Review and Reproduction of

A Multivariate Spatial Skew-t Process for Joint Modeling of Extreme Precipitation Indexes

Course Project for **MTH643A: Spatial Statistics**

Submitted by: Arijit Dey (221281), Pratyusha Sarkar (221373), Ringan Majumdar (221390), and Swarnajit Podder (221453)

Supervisor: Dr. Arnab Hazra

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Date of Submission:

November 14, 2023

Academic Year 2023–24

Acknowledgement

We want to express our heartfelt thanks to our project guide, Dr. Arnab Hazra. His guidance, support, consistent efforts, and encouragement were instrumental in helping us tackle the challenges we encountered during this project. We are also thankful for his assistance in enhancing our understanding of the subject, in and out of the classroom.

–Team 3 (Arijit, Pratyusha, Ringan, Swarnajit)

Declaration

We hereby declare that the work presented in the project report entitled "Review and Reproduction of **A Multivariate Spatial Skew-t Process for Joint Modeling of Extreme Precipitation Indexes**" contains our own efforts. The codes and the plots were reproduced independently. The project's documentation includes a language unique to our team.

Abstract

To study trends in extreme precipitation across India over the years 1901-2022, 10 climate indices have been examined that serve as indicators of extreme precipitation. Some of them are like : annual maximum of consecutive 5-day average (P), Annual sum of P when P > 95th percentile, Maximum annual number of consecutive wet days etc. These indices not only vary over different spatial locations but also are dependent among themselves. To study the behaviour of the indices, consider a gridded dataset on daily rainfall provided by Indian Meteorological Department (IMD), Pune (https://www.imdpune.gov.in/cmpg/Griddata/Rainfall_1_NetCDF.html) is considered. In this paper, a multivariate spatial skew-t process is proposed to model the extreme precipitation indices jointly and its theoretical properties are being discussed also. In a Bayesian framework based study, it is found that the proposed model is performing quite better than other commonly used processes like multivariate spatial Gaussian Process, multivariate spatial t-processes. We use the recommended model to calculate the average changes in extreme rainfall indicators every 25 years across India and discover significant local variations.

Contents

1 Introduction

Extreme precipitation stands out as a significant climate factor ((IPCC et al., 2007)), and research works into its long-term variations in occurrence, strength, and duration are essential for sustainable development. To evaluate the impacts of climate change, the World Meteorological Organization (WMO) Commission for Climatology (CCl)/CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI) suggested a series of indices that describe climate extremes. Based on the daily observations of temperature and precipitation available at the National Climatic Data Center (NCDC)s Global Historical Climatology Network (GHCN) Daily data set, Donat et al. (2013) derive a suite of gridded data products called GHCNDEX (`http://www.climdex.org/gewocs.html`) that covers 27 climate indexes of which 10 explain extreme precipitation. To draw inferences about spatiotemporal trends, (Donat et al., 2013) calculate linear trends using the Sens trend estimator ((Sen, 1968)) separately for each grid location and use the Mann-Kendall test of significance (Kendall, 1955). The results show fewer significant changes in precipitation compared to the temperature indexes. However, this analysis is questionable, as it completely ignores spatial and mutual dependence. For a proper analysis, it is imperative to analyze the indexes jointly and account for the dependence exhibited by the data.

Multivariate spatial modeling generally relies on the assumption that the data is coming from a Gaussian process (GP) ((Gelfand et al., 2010)) because of the appealing theoretical characteristics of GPs, their straightforward implementation in high-dimensional spaces, and their flexibility in modeling. Nonetheless, Gaussian processes (GPs) face criticism in modeling spatial extremes, primarily due to the asymptotic independence that exists between any two spatial locations, except in the trivial scenario of exact dependence ((Davison et al., 2013)). In the case of a multivariate GP (MGP), asymptotic dependence across the components are similarly zero. Therefore, employing a geostatistical approach with a Multivariate Gaussian Process (MGP) is questionable when aiming to model multivariate spatial extremes in the presence of asymptotic dependence.

In this paper, we put forth a category of multivariate skew-t processes (MSTPs) through an extension of the univariate spatial skew-t process of (Padoan, 2011) and (Morris et al., 2017). A skew-t distribution is selected for its adaptability in modeling asymmetry and heavy-tailed data. We construct a spatial skew-t process considering separable covariance structure across the space and across the indexes ((Banerjee and Gelfand, 2002)) along with random mean and scale. We conduct a numerical comparison of the performance of MGP, the multivariate symmetric t process (MTP), MSTP, and their respective univariate counterparts in trend estimation.

