایی مانند چولگی، ناهمگنی و ساختارهای پیچیده همبستگی هستند که مدل‌سازی آن‌ها را با چالش‌های جدی مواجه می‌سازد ساهو (۲۰۲۲). یکی از رویکردهای متداول برای تحلیل این داده‌ها استفاده از میدان‌های تصادفی گاوسی است که به دلیل انعطاف‌پذیری در مدل‌سازی همبستگی‌ها کاربرد گسترده‌ای یافته‌اند (کُرسی و ویکل، ۲۰۱۱؛ ساهو، ۲۰۲۲). این مدل‌ها با محدودیت‌هایی همچون عدم شناسایی‌پذیری پارامترها، هزینه محاسباتی بالا در چارچوب بیزی (بویلاکوا و همکاران، ۲۰۲۰؛ موزن و همکاران، ۲۰۲۳؛ سعد و همکاران، ۲۰۲۴) و اتکای شدید به فرض گاوسی بودن داده‌ها مواجه‌اند؛ در حالی‌که در بسیاری از کاربردهای واقعی، این فرض برقرار نیست گنتون (۲۰۰۴). برای رفع این محدودیت، خانواده‌ای از مدل‌های چوله معرفی شده‌اند. به‌عنوان مثال، توزیع چوله گاوسی توسط آزالینی و کاپیتانیو (۲۰۰۵) و توزیع چوله تی توسط آرانو و همکاران (۲۰۲۱) معرفی شدند که امکان مدل‌سازی عدم تقارن و رفتار دنباله‌های سنگین را فراهم می‌سازند. در ادامه، میدان‌های تصادفی چوله گاوسی به‌عنوان تعمیمی از میدان‌های گاوسی پیشنهاد شدند.

آلارد و ناویو (۲۰۰۷) و ریستاد و اومره (۲۰۱۴) برای نخستین بار میدان تصادفی چوله گاوسی تقریباً مانا را معرفی کردند و سپس نسخه‌های مانای این مدل توسعه داده شد (کریمی و حسینی، ۱۴۰۰). این ایده‌ها در چارچوب بیزی تقریبی برای برآورد مدل‌های آمیخته خطی تعمیم‌یافته فضایی نیز به‌کار گرفته شدند (حسینی و کریمی، ۱۴۰۳). در همین راستا، یک زیررده انعطاف‌پذیر از توزیع چوله گاوسی بسته^۱ (CSN) معرفی شد که به‌واسطه آن می‌توان میدان‌های تصادفی فضایی‌ای تعریف کرد که پارامترهایشان قابل شناسایی و تفسیر باشند (مارکز و گنزالس، ۲۰۲۲). این زیررده که با عنوان توزیع چوله گاوسی بسته منعطف^۲ (FCSN) شناخته می‌شود، ساختاری مشابه سایر توزیع‌های چوله دارد (حسینی و کریمی، ۲۰۲۱)، اما انعطاف بیشتری در مدل‌سازی ساختارهای همبستگی فضایی فراهم می‌آورد. به‌دنبال آن، کریمی و حسینی (۱۴۰۲) ویژگی‌های میدان‌های فضایی چوله بر اساس توزیع FCSN را بررسی و مدلی مناسب برای داده‌های فضایی نامتقارن ارائه کردند. همچنین، کریمی و حسینی (۱۴۰۳) با بهره‌گیری از این چارچوب، تحلیل بیزی سلسله‌مراتبی رگرسیون فضایی را مطرح کرده و با استفاده از رهیافت‌های بیز تقریبی، هزینه محاسباتی برآورد را کاهش دادند. با وجود این تلاش‌ها، برآورد پارامترهای مدل‌های چوله همچنان پیچیده است و روش‌های کلاسیک مانند زنجیر مارکوف مونت‌کارلو (MCMC) یا مونت‌کارلو همیلتونی (نیل، ۲۰۱۱؛ هوفمن و گلن، ۲۰۱۴) اگرچه از نظر تنوری دقیق‌اند، اما در ابعاد بالا زمان‌بر بوده و به تنظیمات ظریف نیاز دارند بتانکورت (۲۰۱۷). از سوی دیگر، تکیه بر توابع همبستگی از پیش تعیین‌شده (مانند ماترن یا نمایی) ممکن است توانایی مدل در ثبت ساختارهای غیرایستا یا روابط غیرخطی را محدود سازد (بویلاکوا و همکاران، ۲۰۲۰). برای

¹Closed Skew Normal²Flexible Closed Skew Normal

غلبه بر این محدودیت‌ها، در دهه اخیر شبکه‌های عصبی بیزی^۱ (BNNs) به عنوان ابزاری توانمند معرفی شده‌اند (بلاندل و همکاران، ۲۰۱۵؛ جاسپن و همکاران، ۲۰۲۲). این شبکه‌ها با اعمال پیشین‌های احتمالاتی روی وزن‌ها، قابلیت یادگیری روابط غیرخطی پیچیده را بدون نیاز به فرض ساختار همبستگی خاص فراهم می‌کنند و در عین حال، امکان کمی‌سازی عدم قطعیت را نیز مهیا می‌سازند. اما محاسبه پسین در این شبکه‌ها به‌طور تحلیلی ممکن نیست و باید تقریب زده شود. در این میان، استنباط واریاسیونی^۲ (VI) به عنوان جایگزینی سریع‌تر و مقیاس پذیرتر نسبت به MCMC مطرح شده است (بلای و همکاران، ۲۰۱۷). این روش با بیشینه‌سازی حد پایین شواهد^۳ (ELBO) علاوه بر سرعت محاسباتی بالا، امکان تحلیل مجموعه داده‌های بزرگ و پیچیده را نیز فراهم می‌سازد.

اگرچه شبکه‌های عصبی بیزی چارچوب کلی برای بیزی‌سازی وزن‌های شبکه ارائه می‌دهند، اما در داده‌های فضایی-زمانی نامنظم، اصطلاح و چارچوب جدیدتری با نام میدان عصبی بیزی^۴ (BNF) مطرح شده است. در این رویکرد، میدان تصادفی به‌طور مستقیم روی مختصات پیوسته فضا-زمان تعریف می‌شود.

در این مقاله مدل BNF مورد استفاده قرار می‌گیرد و برای بازنمایی چولگی داده‌ها، مؤلفه تصادفی با توزیع FCSN مدل‌سازی خواهد شد. در این ترکیب، بخش میانگین ساختاری که توسط شبکه عصبی بیزی روی مختصات و متغیرهای توضیحی یاد گرفته می‌شود از بخش خطا که ویژگی‌های چوله را در بر می‌گیرد جدا شده است. این جداسازی هم تفسیر نتایج را ساده‌تر می‌کند و هم باعث می‌شود همبستگی‌ها در حضور داده‌های نامتوازن بهتر برآورد شوند. برآورد بیزی پارامترها از طریق VI و بیشینه‌سازی ELBO انجام می‌شود (بیگام و همکاران، ۲۰۱۹).

ساختار مقاله به این صورت است، ابتدا به معرفی کوتاه چارچوب BNF و روش استنباط VI می‌پردازیم و تفاوت آن را با BNN عمومی توضیح می‌دهیم. سپس مدل پیشنهادی یعنی ترکیب BNF با توزیع FCSN شرح داده می‌شود. بعد از آن جزئیات استنباط و پیاده‌سازی ارائه خواهد شد. در ادامه نتایج مطالعه شبیه‌سازی و یک مطالعه موردی گزارش می‌شود. در پایان نیز جمع‌بندی و پیشنهادهایی برای پژوهش‌های آینده مطرح خواهد شد.

۲ شبکه‌های عصبی بیزی، میدان عصبی بیزی و استنباط واریاسیونی

ایده اصلی در شبکه عصبی بیزی این است که به جای ثابت گرفتن وزن‌های شبکه، برای آن‌ها توزیع‌های پیشین در نظر بگیریم. به این ترتیب، پس از مشاهده داده‌ها، به جای یک جواب نقطه‌ای، یک توزیع پسین روی وزن‌ها (و به تبع آن، روی خروجی‌ها) به دست می‌آید. این کار امکان کمی‌سازی عدم قطعیت را فراهم می‌کند و برای داده‌های پرنویز، کم‌نمونه یا با ساختارهای پیچیده سودمند است. به‌طور خلاصه، در BNN ورودی‌ها به شبکه وارد می‌شوند، وزن‌ها تصادفی (با پیشین و پسین) مدل می‌شوند و خروجی نهایی به صورت یک مقدار احتمالاتی گزارش می‌شود. کاربرد شبکه‌های عصبی بیزی در داده‌های فضایی-زمانی اهمیت بسزایی دارد، زیرا این نوع داده‌ها اغلب دارای ساختارهای همبستگی