Subsequently, we utilize MSTPs to make inferences regarding long-term trends in extreme precipitation indices.

The paper is organized as follows. In Section 2, we present the CLIMDEX/GHCNDEX data description and preliminary analysis. Section 3 outlines the modeling process using the proposed MSTPs. Section 4 delves into Bayesian computation, and in Section 5, we employ our approach to analyze the CLIMDEX indexes.

2 The Indices and Exploratory Data Analysis

2.1 Description of the Data and the Indices

Contributors: Pratyusha, Swarnajit

The CLIMDEX/GHCNDEX data repository (http://www.climdex.org/gewocs.html) includes 10 indices that represent extreme precipitation (Table 2.1).

Abbreviation	Description
Rx5day	Annual maximum of consecutive 5-day average P
R99p	Annual sum of P when P > 99th percentile
Rx1day	Annual maximum of P
R95p	Annual sum of P when P > 95th percentile
R95pT	Annual count of days when P > 95th percentile
SDII	Annual total P divided by the number of days with P $\geq$ 1 mm
CWD	Maximum annual number of consecutive wet days (i.e., P $\geq$ 1 mm)
R10mm	Annual number of days with P $\geq$ 10 mm
PRCPTOT	Annual total precipitation from days with P $\geq$ 1 mm
R20mm	Annual number of days with P $\geq$ 20 mm

Table 1: Description of the CLIMDEX climate indexes. P is daily precipitation (mm). The table is reproduced from http://etccdi.pacificclimate.org/list_27_indices.shtml.

Each annual index is calculated over the period of 122 years from the year 1901 to 2022 on the $1^\circ \times 1^\circ$ grid. The primary data source is the Yearly Gridded Rainfall (1.0 x 1.0) data provided by Indian Meteorological Department (IMD), Pune (https://www.imdpune.gov.in/cmpg/Griddata/Rainfall_1_NetCDF.html).

In this paper, we consider 357 grid locations across the mainland of India (excluding Lakshadweep and Andaman and Nicobar Islands). The mainland of India is divided into 34 climate