¹Bayesian Neural Networks

²Variational Inference

³Evidence Lower Bound

⁴Bayesian Neural Field

پیچیده، ناهمگنی و گاهی چولگی هستند. BNNs با انعطاف بالا و توانایی یادگیری روابط غیرخطی می توانند این ساختارها را به طور کارآمد مدل سازی کنند. به عنوان نمونه، چارچوب میدان عصبی بیزی به عنوان مدلی مقیاس پذیر برای تحلیل داده های فضایی-زمانی معرفی شده است که ترکیبی از شبکه های عمیق و استنباط بیزی سلسله مراتبی را به کار می گیرد و برای پیش بینی و درون یابی در حوزه هایی مانند کیفیت هوا و سلامت عمومی نتایج بسیار دقیقی ارائه کرده است (سعد و همکاران، ۲۰۲۴). همچنین، در مدل های شبکه های عصبی بیزی فضایی^۱ (SBNNs) لایه های ویژه ای برای کدگذاری وابستگی های مکانی طراحی شده اند (زمیت و همکاران، ۲۰۲۴). در کنار این ها، توسعه شبکه های بازگشتی بیزی نیز نشان داده است که این چارچوب می تواند در مدل سازی سری های زمانی مکانی و تحلیل پویایی های غیرخطی همچون داده های آب و هوایی یا انتشار بیماری ها بسیار اثربخش باشد (فورتوناتو و همکاران، ۲۰۱۷). در مجموع، شبکه های عصبی بیزی ابزاری قدرتمند برای مدل سازی داده های فضایی-زمانی هستند، زیرا علاوه بر توانایی یادگیری روابط غیرخطی پیچیده، برآورد عدم قطعیت معتبر را نیز فراهم می کنند. این ویژگی موجب می شود که BNNs نه تنها در تحلیل های پیش بینی و درون یابی، بلکه در تصمیم گیری های حساس مبتنی بر داده های مکانی-زمانی نیز به عنوان روشی کارآمد و قابل اعتماد شناخته شوند. در داده های فضایی-زمانی چارچوب میدان عصبی بیزی دقیق تر و مناسب تر است (سعد و همکاران، ۲۰۲۴). در BNF میدان تصادفی فضایی-زمانی به طور مستقیم با یک نگاشت عصبی بیزی به صورت

$$F(s, t) = f_{\theta}(s, t, x(s, t)), \quad Y(s, t) | F(s, t) \sim \text{Dist}(g(F(s, t)), \gamma),$$

تعریف می شود، که در آن (s, t) مختصات مکانی-زمانی، $x(s, t)$ متغیرهای کمکی، θ پارامترهای شبکه (با پیشین های بیزی)، g تابع لینک و Dist توزیع مشاهدات است. تفاوت مفهومی با BNN عمومی این است که در BNF نگاه میدانی داریم، مدل از ابتدا برای بازنمایی یک میدان تصادفی روی فضا-زمان طراحی شده است. هدف در چارچوب بیزی محاسبه توزیع پسین

$$p(\theta | D) = \frac{p(D | \theta) p(\theta)}{p(D)},$$

است که نشان دهنده عدم قطعیت در پارامترها (θ) پس از مشاهده داده ها (D) می باشد. با این حال، محاسبه مستقیم این توزیع در عمل ناممکن است؛ زیرا $p(D) = \int p(D | \theta) p(\theta) d\theta$ شامل یک انتگرال پربعد و پیچیده است. به همین دلیل، استنباط واریاسیونی معرفی می شود تا این مشکل برطرف شود. ایده اصلی آن است که به جای محاسبه مستقیم پسین، یک خانواده ساده تر از توزیع ها با نماد $q_{\phi}(\theta)$ در نظر گرفته شود که توسط پارامترهای آزاد ϕ مشخص می گردد. سپس، این توزیع تقریبی به گونه ای آموزش داده می شود که تا حد امکان به پسین واقعی نزدیک باشد. برای این منظور از معیار واگرایی کولبک-لایبلر^۲ (KL) استفاده می شود. از آنجا که محاسبه مستقیم KL با پسین واقعی

¹Spatial Bayesian Neural Networks

²Kullback-Leibler Divergence

امکان‌پذیر نیست، یک کران پایین موسوم به کران پایین شواهد یا ELBO تعریف می‌شود. بیشینه‌سازی این کران پایین معادل با کمینه‌سازی فاصله KL بین توزیع تقریبی $q_\phi(\theta)$ و پسین واقعی است. بنابراین، داده‌ها D مستقیماً به پسین واقعی متصل هستند اما محاسبه آن امکان‌پذیر نیست. در عوض، یک توزیع تقریبی $q_\phi(\theta)$ معرفی می‌شود که با استفاده از داده‌ها و بهینه‌سازی ELBO به پسین نزدیک می‌گردد. نتیجه، تقریب محاسباتی کارآمدی از پسین است که در مقایسه با روش‌های نمونه‌گیری مانند MCMC سرعت و مقیاس‌پذیری بیشتری دارد و برای مدل‌های پیچیده و داده‌های حجیم بسیار مناسب است.

۳ مدل فضایی-زمانی چوله منعطف

در این بخش چارچوب کامل مدل معرفی می‌شود. ایده اصلی آن است که میانگین میدان فضا-زمانی با استفاده از چارچوب میدان عصبی بیزی بازنمایی گردد و مولفه تصادفی که عدم تقارن و ناهمگنی را بازتاب می‌دهد، با توزیع FCSN مدل‌سازی شود. این ترکیب موجب می‌شود هم روابط غیرخطی پیچیده مکانی-زمانی به‌صورت انعطاف‌پذیر مدل‌سازی شوند و هم رفتار چوله به‌درستی بازنمایی شوند. برآورد پارامترهای هر دو بخش به‌صورت یکپارچه در چارچوب بیزی و با استنباط واریاسیونی انجام می‌گیرد.

۳.۱ میدان تصادفی چوله گاوسی بسته منعطف

توزیع CSN یک تعمیم انعطاف‌پذیر از توزیع گاوسی چندمتغیره است که امکان مدل‌سازی داده‌های نامتقارن را فراهم می‌سازد. فرض کنید $U = (U_1', U_2')' \in \mathbb{R}^{p+q}$ یک بردار تصادفی با توزیع گاوسی چندمتغیره به‌صورت $U \sim \mathcal{N}_{p+q} \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right)$ باشد، که در آن $\mu_1 \in \mathbb{R}^p$ ، $\mu_2 \in \mathbb{R}^q$ و ماتریس کوواریانس به‌صورت بلوکی تعریف شده است. بنابراین بردار تصادفی $V = [U_1 \mid U_2 \leq 0]$ دارای توزیع CSN است که با $\text{CSN}_{p,q}(\mu_1, \Sigma_{11}, \Sigma_{21}\Sigma_{11}^{-1}, \mu_2, \Sigma_{21})$ نشان داده می‌شود و چگالی آن به صورت

$$f_V(v) = k \phi_p(v; \mu_1, \Sigma_{11}) \Phi_q(0; \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(v - \mu_1), \Sigma_{21}),$$

است، که در آن $\Sigma_{21} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ ، $\phi_p(\cdot)$ چگالی گاوسی چندمتغیره، $\Phi_q(\cdot)$ توزیع تجمعی گاوسی چندمتغیره، و ثابت نرمال‌سازی $k^{-1} = \Phi_q(0; \mu_2, \Sigma_{22})$ است. بر مبنای این ساختار، مارکز و گزالس (۲۰۲۲) زیررده‌ای انعطاف‌پذیر از CSN را معرفی کرد که با عنوان توزیع چوله گاوسی بسته منعطف شناخته می‌شود. به‌طور مشخص، اگر

$$Z_n \sim \text{CSN}_{n,n}(\mu - b\delta\sigma\tau \Sigma_n^{1/2} \mathbf{1}_n, \sigma^2\tau^2\Sigma_n, \frac{\lambda}{\sigma\tau}\Sigma_n^{-1/2}, \mathbf{0}, I_n),$$

با مقادیر $b = \sqrt{2/\pi}$, $\delta = \lambda(1 + \lambda^2)^{-1/2}$, $\tau = (1 - b^2\delta^2)^{-1/2}$, $\sigma > 0$ و $\lambda \in \mathbb{R}$, آنگاه $Z_n \sim \text{FCSN}_n(\mu, \Sigma_n, \sigma, \lambda)$ و داریم $E(Z_n) = \mu$ و $\text{Var}(Z_n) = \sigma^2 \Sigma_n$. چگالی Z_n به صورت

$$f_{Z_n}(z) = \tau^n \phi_n(z; \mu - b\delta\sigma\tau \Sigma_n^{-1/2} \lambda_n, \sigma^2 \tau^2 \Sigma_n) \\ \times \Phi_n\left(\frac{\lambda}{\sigma\tau} \Sigma_n^{-1/2}(z - \mu) + b\delta\lambda_n; \mathbf{0}, I_n\right)$$

بیان می‌شود؛ که در آن λ شدت چولگی را کنترل می‌کند. برای مدل‌سازی داده‌های فضایی-زمانی، فرض کنید $Z_n = (Z(s_1, t_1), \dots, Z(s_n, t_n))'$ بردار مقادیر میدان تصادفی در نقاط (s_i, t_i) باشد. در این صورت $Z_n \sim \text{FCSN}_n(X\beta, \Sigma_n, \sigma, \lambda)$ که در آن X ماتریس طراحی $n \times p$, $\beta \in \mathbb{R}^p$ ضرایب رگرسیونی، $\sigma > 0$ پارامتر مقیاس و $\lambda \in \mathbb{R}$ پارامتر چولگی است. در این مدل، Σ_n ماتریس کوواریانس بوده و به صورت $\Sigma_n = \sigma^2 C_n$ تعریف می‌شود؛ که در آن C_n ماتریس همبستگی فضایی-زمانی است. مولفه‌های C_n به شکل $c_{ij} = \rho(\|s_i - s_j\|, |t_i - t_j|; \varphi_s, \varphi_t, \nu_s, \nu_t)$ تعریف می‌شوند؛ که در آن $\rho(h, t; \varphi_s, \varphi_t, \nu_s, \nu_t)$ یک تابع همبستگی معتبر تفکیک‌پذیر به صورت

$$\rho(h, t; \varphi_s, \varphi_t, \nu_s, \nu_t) = \rho_s(h; \varphi_s, \nu_s) \rho_t(t; \varphi_t, \nu_t),$$

است و به صورت استاندارد (تابع مترن) داریم

$$\rho_s(h; \varphi_s, \nu_s) = \frac{1}{\sqrt{\nu_s} \Gamma(\nu_s)} \left(\frac{h}{\varphi_s}\right)^{\nu_s} K_{\nu_s}\left(\frac{h}{\varphi_s}\right), \\ \rho_t(t; \varphi_t, \nu_t) = \frac{1}{\sqrt{\nu_t} \Gamma(\nu_t)} \left(\frac{t}{\varphi_t}\right)^{\nu_t} K_{\nu_t}\left(\frac{t}{\varphi_t}\right),$$

که در آن $K_\nu(\cdot)$ تابع بسل اصلاح‌شده نوع دوم است (ساحو، ۲۰۲۲). این جداسازی نقش‌ها بدین معناست که C_n ساختار همبستگی مقیاس‌واحد را تعیین می‌کند و $\Sigma_n = \sigma^2 C_n$ کوواریانس کامل را می‌سازد. چنین پارامتردهی‌ای هم تفسیرپذیری را افزایش می‌دهد و هم برآورد را در حضور چولگی λ پایدارتر می‌کند. تابع مترن دو پارامتر کلیدی دارد که رفتار میدان فضایی-زمانی را کنترل می‌کنند. لذا استفاده از مترن با پارامترهای $(\varphi_\bullet, \nu_\bullet)$ انعطاف کافی برای توصیف همواری‌ها و مقیاس‌های متفاوت را فراهم می‌سازد و نسبت به همبستگی‌های ساده‌تری مانند نمایشی $(\nu_s = \nu_t = 0.5)$ دقت بیشتری در بازنمایی ساختارهای فضایی-زمانی ارائه می‌کند (راویشنکر و همکاران، ۲۰۲۳). در تابع مترن، پارامتر ν درجه همواربودن میدان را مشخص می‌کند و پارامتر φ تعیین‌کننده برد یا مقیاس همبستگی است. هر دو پارامتر مستقیماً شکل و رفتار ساختار کوواریانس را کنترل می‌کنند. هرچه φ بزرگ‌تر باشد، همبستگی بین نقاط مکانی-زمانی با فاصله بیشتری پایدار می‌ماند و هرچه φ کوچک‌تر باشد، همبستگی خیلی سریع با فاصله از بین می‌رود. میزان همواری یا صافی فرآیند را تعیین می‌کند. به علاوه هرچه ν بزرگ‌تر باشد، میدان

هموارتر، مشتق‌پذیرتر و با تغییرات ملایم‌تر است و مقادیر کوچک ν میدان را ناهموار و شبیه نویز با تغییرات شدید می‌کند. پس این پارامتر کنترل‌کننده درجه همواربودن فرآیند است.

۳.۲ مدل رگرسیون فضایی-زمانی با چارچوب ترکیبی BNF-FCSN

برای داده‌های $Y_n = (Y(s_1, t_1), \dots, Y(s_{n_s}, t_{n_t}))'$ در موقعیت‌های فضایی-زمانی به صورت

$$Y(s_i, t_j) = m(s_i, t_j) + Z(s_i, t_j), \quad i = 1, \dots, n_s, \quad j = 1, \dots, n_t, \quad (1)$$

تعریف می‌شود، که در آن میانگین میدان است که توسط یک نگاشت میدان عصبی بیزی به صورت

$$m(s, t) = f_\theta(s, t, x(s, t)), \quad \theta \sim p(\theta),$$

مدل می‌شود، که در آن f_θ یک شبکه عصبی با وزن‌های بیزی و $x(s, t)$ متغیرهای توضیحی است. $Z(s_i, t_i)$ مؤلفه تصادفی فضایی زمانی است که برای ثبت عدم تقارن و دنباله‌های سنگین با توزیع FCSN به صورت $Z_n \sim \text{FCSN}_n(\mathbf{0}, \Sigma_n, \sigma, \lambda)$ مدل می‌شود. بر این اساس، توزیع مشترک پاسخ‌ها Y_n به صورت

$$f_\theta(y) = \tau^n \phi_n(y; m_n - b\delta\sigma\tau C_n^{-1/2} \mathbf{1}_n, \sigma^2 \tau^2 C_n) \times \Phi_n\left(\frac{\lambda}{\sigma\tau} C_n^{-1/2} (y - m_n) + b\delta\lambda \mathbf{1}_n; \mathbf{0}, I_n\right), \quad (2)$$

نوشته می‌شود که در آن $m_n = (m(s_1, t_1), \dots, m(s_{n_s}, t_{n_t}))'$ خروجی BNF در نقاط داده است. با این چارچوب ترکیبی میانگین میدان به جای یک رابطه خطی ساده، توسط BNF مدل می‌شود و بنابراین الگوهای غیرخطی پیچیده و ناهمگن قابل مدل‌سازی‌اند؛ مؤلفه تصادفی FCSN امکان مدل‌سازی عدم تقارن را فراهم می‌کند؛ استفاده از تابع همبستگی مترن در C_n انعطاف کافی برای ثبت ساختارهای فضایی-زمانی گوناگون را فراهم می‌سازد؛ جدایی میانگین و واریانس از پارامتر چولگی λ موجب پایداری بیشتر برآوردها در داده‌های نامتقارن می‌شود (حسینی و کریمی، ۱۴۰۳). در بخش‌های بعدی، به منظور برآورد مؤثر پارامترها، از استنباط واریاسیونی در چارچوب بیزی استفاده خواهد شد.

۴ برآورد بیزی پارامترهای مدل

برای برآورد پارامترهای مدل رگرسیون فضایی-زمانی چوله تعریف‌شده در مدل (۱)، یک چارچوب بیزی مبتنی بر میدان عصبی بیزی و روش استنباط واریاسیونی ارائه می‌شود. به منظور مقایسه این روش با روش‌های معمول، ابتدا

روش های استاندارد مونت کارلویی زنجیر مارکوفی مانند متروپلیس-هستینگس و گیبز به کار گرفته شده اند. هدف، برآورد بردار پارامترها $\theta = (\beta, \sigma, \varphi_s, \varphi_t, \nu_s, \nu_t, \lambda)'$ با استفاده از توزیع های پیشین مناسب و ارزیابی کارایی دو روش است. چگالی پسین به صورت $p(\theta|y) \propto f_\theta(y|\theta) \cdot p(\theta)$ تعریف می شود، که در آن $f_\theta(y|\theta)$ تابع چگالی احتمال داده ها (۲) و $p(\theta)$ توزیع پیشین پارامترها به صورت $p(\theta) = p(\beta)p(\sigma)p(\lambda)p(\varphi_s)p(\nu_s)p(\varphi_t)p(\nu_t)$ است. در این مطالعه پیشین ها به صورت

- $\beta \sim \mathcal{N}_p(\mathbf{0}, \text{diag}(\sigma_{\beta_1}^2, \dots, \sigma_{\beta_p}^2))$ ، توزیع گاوسی چندمتغیره با واریانس های مستقل،
- $\sigma \sim \text{Half-Cauchy}(\cdot, \gamma_\sigma)$ ، برای تضمین مثبت بودن σ ،
- $\lambda \sim \mathcal{N}(\cdot, \sigma_\lambda^2)$ ، محدود به بازه $[-5, 5]$ ، برای کنترل چولگی،
- $\varphi_s \sim \Gamma(2, \cdot/5)$ ، محدود به بازه $(\cdot, 10]$ ، برای پارامتر مقیاس فضایی،
- $\varphi_t \sim \Gamma(3, \cdot)$ ، محدود به بازه $(\cdot, 5]$ ، برای پارامتر مقیاس زمانی،
- $\nu_s \sim \Gamma(1.5, \cdot/2)$ ، محدود به بازه $(\cdot/5, 3]$ ، برای پارامتر همواری فضایی،
- $\nu_t \sim \Gamma(1.5, \cdot/2)$ ، محدود به بازه $(\cdot/5, 3]$ ، برای پارامتر همواری زمانی،

در نظر گرفته شده اند. **روبو (۲۰۲۰)** استفاده از این پیشین ها برای مدل های فضایی-زمانی بیزی توصیه می کند، که در این مطالعه نیز از پیشین های پیشنهادی در منبع مذکور استفاده شده است. لازم به ذکر است به منظور بررسی پایداری مدل نسبت به انتخاب پیشین ها، تحلیل حساسیت نیز انجام شد. در این تحلیل، چندین مجموعه ابرپارامتر گاما و بازه های متفاوتی برای پارامتر چولگی امتحان گردید. نتایج به دست آمده نشان داد که شاخص های کلیدی پیش بینی و ساختار میانگین تخمینی میدان فضایی-زمانی نسبت به این تغییرات پایدار بوده و انتخاب پیشین ها اثر معناداری بر رفتار مدل نهایی ندارد.

۴.۱ برآورد بیزی پارامترها با الگوریتم های MCMC

در حالت پایه، اگر میانگین مدل به صورت خطی X/β باشد، می توان برآورد پارامترهای مدل فضایی-زمانی چوله معطف را با استفاده از الگوریتم های MCMC مانند نمونه گیری گیبز و متروپلیس-هستینگز انجام داد. در این حالت، با نوشتن چگالی های شرطی کامل برای پارامترها $\theta = (\beta, \sigma, \lambda, \varphi_s, \nu_s, \varphi_t, \nu_t)$ می توان در هر گام زنجیره، از توزیع های شرطی مربوطه نمونه گیری کرد. برای پارامترهایی که چگالی شرطی کامل آن ها به صورت تحلیلی در دسترس نیست (مانند λ یا پارامترهای همبستگی)، از گام های متروپلیس-هستینگز استفاده می شود. مراحل کلی الگوریتم ترکیبی عبارتند از:

۱. مقداردهی اولیه $\theta^{(0)}$.