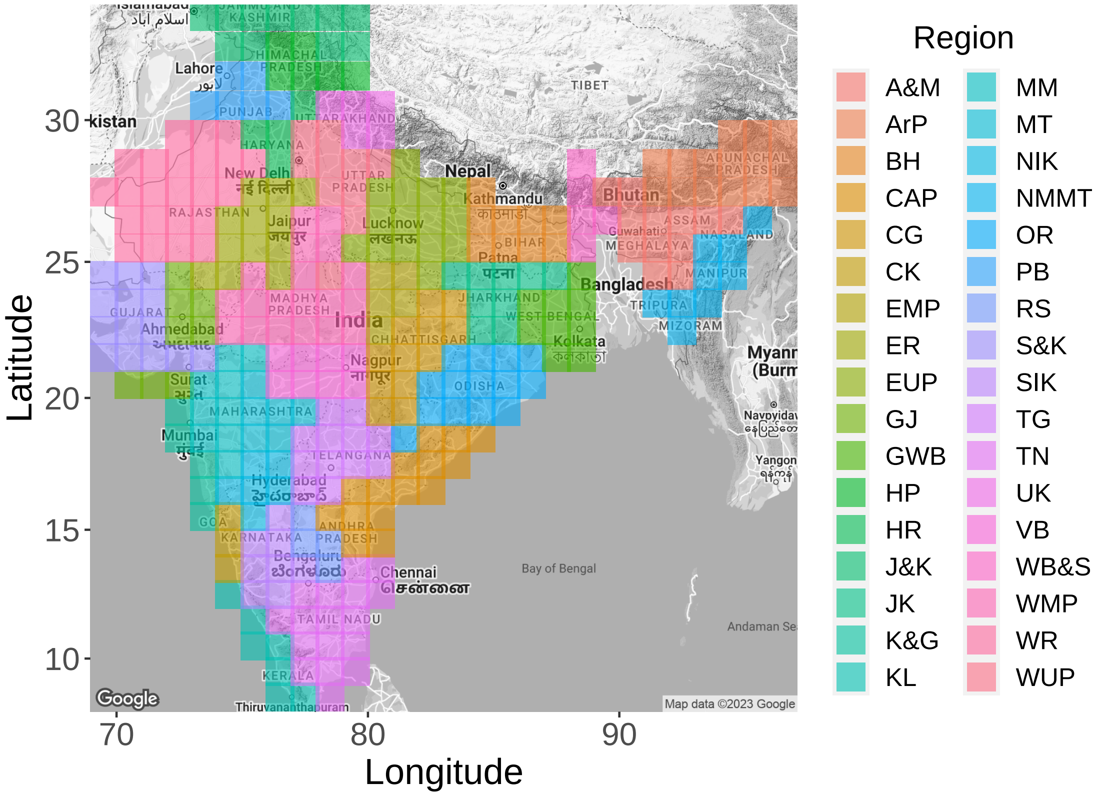

Figure 1: Image depicting the meteorological subdivisions of the 1° × 1° grid covering the mainland of India according to (Guhathakurta and Rajeevan, 2008)

.

regions according to (Guhathakurta and Rajeevan, 2008) presented in Figure 1.: 1. Jammu & Kashmir (J&K), 2. Himachal Pradesh (HP), 3. Punjab (PB), 4. Uttaranchal (UK), 5. Haryana (HR), 6. West Rajasthan (WR), 7. West U.P. (WUP), 8. East Rajasthan (ER), 9. East U.P. (EUP), 10. Bihar (BH), 11. West Bengal & Sikkim (WBS), 12. Assam and Meghalaya (A&M), 13. Arunachal Pradesh (ArP), 14. N.M.M.T. (NMMT), 15. Saurasthra & Kauch (S&K), 16. Gujarat (GJ), 17. West M.P. (WMP), 18. East M.P. (EMP), 19. Chattisgarh (CG), 20. Jharkhand (JK), 21. Ganjetic West Bengal (GWB), 22. Konkan & Goa (K&G), 23. Madhya Maharashtra (MM), 24. Marathwada (MT), 25. Vidarbha (VB), 26. Orissa (OR), 27. Telangana (TG), 28. North Interior Karnataka (NIK), 29. Coastal Karnataka (CK), 30. South Interior Karnataka (SIK), 31. Rayalaseema (RS), 32. Coastal A.P. (CAP), 33. Kerala (KL), 34. Tamilnadu (TN).

We conduct individual analysis for each climatically coherent region, taking into account the heterogeneity of climate anomalies across these regions.

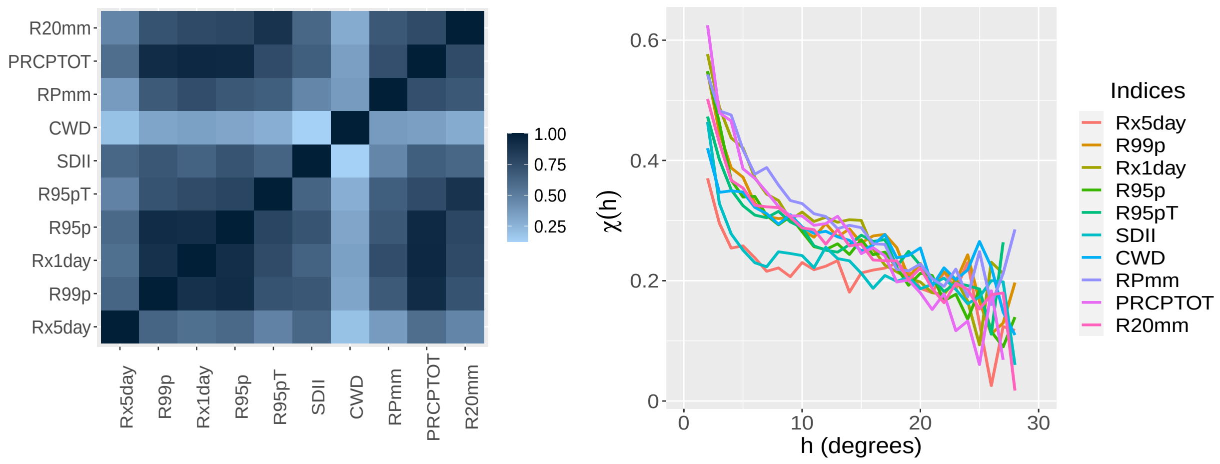

Figure 2: Estimated extremal dependencies between different CLIMDEX indexes (left panel). The spatial extremal dependencies for each CLIMDEX index (right panel)

.