۲. در هر تکرار k ، به ترتیب پارامترها به روزرسانی می شوند:

- (پارامترهای دیگر، $\beta^{(k)} \sim p(\beta|y$

$$\begin{aligned} \bullet \quad \sigma^{(k)} &\sim p(\sigma|\mathbf{y}, \dots) \\ \bullet \quad \lambda^{(k)} &\sim p(\lambda|\mathbf{y}, \dots) \\ \bullet \quad (\varphi_s^{(k)}, \nu_s^{(k)}, \varphi_t^{(k)}, \nu_t^{(k)}) &\sim p(\cdot|\mathbf{y}, \dots) \end{aligned}$$

۳. تکرار مراحل تا رسیدن به همگرایی (با معیارهایی مانند آماره گلمن-روبین).

با این حال، وقتی میانگین مدل به صورت غیرخطی و با چارچوب میدان عصبی بیزی تعریف شود، چگالی‌های شرطی کامل دیگر فرم بسته ندارند و هر گام از MCMC مستلزم محاسبات پرهزینه در شبکه عصبی است. این امر موجب می‌شود اجرای MCMC برای مدل ترکیبی BNF-FCSN در عمل بسیار کند و زمان‌بر باشد. به همین دلیل، اگرچه MCMC مبنای برآورد بیزی کلاسیک است و در مدل‌های ساده‌تر کارایی دارد، اما برای مدل حاضر مقیاس‌پذیری لازم را ندارد. به همین دلیل در بخش بعدی، از رهیافت استنباط واریاسیونی (VI) استفاده می‌کنیم که بر پایه بهینه‌سازی کران پایین شواهد (ELBO) عمل می‌کند و امکان برآورد کارآمد پارامترها در مدل‌های پیچیده و داده‌های بزرگ فضایی-زمانی را فراهم می‌سازد (زمیت و همکاران، ۲۰۲۴؛ مک‌درمات و ویکلی، ۲۰۱۹).

۵ برآورد پارامترها با استفاده از میدان عصبی بیزی و استنباط واریاسیونی

در این بخش برآورد بیزی برای مدل ترکیبی ارائه می‌شود؛ مدلی که در آن میانگین میدان فضایی-زمانی توسط میدان عصبی بیزی (BNF) بازنمایی می‌شود و مولفه تصادفی با توزیع FCSN مدل شده است. هدف برآورد توزیع پسین پارامترهای $\theta = (\theta_{\text{BNF}}, \sigma, \lambda, \varphi_s, \nu_s, \varphi_t, \nu_t)$ است؛ به طوری که θ_{BNF} مجموعه وزن‌ها و اربیبی‌های شبکه عصبی بیزی را دربر می‌گیرد. برای غلبه بر مشکل MCMC از VI استفاده می‌کنیم. ایده اصلی آن است که به جای پسین واقعی، یک توزیع ساده‌تر $q(\theta; \psi)$ (مثلاً گاوسی چندمتغیره با کوواریانس کم‌رتبه یا قطری) در نظر گرفته می‌شود و پارامترهای آزاد ψ به گونه‌ای یاد گرفته می‌شوند که $q(\theta; \psi)$ تا حد ممکن به پسین واقعی نزدیک باشد. این کار با بیشینه‌سازی کران پایین شواهد

$$\text{ELBO}(\psi) = \mathbb{E}_{q(\theta; \psi)}[\log p(\mathbf{Y}_n | \theta)] - \text{KL}(q(\theta; \psi) \| p(\theta)),$$

انجام می‌شود. در چارچوب این مطالعه، توزیع تقریبی $q(\theta; \psi)$ به صورت ترکیبی تعریف شده است. برای وزن‌ها و اربیبی‌های شبکه عصبی از تقریب میانگین میدان^۱ استفاده شده است که در آن کوواریانس‌ها به صورت قطری در نظر گرفته می‌شوند. در این حالت، عبارت $\text{KL}(q\|p)$ برای این بخش به صورت جمعی نوشته می‌شود. با این حال، پارامترهای فضایی-زمانی با یک توزیع گاوسی چندمتغیره کم‌رتبه تقریب زده شده‌اند که وابستگی‌های درونی را حفظ می‌کند؛ بنابراین برای این بخش، عبارت KL به صورت کامل و متناسب با ساختار کم‌رتبه محاسبه می‌شود. به این ترتیب نوشتن KL به صورت جمعی تنها برای جزء میانگین میدان معتبر است و پیاده‌سازی عملی KL برای کل مدل

¹Mean Field

مطابق با ساختار واقعی q انجام شده است. برای محاسبه گرادیان ELBO، از ترفند بازپارامتری سازی استفاده می شود. اگر $\theta_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ می توان نوشت $\theta_i = \mu_i + \sigma_i \epsilon_i, \epsilon_i \sim \mathcal{N}(0, 1)$ که این بازنویسی اجازه می دهد گرادیان های پایدار نسبت به μ_i, σ_i محاسبه شوند و واریانس کاهش یابد (بیگام و همکاران، ۲۰۱۹). در VI به جای تکرارهای سنگین، MCMC از بهینه سازی گرادیانی استفاده می کند؛ بنابراین مقیاس پذیری بالاتر دارد، برای داده های حجیم و نامتقارن فضایی-زمانی عملی است، برآورد عدم قطعیت معتبر ارائه می کند. برای تضمین یک پارچگی روش شناسی و امکان مقایسه مستقیم نتایج، معماری میدان عصبی بیزی به گونه ای طراحی شد که در هر دو مطالعه شبیه سازی و داده های واقعی بدون تغییر مورد استفاده قرار گیرد. در هر دو مطالعه شبیه سازی و مثال داده های واقعی، میانگین میدان فضایی-زمانی $m(s, t)$ توسط یک میدان عصبی بیزی مدل شده است. ورودی شبکه شامل پنج ویژگی عددی: طول جغرافیایی، عرض جغرافیایی، زمان نرمال شده و دو مؤلفه سینوسی و کسینوسی زمان است. معماری شبکه به صورت یک شبکه تمام متصل با دو لایه مخفی در نظر گرفته شده است؛ به طوری که تعداد نورون ها در دو لایه مخفی به ترتیب ۳۲ و ۱۶ است (ساختار ۱ → ۱۶ → ۳۲ → p). تابع فعال سازی در هر دو لایه مخفی، تابع $\tanh$ انتخاب شده است. وزن ها و اریبی های شبکه با پیشین گاوسی $\mathcal{N}(0, 1)$ مدل شده و در چارچوب استنباط واریاسیونی به همراه سایر پارامترهای مدل برآورد می شوند. این معماری در تمام آزمایش ها ثابت نگه داشته شده است تا امکان مقایسه و تکرارپذیری نتایج فراهم شود. فرآیند برآورد با VI برای مدل ترکیبی BNF-FCSN به صورت زیر پیاده سازی می شود:

الگوریتم ۱. استنباط واریاسیونی برای مدل BNF-FCSN

۱. مقداردهی اولیه:

- تعریف توزیع تقریبی $q(\theta; \psi^{(0)})$ با پارامترهای اولیه $(\mu_i^{(0)}, \sigma_i^{(0)})$.
- تنظیم نرخ یادگیری (α) و تعداد تکرارها (T) .

۲. برای $t = 1, \dots, T$:

- نمونه گیری $(\theta^{(t)}; \psi^{(t)}) \sim q(\theta; \psi^{(t)})$ با بازپارامتری سازی:

$$\theta_i^{(t)} = \mu_i^{(t)} + \sigma_i^{(t)} \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, 1).$$

- محاسبه لگاریتم درست نمایی با استفاده از شبکه BNF (برای میانگین m_n) و چگالی FCSN (برای مولفه تصادفی): $\log p(\mathbf{Y}_n | \theta^{(t)})$.
- محاسبه واگرایی کولبک-لیبلر بین توزیع تقریبی و پیشین ها:

$$KL(q(\theta; \psi^{(t)}) \| p(\theta)) = \sum_{i=1}^k KL(\mathcal{N}(\mu_i^{(t)}, \sigma_i^{(t)2}) \| p(\theta_i)).$$

• محاسبه گرادیان کران پایین شواهد :

$$\nabla_{\psi} ELBO = \nabla_{\psi} \{ \mathbb{E}_{q(\theta; \psi)} [\log p(\mathbf{Y}_n | \theta)] - KL(q(\theta; \psi) \| p(\theta)) \},$$

و به‌روزرسانی پارامترها با الگوریتم آدام: $\psi^{(t+1)} \leftarrow \psi^{(t)} + \alpha \nabla_{\psi} ELBO$

۳. خروجی: توزیع تقریبی $q(\theta; \psi^*)$ و برآورد میانگین μ_i برای هر پارامتر.

تمام محاسبات عددی و پیاده‌سازی مدل در محیط پایتون انجام شده است. اجرای مدل‌ها روی یک رایانه با مشخصات پردازنده 2.7 GHz Intel Core i7، حافظه اصلی 8 GB 1333 MHz DDR3 RAM، سیستم‌عامل macOS Sierra 10.12.6، کارت گرافیک Intel HD Graphics 3000 (512 MB)، محیط اجرا Python به همراه کتابخانه‌های PyTorch و Pyro انجام شد. در ادامه، به‌منظور ارزیابی عملی چارچوب پیشنهادی، یک مطالعه شبیه‌سازی و یک مثال داده‌های واقعی مورد بررسی قرار می‌گیرند و مدل پیشنهادی به‌همراه رهیافت استنباط واریاسیونی در هر دو مطالعه پیاده‌سازی می‌شود.

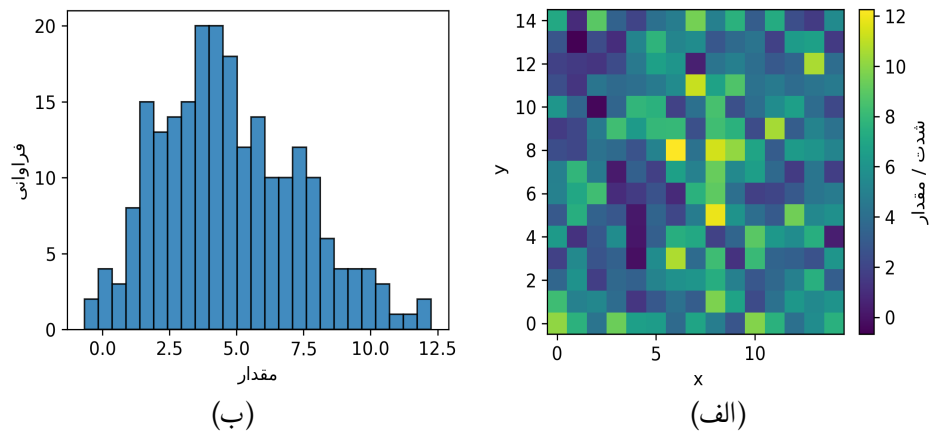
۶ مطالعه شبیه‌سازی

به‌منظور ارزیابی عملکرد چارچوب پیشنهادی، یک مطالعه شبیه‌سازی طراحی گردید. داده‌ها بر روی یک شبکه فضایی منظم 15×15 شامل ۲۲۵ موقعیت مکانی و طی $T = 8$ لحظه زمانی در بازه $[1, 8]$ تولید شدند، به‌طوری که در مجموع $n = 225 \times 8 = 1800$ مشاهده ایجاد گردید. ماتریس متغیرهای توضیحی به‌صورت $\mathbf{X} = [\mathbf{x}_n, \mathbf{y}]$ شامل یک ستون ثابت (عرض از مبدأ) و مختصات نرمال‌شده فضایی $(\mathbf{x}, \mathbf{y})$ تعریف شد. برای بررسی عملکرد مدل‌ها، دو سناریوی پارامتری در نظر گرفته شد: چولگی متوسط $\varphi_t = 4, \varphi_s = 1, \sigma = 1, \beta = (1, 2, -1)'$ و سناریوی دوم با چولگی بالا $\nu_t = 1, \nu_s = 1.5, \lambda = 2$ و $\nu_t = 1, \nu_s = 2, \lambda = 4$. شکل ۳ نمونه‌ای از داده‌های شبیه‌سازی شده در سناریوی دوم را نمایش می‌دهد، (الف) نقشه فضایی پاسخ‌ها در لحظه زمانی $t = 1$ و در (ب) توزیع حاشیه‌ای همان مقادیر نشان داده شده است که وجود چولگی مثبت و دم راست سنگین را آشکار می‌سازد. برای بررسی رفتار چوله و انحراف از گاوسی بودن، در هر سناریو مقادیر پاسخ در زمان ثابت $t = 1$ از میان ۲۰ تکرار مستقل شبیه‌سازی و ۲۲۵ موقعیت مکانی تجمیع شده و یک نمونه یک‌بعدی با اندازه ۴۵۰۰ تشکیل گردید. برای این نمونه، ضریب چولگی برابر با ۰.۵۶۰ به‌دست آمد که نشان‌دهنده چولگی مثبت و وجود دم راست سنگین است.

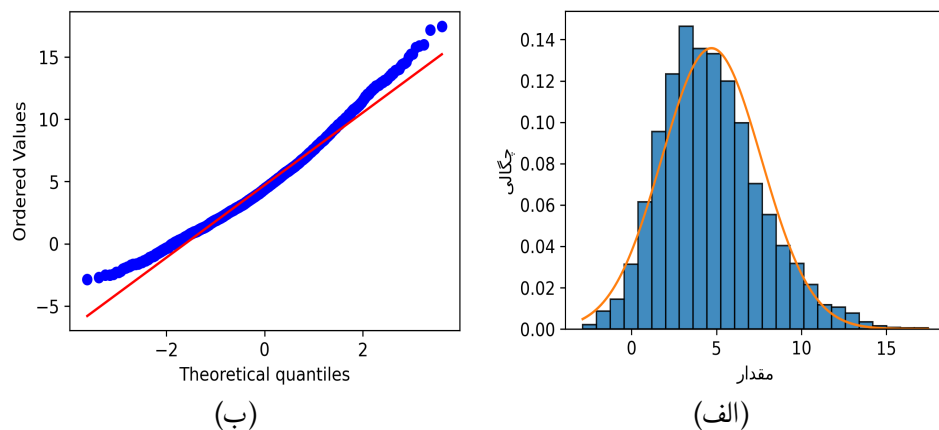
آزمون شاپیرو-ویلک ($SW = 0.982, p = 1.576 \times 10^{-23}$) و آزمون کلموگروف-اسمیرنوف $KS = 0.045, p = 2.667 \times 10^{-8}$) در برابر گاوسی انجام شد که با توجه به نتایج هر دو در سطح پنج درصد فرض گاوسی بودن حاشیه‌ای را رد کردند. این نتایج به همراه نمودار QQ شکل (۲) تأیید می‌کنند که داده‌های شبیه‌سازی شده دارای

۹۲ مدل‌سازی بیزی داده‌های فضایی-زمانی با میدان عصبی بیزی

رفتار چوله و انحراف معنادار از توزیع گاوسی هستند و استفاده از توزیع FCSN برای مولفه تصادفی مدل از نظر آماری کاملاً موجه است.



شکل ۱. نمونه‌ای از داده‌های شبیه‌سازی‌شده در سناریوی چولگی بالا الف- نقشه فضایی پاسخ‌ها در لحظه زمانی $t = 1$ ، ب- توزیع حاشیه‌ای مقادیر.



شکل ۲. الف- هیستوگرام، ب- نمودار چندک-چندک برای نمونه ۴۵۰۰ تایی داده‌های شبیه‌سازی‌شده در زمان $t = 1$ ، نشان‌دهنده چولگی مثبت و انحراف واضح از گاوسی بودن.

برای برآورد بیزی پارامترها دو رویکرد متفاوت مورد مقایسه قرار گرفتند. در رویکرد نخست، مدل فضایی-زمانی با فرض خطای چوله گاوسی بسته منعطف و ساختار همبستگی مترن، با استفاده از الگوریتم‌های MCMC برازش داده شد. در رویکرد دوم، مدل پیشنهادی با بهره‌گیری از BNF و توزیع FCSN پیاده‌سازی شد. در این چارچوب، میانگین

میدان مستقیماً توسط BNF مدل می‌شود و اجزای نامتقارن و ناهمگن از طریق FCSN مدل‌سازی می‌گردند. برای برآورد پارامترها از استنباط واریاسیونی (VI) با بهینه‌سازی ELBO استفاده شد که امکان مقیاس‌پذیری محاسباتی را فراهم می‌کند. طرح شبیه‌سازی طی ۲۰ بار تکرار مستقل اجرا شد و عملکرد دو رویکرد با استفاده از معیارهای میانگین مربعات خطا (MSE)، پوشش فواصل باورمندی ۹۵٪ (CP) و زمان محاسباتی (بر حسب ثانیه) مقایسه گردید. نتایج در جدول ۱ گزارش شده است. نتایج نشان داد که مدل پایه در هر دو سناریو خطاهای بالاتری داشت و پوشش

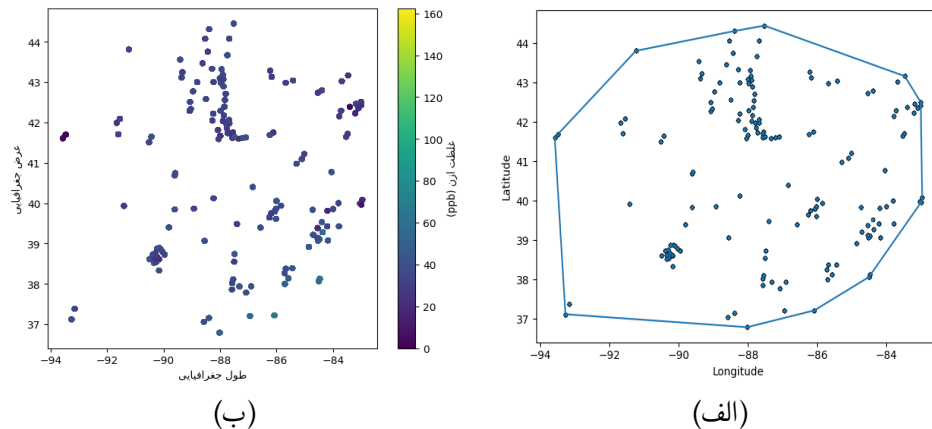
جدول ۱. نتایج شبیه‌سازی در دو حالت چولگی متوسط و بالا.