2.2 Necessity of a Multivariate Spatial Model

Contributors: Ringan, Swarnajit

In order to justify the necessity of a multivariate model, we calculate empirical estimates of the extremal dependence between indices and spatial locations. The extremal dependence between two variables Y_1 and Y_2 is often measured using χ-measure ((Sibuya et al., 1960)) given by

$$\chi = \lim_{u \to 1} Pr[Y_1 > F_1^{-1}(u) | Y_2 > F_2^{-1}(u)] \tag{1}$$

where F_1 and F_2 are marginal distribution functions of Y_1 and Y_2 respectively. If the value of χ is close to 1, that indicates strong asymptotic dependence, while a value near 0 indicates no asymptotic dependence. The measure χ is estimated empirically using a measure F-madogram ((Cooley et al., 2006)), which is defined as $\nu_F = \frac{1}{2}E[|F_1(Y_1) - F_2(Y_2)|]$ based on replications of Y_1 and Y_2 and their corresponding empirical distribution functions. Using F-madogram χ is estimated as :

$$\chi = 2 - \frac{1 + 2\nu_F}{1 - 2\nu_F}$$

.

The cross-index χ-measure between two indices p_1 and p_2 by-

$$\chi_{p_1,p_2} = \lim_{u \to 1} Pr[Y_{tp_1}(s) > F^{-1}_{Y_{tp_1}(s)}(u) | Y_{tp_2}(s) > F^{-1}_{Y_{tp_2}(s)}(u)] \tag{2}$$

where $Y_{tp}(s)$ denote the observation at as spatial location s and at time t and $F_{Y_{tp}(s)}$ denotes the marginal distribution function of $Y_{tp}(s)$.

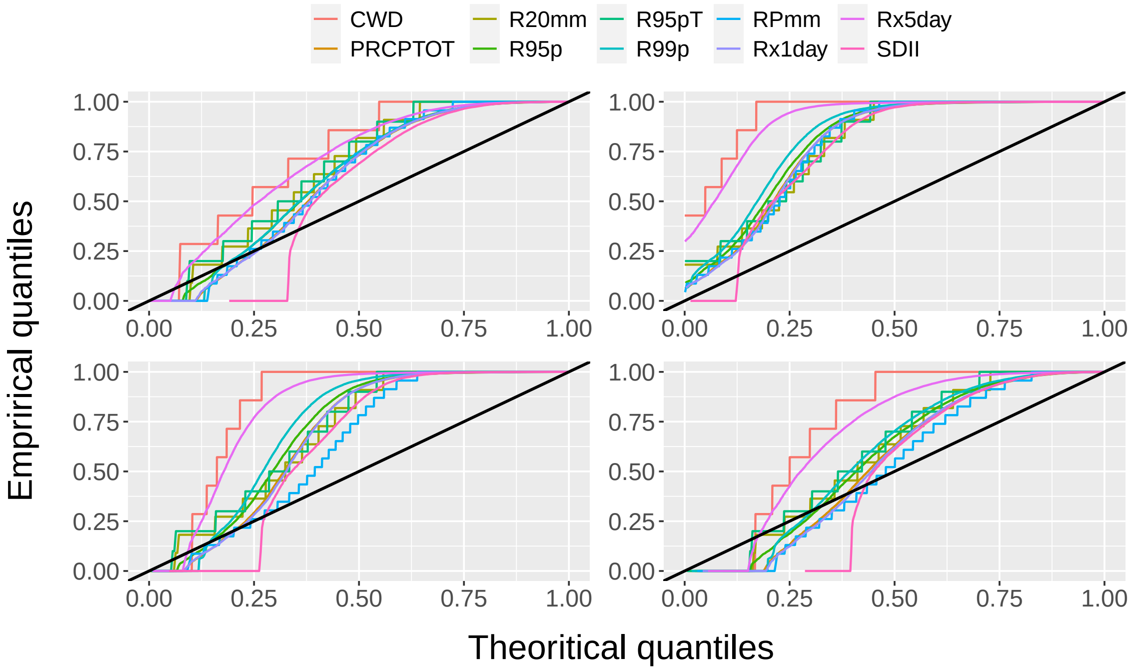

Figure 3: Quantile-Quantile Plot (qqplot) based on fitting different univariate distributions: Normal (top left), t (top right), skew-t (bottom left), GEV (bottom right)

.