سناریو	مدل	متوسط RMSE	متوسط بازه باورمندی	متوسط زمان (s)
۱	BNF-FCSN+VI	۰/۲۹	۹۶/۸۱	۹۰/۶۴
۱	FCSN+MCMC	۰/۷۲	۹۵/۴۰	۱۴۷/۵۰
۲	BNF-FCSN+VI	۰/۶۸	۹۶/۷۳	۱۰۰/۶۳
۲	FCSN+MCMC	۰/۹۵	۹۵/۲۴	۱۵۶/۸۲

فواصل باورمندی آن کمتر بود. در مقابل، مدل پیشنهادی خطای میانگین مربعات کمتری ارائه داد، پوشش فواصل باورمندی آن نزدیک به ۹۵٪ بود و زمان محاسباتی کمتر از MCMC برآورد شد. همان‌گونه که مشاهده می‌شود، هر دو روش پوششی نزدیک به سطح اسمی ۹۵٪ ارائه می‌دهند؛ به‌طوری‌که در سناریوی اول پوشش روش VI برابر با ۹۶/۸۱٪ و پوشش روش MCMC برابر با ۹۵/۴۰٪ است. این الگو در سناریوی دوم نیز تکرار می‌شود. بنابراین، نتایج نشان می‌دهند که رهیافت واریاسیونی، علاوه بر کاهش قابل‌توجه زمان اجرا، از نظر کمی‌سازی عدم‌قطعیت نیز عملکردی معتبر و قابل مقایسه با روش MCMC دارد و پوشش فاصله‌های باورمندی آن به سطح اسمی هدف نزدیک است. این نتایج بیان‌گر آن است که ترکیب میدان عصبی بیزی با توزیع چوله منقطع و استنباط واریاسیونی می‌تواند هم دقت پیش‌بینی و هم کارایی محاسباتی را به‌طور چشمگیری بهبود بخشد.

۷ مطالعه موردی با داده‌های واقعی ازن

به‌منظور ارزیابی عملی چارچوب پیشنهادی، مجموعه داده ozone2 از بسته fields در نرم‌افزار R مورد استفاده قرار گرفت. این داده شامل اندازه‌گیری‌های روزانه غلظت ازن سطح زمین (بر حسب ppb) در ایالت‌های شمال‌شرقی ایالات متحده طی یک دوره ۸۹ روزه در تابستان و در ۱۵۳ ایستگاه پایش جوی است. چنین ساختاری داده‌ای طولی-فضایی ایجاد می‌کند که هر مشاهده ترکیبی یکتا از مکان و زمان را نشان می‌دهد. موقعیت مکانی ایستگاه‌ها به‌صورت طول و عرض جغرافیایی در دسترس بوده و در نقشه منطقه مطالعه (شکل ۴) نمایش داده شده است. این نقشه گستردگی فضایی ایستگاه‌ها و ناهمگنی جغرافیایی آن‌ها را نشان می‌دهد و مبنایی برای تحلیل همبستگی‌های مکانی فراهم می‌سازد. برای آماده‌سازی داده‌ها، مشاهدات از حالت ماتریسی اولیه به قالب طولی تبدیل شدند تا هر ردیف داده معرف یک ترکیب یکتای روز-ایستگاه باشد. متغیر پاسخ، یعنی غلظت ازن، همراه با مجموعه‌ای از متغیرهای

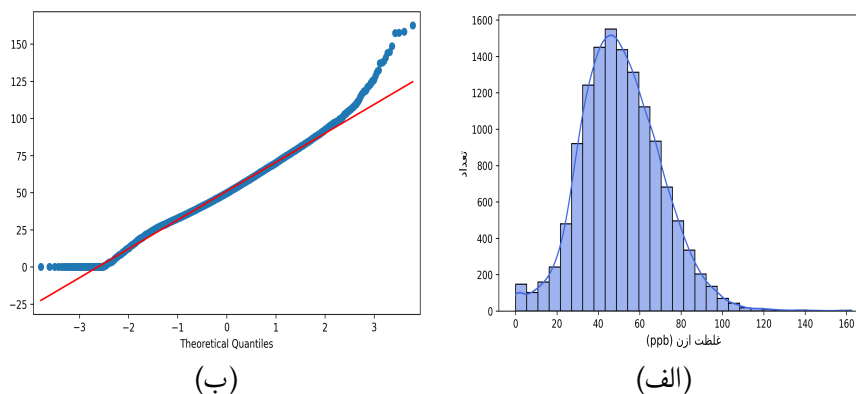


شکل ۳. تحلیل اکتشافی داده‌های واقعی ازن: الف- موقعیت مکانی ۱۵۳ ایستگاه پایش ازن، ب- نقشه غلظت میانگین ازن طی ۸۹ روز.

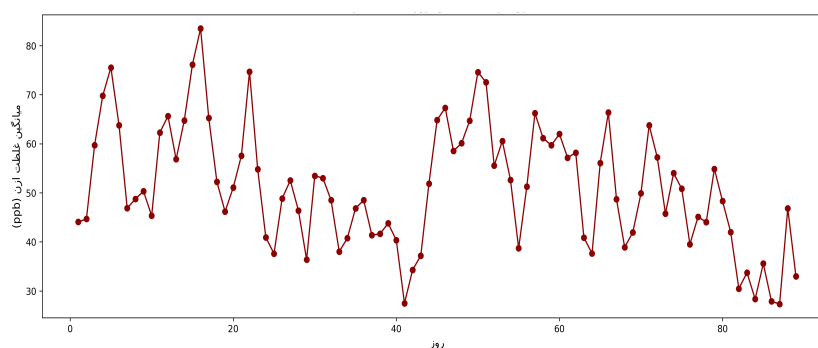
کمکی به مدل افزوده شد. مختصات جغرافیایی ایستگاه‌ها به‌عنوان ویژگی‌های فضایی، شاخص زمان نرمال‌شده در بازه $[0, 1]$ برای نمایش روند طولی و دو مؤلفه سینوسی و کسینوسی $\sin(2\pi t)$ و $\cos(2\pi t)$ برای ثبت الگوهای فصلی مورد استفاده قرار گرفتند. همچنین، یک مؤلفه تصادفی در سطح ایستگاه برای لحاظ‌کردن ناهمگنی ثابت میان مکان‌ها تعریف شد. نمودار ۴ (الف) توزیع کلی غلظت ازن را نمایش می‌دهد. همان‌طور که دیده می‌شود، داده‌ها دارای دم راست سنگین هستند و از توزیع گاوسی فاصله دارند. این ویژگی نشان‌دهنده وجود چولگی مثبت و احتمال وقوع مقادیر بسیار بزرگ در مقایسه با توزیع گاوسی است. در نمودار ۴ (ب) نیز این موضوع تأیید می‌شود. نقاط به‌ویژه در بخش‌های دم راست به‌طور مشخص از خط قطری فاصله گرفته‌اند که نشان‌دهنده انحراف از گاوسی است. این شواهد لزوم استفاده از توزیع‌های چوله مانند FCSN را برای مدل‌سازی مناسب‌تر داده‌ها تقویت می‌کنند. نمودار ۵ روند روزانه نشان می‌دهد که نوسانات زمانی و دوره‌های اوج آلودگی نیز بخش مهمی از ساختار فضایی-زمانی داده‌ها هستند. ترکیب این سه نمودار اکتشافی تصویری روشن از ماهیت نامتقارن، نایستا و پرنوسان داده‌های ازن ارائه می‌دهد. بر همین اساس، دو مدل برازش داده شد، مدل اول با فرض مؤلفه تصادفی گاوسی کلاسیک و مدل دوم با فرض مؤلفه تصادفی FCSN که پارامتر λ را برای چولگی در نظر می‌گیرد. هر دو مدل در چارچوب میدان عصبی بیزی و با بهره‌گیری از استنباط واریاسیونی پیاده‌سازی شدند. جدول ۱ نتایج حاصل از برازش دو مدل روی داده واقعی ozone2 را نشان می‌دهد.

همان‌طور که ملاحظه می‌شود، مدل پیشنهادی BNF-VI-FCSN در تمام معیارها عملکرد بهتری نسبت به مدل گاوسی (BNF-VI-N) دارد. مقادیر خطای پیش‌بینی، ریشه میانگین مربعات خطا^۱ (RMSE) و میانگین

¹Root Mean Squared Error



شکل ۴. تحلیل اکتشافی داده‌های واقعی ازن: الف- توزیع حاشیه‌ای و چگالی هموارشده، ب- انحراف از نرمالیت حاشیه‌ای در QQ



شکل ۵. روند روزانه و نوسانات زمانی در بازه ۸۹ روز.

قدر مطلق خطا^۱ (MAE) در مدل FCSN تقریباً نصف مقادیر متناظر در مدل گاوسی است ($RMSE = ۱۵/۲۶$) در برابر $۳۰/۷۹$ و $MAE = ۱۱/۸۶$ در برابر $۲۶/۴۵$) که بیانگر دقت بالاتر در پیش‌بینی است. از سوی دیگر، معیار CRPS نیز کاهش چشمگیری داشته ($۸/۴۸$ در مقابل $۱۹/۶۹$) که نشان‌دهنده بهبود کیفیت توزیع پیش‌بینی است. یکی از شاخص‌های کلیدی در مقایسه مدل‌ها، معیار CRPS^۲ است. این شاخص فاصله بین توزیع پیش‌بینی مدل و مقدار واقعی مشاهده‌شده را در نظر می‌گیرد و ترکیبی از دقت (نزدیک بودن پیش‌بینی‌ها به مشاهدات) و کالیبراسیون (درستی کمی‌سازی عدم قطعیت) را می‌سنجد. مقدار CRPS هرچه کوچک‌تر باشد، نشان‌دهنده پیش‌بینی دقیق‌تر و منسجم‌تر است. در واقع، CRPS تعمیم‌یافته MAE برای پیش‌بینی‌های احتمالاتی محسوب می‌شود که نه تنها میانگین، بلکه کل توزیع پیش‌بینی را ارزیابی می‌کند. در این مطالعه، کاهش قابل‌توجه CRPS در مدل FCSN

^۱Mean Absolute Error

^۲Continuous Ranked Probability Score

جدول ۲. مقایسه عملکرد مدل روی داده واقعی ازن.