For a particular index p (with the isotropic assumption for the spatial process), the spatial dependence between two locations $\boldsymbol{s}$ and $\boldsymbol{s}+\boldsymbol{h}$ is given by -

$$\chi_p(h) = \lim_{u \to 1} Pr[Y_{tp}(\boldsymbol{s}+\boldsymbol{h}) > F^{-1}_{Y_{tp}(\boldsymbol{s}+\boldsymbol{h})}(u) | Y_{tp}(\boldsymbol{s}) > F^{-1}_{Y_{tp}(\boldsymbol{s})}(u)] \tag{3}$$

where $h = \|h\|$, the Euclidean distance between $\boldsymbol{s}$ and $\boldsymbol{s}+\boldsymbol{h}$.

In Fig. 2, the extremal dependencies among different indices are shown in the left panel. The plot provides indications of pronounced extremal dependence among the initial six indices and the final three indices. Also, assuming isotropic correlation across the regions, we estimate the $\chi_p(h)$ using F-madogram for each index separately averaging across the pair of sites within each region and also averaging across the regions. In the right panel, the plot of $\chi_p(h)$ vs h is shown which shows a decreasing pattern for each index but none of them reaches 0 even for high values of h. These motivates a joint analysis of the indexes using a multivariate spatial model.

2.3 Comparison among potential fit of different Parametric Models

Contributors: Ringan, Swarnajit

We employ maximum likelihood estimation to separately fit normal, t, skew-t, and GEV distributions for every grid location and each CLIMDEX index. This allows us to assess the suitability of these parametric models. In fig. 3 we draw quantile-quantile plots (qqplots) i.e. theoretical quantiles versus the fitted data quantiles for those four different distributions, after combining across locations and time. The graphs show that the skew-t and GEV distributions tend to provide a somewhat better fit compared to the normal and t distributions. However, following the original paper, we fit a skew-t model which imposes a significantly lighter computational load.

To conduct a more comprehensive examination of the skew-t model's parameterization, we perform a non-spatial analysis, i.e. we separately evaluate and compare the estimated parameters for each index at every grid location and average them out across the zones. A random variable Y_t follows a univariate skew-t distribution with parameters $(\mu(t), \lambda, a, b)$ if

$$Y_t = \mu(t) + \sigma_t|z_t|\lambda + \sigma_t\epsilon_t$$

where $\epsilon_t \sim N(0,b), z_t \sim N(0,1)$ and $\sigma_t^2 \sim IG(a/2, a/2)$. The trend component μ_t is taken to be a linear combination of 5 cubic splines (one for approximately 25 years). Table 2 lists the mean and coefficient of variation (CV) of the parameters λ, a, b for each index. Instead of λ, $\lambda^* = \lambda/\sqrt{b}$ is listed as it is used to capture the properties of extremal dependence.

	Rx5day	R99p	Rx1day	R95p	R95pT	SDII	CWD	R10 mm	PRCPTOT	R20mm
Mean										
λ^*	0.32	0.03	0.04	0.03	0.33	0.29	0.49	0.23	0.03	0.31
a	3.71	4.08	4.62	4.85	6.65	4.41	3.28	4.86	4.59	6.33
b	10.23	393. 98	198.79	492.30	3.52	5.84	3.04	6.02	544.21	3.77
CV										
λ^*	0.25	0.34	0.34	0.33	0.33	0.35	0.37	0.42	0.34	0.33
a	0.19	0.29	0.26	0.31	0.38	0.29	0.31	0.28	0.29	0.34
b	0.18	0.19	0.17	0.19	0.15	0.16	0.21	0.14	0.17	0.15

Table 2: The estimates of the parameters λ^*, a, b obtained by fitting univariate skew-t distribution at each location.

From the table we can observe that, the parameter b vary too much across different indices (between 3.04, for CWD and 544.21 for PRCPTOT) whereas the variations of the other two parameters are not that much across the indices. So, modelling the scale parameter b varying across spaces is needed.

3 Methodology

Contributors: Arijit, Swarnajit

Let, $Y_{tp}(\boldsymbol{s})$ be the observation corresponding to p-th variable at time t and a spatial location $\boldsymbol{s}$, where $p \in \{1, \ldots, P = 10\}$, $t \in \{1, \ldots, T\}$, and $\boldsymbol{s} \in \mathcal{D} \subseteq \mathbb{R}^2$. We denote $\boldsymbol{Y}_t(\boldsymbol{s}) = [Y_{t1}(\boldsymbol{s}), Y_{t2}(\boldsymbol{s}), \ldots, Y_{tP}(\boldsymbol{s})]$ to be the vector of length P denoting the observation corresponding to all the indices for a location $\boldsymbol{s}$ and for a time point t. In this paper, we are assuming it to be distributed as P-variate skew-t distribution. Hence, following the hierarchical structure of a skew-t distribution described in A.1, we model $\boldsymbol{Y}_t(\boldsymbol{s})$ as

$$\boldsymbol{Y}_t(\boldsymbol{s}) = \boldsymbol{\mu}_t(\boldsymbol{s}) + \sigma_t|z_t|\boldsymbol{\lambda} + \sigma_t\boldsymbol{\epsilon}_t(\boldsymbol{s}), \tag{4}$$

where $\boldsymbol{\mu}_t(\boldsymbol{s}) = [\mu_{t1}(\boldsymbol{s}), \mu_{t2}(\boldsymbol{s}), \ldots, \mu_{tP}(\boldsymbol{s})]'$ gives us the mean structure, $z_t \overset{\text{iid}}{\sim} N(0,1)$, $\sigma_t^2 \overset{\text{iid}}{\sim}$ Inverse-Gamma$(a/2, a/2)$ and $\boldsymbol{\lambda}$ denotes the vector of skewness parameters.

To accommodate the spatial dependence in the observations, we assume the error process $\boldsymbol{\epsilon}_t(\boldsymbol{s}) = \left[\epsilon_{t1}(s), \epsilon_{t2}(s), \ldots, \epsilon_{tP}(s)\right]'$ to follow a P-variate zero-mean Gaussian process It is important to note that the spatial dependence doesn't transcend to the time components, i.e., we take $\boldsymbol{\epsilon}_t(\boldsymbol{s})$'s to be i.i.d. over t. Now if we denote $\boldsymbol{\epsilon}(t) = [\boldsymbol{\epsilon}_t(\boldsymbol{s}_1)', \boldsymbol{\epsilon}_t(\boldsymbol{s}_2)', \ldots, \boldsymbol{\epsilon}_t(\boldsymbol{s}_n)']'$, a vector of length nP, stacking all the error terms corresponding to each variable and for each observed location $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_n\}$, then it's variance can be computed with a covariance structure that is separable across the space and the indices.

If we take the covariance of the error terms across indices for a particular $\boldsymbol{s}$, $\boldsymbol{\epsilon}_t(s)$ as Σ_I, a $P \times P$ dimensional matrix which is independent of $\boldsymbol{s}$, and the covariance of the error terms across space for a particular p, $\boldsymbol{\epsilon}_{tp} = [\epsilon_{tp}(\boldsymbol{s}_1), \epsilon_{tp}(\boldsymbol{s}_2), \ldots, \epsilon_{tp}(\boldsymbol{s}_n)]'$ as Σ_S, a $n \times n$ dimensional matrix which is independent of p, then with the separability assumption, the variance of $\boldsymbol{\epsilon}_t$ can be written as $\Sigma_S \otimes \Sigma_I$.

The covariance matrix Σ_S is constructed using the spatial covariance of the pairwise components of $\boldsymbol{\epsilon}_{tp}$, which are assumed to follow an isotropic Matérn structure, given as

$$\text{cor}[\epsilon_{tp}(\boldsymbol{s}_i), \epsilon_{tp}(\boldsymbol{s}_j)] = \frac{\gamma}{\Gamma(\nu)2^{\nu-1}}\left(\frac{h}{\rho}\right)^{\nu} K_\nu\left(\frac{h}{\rho}\right) + (1-\gamma)I(h=0), \quad i, j \in 1, \ldots, n,$$

where h is the distance between $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ given as $\|\boldsymbol{s}_i - \boldsymbol{s}_j\|$, $\rho > 0$ and $\nu > 0$ are the range and smoothness parameter of the Matérn correlation structure respectively, $\gamma \in [0,1]$ is the ratio of spatial to total variation, K_ν is the Modified Bessel function of degree ν and $I(u, v) = 1$ if $u = v$, otherwise 0.

We take $\boldsymbol{\mu}_t(s)$ to be a multivariate spatiotemporal mean process. For the purpose of a trend analysis, we assume that $\mu_{tp}(s)$'s are smooth functions of t for each $\boldsymbol{s}$ and p. We take their form as

$$\mu_{tp}(\boldsymbol{S}) = \sum_{l=1}^{L} \beta_{lp}(\boldsymbol{s}) B_l(t)$$

where $B_l(t)$ are known cubic B-spline functions defined over the time component and in the range $[0, T]$ and $\beta_{lp}(\cdot)$ are spatially varying spline coefficients.