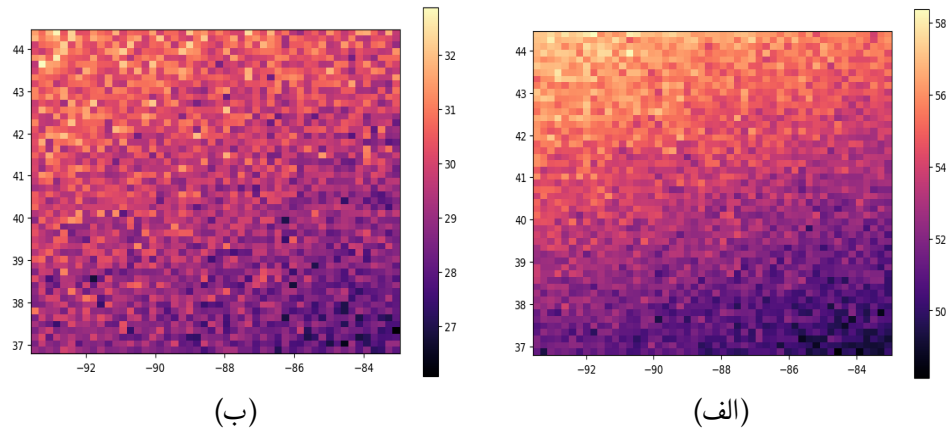
مدل	RMSE	MAE	CRPS	CP	PI	زمان
BNF-VI-N	۳۰/۷۹۱	۲۶/۴۴۷	۱۹/۶۹۱	۶۴/۵۵۶	۶۱/۴۶۶	۲۷۷/۷۹۲
BNF-VI-FCSN	۱۵/۲۵۶	۱۱/۸۵۸	۸/۴۸۴	۹۶/۲۰۵	۶۲/۶۵۰	۳۰۲/۲۶۴

نسبت به مدل گاوسی بیانگر آن است که چارچوب پیشنهادی توانسته است هم از نظر دقت نقطه‌ای و هم از نظر ارزیابی عدم قطعیت، عملکردی برتر ارائه دهد. مهم‌تر از همه، پوشش بازه‌های باورمندی ۹۵ درصد (CP) در مدل FCSN برابر با ۹۶٪ است، در حالی که مدل گاوسی تنها پوشش ۶۴٪ داشته است. این موضوع نشان می‌دهد که مدل FCSN عدم قطعیت را به صورت معتبرتر و دقیق‌تر کمی‌سازی می‌کند. به عبارت دیگر مقایسه بازه‌های باورمندی نشان می‌دهد که مدل گاوسی، عدم قطعیت را به طور قابل توجهی کم‌برآورد می‌کند؛ این رفتار ناشی از ناتوانی مدل گاوسی در بازنمایی دُم‌های سنگین و چولگی ساختاری داده‌های ازن است. در مقابل، مدل FCSN با وارد کردن مؤلفه چولگی قادر است رفتار نامتقارن و دم راست سنگین داده‌ها را به طور صریح مدل کند. این ویژگی موجب تولید برآوردهای عدم قطعیت واقع‌بینانه‌تر شده و در نتیجه، پوشش بازه ۹۵٪ به مقدار اسمی خود همگرا می‌شود. این تفاوت چشمگیر نشان می‌دهد که مدل‌سازی چولگی، نقش مستقیم و تعیین‌کننده‌ای در ارزیابی صحیح عدم قطعیت پیش‌بینی برای داده‌های ازن دارد.

در نهایت، میانگین پهنای بازه پیش‌بینی^۱ (PI) در دو مدل تقریباً مشابه است (۶۲/۶۵ در مقابل ۶۱/۴۷)، به این معنا که بهبود پوشش در مدل FCSN بدون افزایش هزینه پهنای بازه به دست آمده است. زمان اجرای دو مدل نیز اختلاف زیادی نداشته و افزایش اندک زمان در مدل FCSN (حدود ۲۵ ثانیه) در برابر بهبود قابل توجه در معیارهای دقت و عدم قطعیت کاملاً قابل توجیه است. به طور کلی، نتایج جدول ۱ نشان می‌دهند که ترکیب چارچوب میدان عصبی بیزی با توزیع چوله گاوسی بسته منعطف ($BNF-VI-FCSN$) برای داده‌های واقعی فضایی-زمانی نامتقارن، مدلی قدرتمند و معتبر محسوب می‌شود. به منظور ارزیابی رهیافت MCMC از الگوریتم مونت کارلویی همیلتونی استفاده شد و مدل گاوسی فضایی-زمانی با میانگین خطی و کوواریانس نمایی تفکیک‌پذیر بر روی یک زیرنمونه $n = ۲۰۰۰$ از داده واقعی ozone2 اجرا شد. نتایج نشان داد که برآورد MCMC به شدت دچار کم‌برآوردی عدم قطعیت است. پوشش صفر درصد و پهنای بازه تقریباً نقطه‌ای که نشان می‌دهد الگوریتم‌های MCMC در این ابعاد داده و با ساختار کوواریانس تفکیک‌پذیر قادر به تولید پسین معنادار نیست و در عمل دچار فروپاشی عدم قطعیت می‌شود. این رفتار برای مدل FCSN خطی نیز (با شدت بیشتر) مشاهده شد.

بر اساس شکل ۶، الگوی عدم قطعیت در دو مدل تفاوت معناداری دارد. در مدل $BNF-Normal$ (ب)، عدم قطعیت تقریباً به صورت همگن در سراسر منطقه توزیع شده و الگوهای مکانی بارزی مشاهده نمی‌شود. این موضوع نشان می‌دهد که مدل گاوسی قادر به بازنمایی ویژگی‌های ناهمگن داده‌های واقعی نیست. در مقابل، مدل

¹Prediction Interval Width



شکل ۶. نقشه عدم قطعیت پیش‌بینی غلظت ازن در منطقه مطالعه: الف- مدل پیشنهادی BNF-FCSN ب- مدل BNF- Normal

BNF-FCSN (الف) علاوه بر کاهش سطح کلی عدم قطعیت، الگوهای مکانی متفاوتی را نیز ثبت کرده است؛ به‌ویژه در نواحی شمالی و شرقی، میزان عدم قطعیت بالاتر بوده و این موضوع با تراکم ایستگاه‌های کمتر و تغییرات فضایی بیشتر همخوانی دارد. این یافته‌ها نشان می‌دهند که لحاظ کردن چولگی در چارچوب پیشنهادی، توانایی مدل را در کمی‌سازی معتبرتر عدم قطعیت و شناسایی نواحی پر ریسک افزایش می‌دهد. کدهای پیاده‌سازی مدل پیشنهادی در مخزن [GitHub](#) در دسترس هستند.

بحث و نتیجه‌گیری

هدف اصلی این پژوهش، توسعه و ارزیابی چارچوبی نوین با عنوان میدان عصبی بیزی برای تحلیل داده‌های فضایی-زمانی نامتقارن بود. در این چارچوب، میدان پیوسته فضا-زمان به‌طور مستقیم مدل‌سازی می‌شود و با ترکیب آن با توزیع چوله گاوسی بسته منعطف امکان بازنمایی هم‌زمان ناهمگنی، چولگی و ساختارهای همبستگی پیچیده فراهم گردید. بهره‌گیری از استنباط واریاسیونی نیز سبب شد که این چارچوب علاوه بر دقت بالا، از نظر محاسباتی مقیاس‌پذیر و کارآمد باشد. نتایج مطالعه شبیه‌سازی نشان داد که رهیافت پیشنهادی نسبت به رهیافت کلاسیک مثل MCMC برتری محسوسی دارد. میانگین مربعات خطا برای تمام پارامترها در مدل پیشنهادی کمتر بود و پوشش فواصل باورمندی ۹۵٪ به سطح اسمی نزدیک‌تر قرار گرفت. این یافته‌ها بیانگر آن است که چارچوب BNF می‌تواند به‌طور انعطاف‌پذیر روابط غیرخطی و ساختارهای نامتقارن را در داده‌های فضایی-زمانی بازنمایی کند، در حالی که مدل‌های سنتی مبتنی بر فرض گاوسی چنین دقتی را ارائه نمی‌دهند. علاوه بر این، تحلیل داده واقعی ازن در ایالت‌های شمال شرقی آمریکا کارایی عملی مدل پیشنهادی را تأیید کرد. مقایسه دو مدل نشان داد که BNF-FCSN نسبت به