From the above structure, it is evident that we are taking the same basis functions for each $\boldsymbol{s}$ and p. For the coefficients, we take spatial Gaussian prior on $\boldsymbol{\beta}(\cdot)$, where $\boldsymbol{\beta}(\boldsymbol{s}) = [\boldsymbol{\beta}_1(\boldsymbol{s})', \boldsymbol{\beta}_2(\boldsymbol{s})', \ldots, \boldsymbol{\beta}_P(\boldsymbol{s})']'$ and $\boldsymbol{\beta}_p(\boldsymbol{s}) = [\beta_{1,p}(\boldsymbol{s}), \beta_{2,p}(\boldsymbol{s}), \ldots, \beta_{L,p}(\boldsymbol{s})]'$. In a similar manner to the spatial error process, here also, we assume the covariance of $\boldsymbol{\beta} = [\boldsymbol{\beta}(\boldsymbol{s_1})', \boldsymbol{\beta}(\boldsymbol{s_2})', \ldots, \boldsymbol{\beta}(\boldsymbol{s_n})']'$ to be separable across space, indices, and splines, the main reason being the computational ease.

For each $\boldsymbol{s}$ and each p, the covariance of $[\beta_{1p}(\boldsymbol{s}), \beta_{2p}(\boldsymbol{s}), \ldots, \beta_{Lp}(\boldsymbol{s})]'$ is given by a $L \times L$ matrix Σ_B, which is independent of $\boldsymbol{s}$ and p and controls the correlation between the spline coefficients. Furthermore, we make the assumption that the spatial correlation structure and the correlation among the components of $\boldsymbol{\beta}$ are identical to those observed in the error process. That is, for a particular l and p, the covariance of $[\beta_{lp}(\boldsymbol{s_1}), \beta_{lp}(\boldsymbol{s_2}), \ldots, \beta_{lp}(\boldsymbol{s_n})]'$ is given by Σ_S. On the other hand, for particular value of $\boldsymbol{s}$ and l, covariance of $[\beta_{l1}(\boldsymbol{s}), \beta_{l2}(\boldsymbol{s}), \ldots, \beta_{lP}(\boldsymbol{s})]'$ is given by Σ_I. And with the separable structure assumption, the covariance of $\boldsymbol{\beta}$ is given as $\Sigma_S \otimes \Sigma_I \otimes \Sigma_B$. As for the mean, all the components of $\boldsymbol{\beta}$ with index p are assumed to have mean $\mu_{\boldsymbol{\beta}p}$. Thus, if $\boldsymbol{\mu_\beta} = [\mu_{\boldsymbol{\beta}1}, \mu_{\boldsymbol{\beta}2}, \ldots, \mu_{\boldsymbol{\beta}P}]'$, the mean of $\boldsymbol{\beta}$ is given as $\mathbf{1}_n \otimes \mu_{\boldsymbol{\beta}} \otimes \mathbf{1}_L$. So finally the spatial Gaussian prior for $\boldsymbol{\beta}$ is given by

$$\boldsymbol{\beta} \sim \mathrm{N}_{nLp}(\mathbf{1}_n \otimes \boldsymbol{\mu_\beta} \otimes \mathbf{1}_L, \Sigma_S \otimes \Sigma_I \otimes \Sigma_B). \tag{5}$$

Now, coming back to the distribution of $\boldsymbol{Y}$'s, we mentioned the P-component vector $\mathbf{Y}_t(\boldsymbol{s}) = [Y_{t1}(\boldsymbol{s}), Y_{t2}(\boldsymbol{s}), \ldots, Y_{tP}(\boldsymbol{s})]$ is assumed to follow the P-variate skew-t distribution. Also, after stacking them across n spatial locations, the nP-component vector $\mathbf{Y}_t = [\mathbf{Y}_t(\boldsymbol{s}_1)', \mathbf{Y}_t(\boldsymbol{s}_2)', \ldots, \mathbf{Y}_t(\boldsymbol{s}_n)']'$ is assumed to follow a nP-variate skew-t distribution. With all the covariance matrices elaborated, below we give the distribution of $\boldsymbol{Y}$'s. Following the notation of (Azzalini, 2013), we marginalize over z_t and σ_t^2 from (4), to get the distribution of $\mathbf{Y}_t$ and $\mathbf{Y}_t(\boldsymbol{s}_1)$, which is given as