مدل گاوسی دارای RMSE و MAE کوچکتر و پوشش فواصل باورمندی نزدیک‌تر به سطح ۹۵٪ است. همچنین، نقشه‌های پیش‌بینی و عدم قطعیت آشکار ساختند که مدل FCSN توانسته الگوهای مکانی-زمانی و نواحی پرخطر با غلظت بالای ازن را دقیق‌تر شناسایی کند، در حالی که مدل گاوسی در ثبت دُم‌های سنگین و عدم تقارن داده‌ها ناکام ماند. این نتایج نشان داد که استفاده از FCSN در چارچوب BNF نه تنها برتری محاسباتی، بلکه مزیت‌های تفسیری و کاربردی نیز دارد. از منظر محاسباتی، روش VI در چارچوب BNF زمان اجرانسیبت به MCMC کاهش داد. این مزیت در داده‌های بزرگ مقیاس فضایی-زمانی بسیار مهم است، زیرا الگوریتم‌های MCMC اغلب با کندی همگرایی یا هزینه محاسباتی سنگین مواجه می‌شوند. ویژگی کلیدی BNF آن است که بدون نیاز به تعریف صریح توابع همبستگی، میانگین میدان و ساختارهای نامتقارن را به‌طور مستقیم از داده می‌آموزد و بدین ترتیب انعطاف و قدرت پیش‌بینی بالاتری ایجاد می‌کند. پژوهش حاضر محدودیت‌هایی نیز دارد. نخست آن که هرچند میدان عصبی بیزی در این مطالعه قادر است ساختار میانگین را به‌صورت کاملاً منعطف و غیرایستا مدل‌سازی کند، اما توزیع چوله مورد استفاده، هرچند منعطف است، اما همچنان بر پایه فرض‌های خاصی درباره ایستایی و ساختار همبستگی بنا شده و ممکن است در داده‌های غیرایستا دقت کمتری داشته باشد. این محدودیت مدل در مواجهه با نالیستایی به بخش کوواریانس بازمی‌گردد. در این چارچوب، تابع همبستگی از نوع ماترن تفکیک‌پذیر در فضا - زمان در نظر گرفته شده است، تابعی که ذاتاً ایستا بوده و امکان تغییرپذیری موضعی، مقیاس‌های همبستگی متفاوت در نواحی مختلف، یا ناهمگنی‌های شدید را به‌طور کامل منعکس نمی‌کند. دوم آنکه کیفیت روش واریاسیونی به انتخاب توزیع تقریبی بستگی دارد و در برخی شرایط ممکن است نسبت به MCMC سوگیری نشان دهد.

برای تحقیقات آینده مدل‌های پیشرفته‌تر در چارچوب شبکه‌های عصبی بیزی نظیر مدل‌های بازگشتی، می‌تواند توانایی مدل در ثبت پویایی‌های پیچیده فضایی-زمانی را افزایش دهد. طراحی توزیع‌های تقریبی منعطف‌تر در استنباط واریاسیونی، همچون گاوسی چندلایه یا جریان‌های نرمال‌سازی‌شده، می‌تواند کیفیت تقریب پسین را ارتقا دهد. به‌رحال تمرکز اصلی این پژوهش بر مدل‌سازی عدم قطعیت در داده‌های فضایی-زمانی چوله با استفاده از چارچوب BNF-FCSN-VI بوده است و هدف مقایسه جامع بین تمامی روش‌های موجود برآورد پارامترها نبوده است. روش‌های مبتنی بر درست‌نمایی، مانند MLE یا سایر روش‌های درست‌نمایی مثل رهیافت درست‌نمایی مرکب برای توزیع‌های چوله نظیر FCSN، عمدتاً در تنظیمات ساده‌تر (میانگین خطی و کوواریانس ایستا) توسعه یافته‌اند. تعمیم این روش‌ها به مدل‌های غیرخطی با میانگین میدان عصبی بیزی و داده‌های بزرگ فضایی-زمانی مستلزم کار عددی و نظری قابل توجهی است. بررسی و مقایسه دقیق چنین روش‌هایی با چارچوب BNF-FCSN-VI می‌تواند به عنوان یک موضوع مهم برای پژوهش‌های آینده دنبال شود.

تقدیر و تشکر

نویسندگان مقاله از پیشنهادات و نظرات ارزنده داوران محترم، ویرایش ادبی سردبیر، هیئت تحریریه و ویراستار مجله که در بهبود کیفیت مقاله بسیار مؤثر واقع شد، تقدیر و تشکر می‌نمایند.

مراجع

- کریمی، الف. و حسینی ف.، (۱۴۰۲)، میدان تصادفی چوله نرمال بسته منعطف برای تحلیل داده‌های فضایی چوله، مجله علوم آماری، ۱۷ (۲)، ۳۷۱-۳۸۸.
- کریمی، الف. و حسینی، ف. (۱۴۰۰)، معرفی یک میدان تصادفی مانای چوله گاوسی، مجله علوم آماری، مجله علوم آماری، ۱۵ (۲)، ۵۴۹-۵۶۶.
- کریمی، الف. و حسینی ف.، (۱۴۰۳)، تحلیل بیز مقداری مدل رگرسیون فضایی چوله بر اساس زیر رده منعطفی از توزیع چوله نرمال بسته، مجله علوم آماری، ۱۸ (۲)، ۱۲۷-۱۴۲.
- حسینی، ف.، کریمی، ا. (۱۴۰۳). تحلیل بیزی متغیرهای پنهان در مدل‌های آمیخته خطی تعمیم‌یافته فضایی با میدان تصادفی مانای چوله گاوسی. مجله علوم آماری، ۱۸ (۱)، ۵۷-۷۲.
- Allard, D. and Naveau, P. (2007), A New Spatial Skew-Normal Random Field Model, *Communications in Statistics - Theory and Methods*, **36**, 1821-1834.
- Arellano-Valle, R. B., Ferreira, C. S., and Genton, M. G., (2018). Scale and Shape Mixtures of Multivariate Skew-Normal Distributions *Journal of Multivariate Analysis*, **166**, 98-110.
- Azzalini, A. and Capitanio, A. (1999), Statistical Applications of the Multivariate Skew Normal Distribution, *Journal of the Royal Statistical Society*, **61**(B), 579-602.
- Bevilacqua, M., Caamaño-Carrillo, C., Arellano-Valle, R. B., and Morales-Oñate, V. (2020), Nonstationary and non-Gaussian Modeling of Spatial Data using Flexible Skew-t Processes. *Scandinavian Journal of Statistics*, **48**, 212-245.
- Betancourt, M. (2017), A Conceptual Introduction to Hamiltonian Monte Carlo. *arXiv preprint*, arXiv:1701.02434.
- Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., Singh, R., Szerlip, P. A., Horsfall, P., and Goodman, N. D. (2019), Pyro: Deep Universal Probabilistic Programming. *Journal of Machine Learning Research*, **20**(28), 1-6.

- Blei, D. M., Kucukelbir, A., and McAuliffe, J. D. (2017), Variational Inference: A Review for Statisticians. *Journal of the American Statistical Association*, **112**(518), 859–877.
- Blundell, C., Cornebise, J., Kavukcuoglu, K., and Wierstra, D. (2015), Weight Uncertainty in Neural Networks. *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, 1613–1622.
- Cressie, N. and Wikle, C. K. (2011), *Statistics for Spatio-Temporal Data*. John Wiley and Sons.
- Fortunato, M., Blundell, C., and Vinyals, O. (2017), Bayesian Recurrent Neural Networks. *arXiv preprint* DOI:10.48550/arXiv.1704.02798
- Genton, M. G. (2004), *Skew-Elliptical Distributions and Their Applications: A Journey beyond Normality*, Chapman and Hall/CRC,
- Gómez-Rubio, V. (2020). *Bayesian Inference with INLA (1st ed)*. Chapman and Hall/CRC.
- Hoffman, M. D. and Gelman, A. (2014), No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, **15**(1), 1351-1381.
- Hosseini, F. and Karimi, O., (2021), Approximate Pairwise Likelihood Inference in SGLM Models with Skew Normal Latent Variables, *Journal of Computational and Applied Mathematics*, **398**, 113692.
- Jospin, L. V., Laga, H., Boussaid, F., and Buntine, W. (2022), Hands-on Bayesian Neural Networks — A Tutorial For Deep Learning Users, *IEEE Computational Intelligence Magazine*, **17**(2), 29–48.
- Márquez-Urbina, O. U. and González-Farías, G. (2022), A Flexible Special Case of the CSN for Spatial Modeling and Prediction. *Spatial Statistics*, **47**, 100556.

- McDermott, P.L., and Wikle, C.K. (2019), Bayesian Recurrent Neural Network Models for Forecasting and Quantifying Uncertainty in Spatial-Temporal Data. *Entropy*, **21**(2), 184.
- Mozdzen, A., Cremaschi, A., and Cadonna, A. Guglielmi, A., and Kastner, G. (2022), Bayesian Modeling and Clustering for Spatio-temporal Areal Data: An Application to Italian Unemployment. *Spatial Statistics*, **52**, 100715.
- Neal, R. M. (2011), MCMC using Hamiltonian dynamics. In: *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC.
- Ravishanker, N., Raman, B., and Soyer, R., (2022), *Dynamic Time Series Models using R-INLA: An Applied Perspective*, Chapman and Hall/CRC, 201–245.
- Rimstad, K. and Omre, H. (2014), Skew-Gaussian Random Fields, *Spatial Statistics*, **10**, 43-62.
- Saad, F., Burnim, J., Carroll, C., Patton, B., Koster, U., A. Saurous, R., and Hoffman, M. (2024), Scalable Spatiotemporal Prediction with Bayesian Neural Fields. *Nature Communications*, **15**, 1234.
- Sahu, S. K. (2022), *Bayesian Modeling of Spatio-Temporal Data with R*, Chapman and Hall/CRC.
- Zammit-Mangion, A., Kaminski, M. D., Tran, B.-H., Filippone, M., and Cressie, N. (2024), Spatial Bayesian Neural Networks, *Spatial Statistics*, **60**, 100825.