$$\begin{aligned} &\mathbf{Y}_t(\boldsymbol{s}_i) \sim \mathrm{ST}_P(\boldsymbol{\mu}_t(\boldsymbol{s}_i), \Sigma_I + \boldsymbol{\lambda\lambda}', \Sigma_I^{-1}\boldsymbol{\lambda}, a), \text{ and} \\ &\mathbf{Y}_t \sim \mathrm{ST}_{nP}(\boldsymbol{\mu}_t, \Sigma_S \otimes \Sigma_I + (\mathbf{1}_n\mathbf{1}_n') \otimes \boldsymbol{\lambda\lambda}', (\Sigma_s^{-1}\mathbf{1}_n) \otimes (\Sigma_I^{-1}\boldsymbol{\lambda}), a), \end{aligned} \tag{6}$$

where $\boldsymbol{\mu}_t = [\boldsymbol{\mu}_t(\boldsymbol{s}_1)', \boldsymbol{\mu}_t(\boldsymbol{s}_2)', \ldots, \boldsymbol{\mu}_t(\boldsymbol{s}_n)']'$. We want to emphasize the fact that every distribution is independent and identically distributed along the time components, i.e., the distribution of $\boldsymbol{Y}_t(\boldsymbol{s}_i)$ and $\boldsymbol{Y}_t$ from (6) is true for all $t = 1, 2, \ldots, T$.

Also, with the construction of the basis-spline representation and hierarchical formation of $\boldsymbol{\beta}$, we have $\boldsymbol{\mu}_t = \boldsymbol{X}_t\boldsymbol{\beta}$, where $\boldsymbol{X}_t = I_{nP} \otimes \boldsymbol{x}_t$, with $\boldsymbol{x}_t = [B_1(t), B_2(t), \ldots, B_L(t)]'$ is the vector of B-spline at time t.

4 Computation

Contributors: Arijit, Pratyusha, Swarnajit

We draw inferences about the model parameters by employing Markov chain Monte Carlo (MCMC). We will first describe the prior specifications for all the parameters. Then, combining them with the likelihood, we specify the full posterior distribution of all the parameters together. We will try to use conjugate priors wherever possible. For MCMC, we employ either Gibbs sampler for the parameters with conjugate parameters and Metropolis-Hastings for the parameters with no conjugate priors. For the updation of the parameters, we present full conditional distributions of the individual parameters given the data and the rest of the parameters.

For the purpose of computation, we parametrize our main model from (4), absorbing all the terms containing σ_t into other parameters. It is given as

$$\boldsymbol{Y}_t(\boldsymbol{s}) = \boldsymbol{\mu}_t(\boldsymbol{s}) + |z_t|\boldsymbol{\lambda} + \boldsymbol{\epsilon}_t(\boldsymbol{s}), \tag{7}$$

where $z_t|\sigma_t \sim N(0, \sigma_t^2)$, and now the error process has one additional terms of σ_t^2 in it's covariance matrix, i.e., $\text{Var}(\boldsymbol{\epsilon}_t(\boldsymbol{s})) = \sigma_t^2\, \Sigma_I$, or $\text{Var}(\boldsymbol{\epsilon}_{tp}) = \sigma_t^2\, \Sigma_S$.

Also, rewriting $\boldsymbol{\mu}_t$ in therms of the parameters, the distribution of $\boldsymbol{Y}_t$ is given as

$$\boldsymbol{Y}_t \sim \text{ST}_{nP}(\boldsymbol{X}_t\boldsymbol{\beta}, \Sigma_S \otimes \Sigma_I + (\mathbf{1}_n\mathbf{1}_n') \otimes \boldsymbol{\lambda}\boldsymbol{\lambda}', (\Sigma_s^{-1}\mathbf{1}_n) \otimes (\Sigma_I^{-1}\boldsymbol{\lambda}), a) \tag{8}$$

Notably, here, the list of all the parameters, as well as the latent variables that are needed to update in MCMC, is given as

$$\Theta = \{\boldsymbol{\beta}, \boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\lambda}, \{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T, a, \Sigma_I, \Sigma_B, \rho, \nu, \gamma\}$$

We specify the prior distributions of the parameters used above. We consider the prior $N(\mathbf{0}, 100^2 I_P)$ for $\mu_{\boldsymbol{\beta}}$, which is a conjugate one. For $\boldsymbol{\lambda}$, we take the conjugate prior as $\boldsymbol{\lambda} \sim N(\mathbf{0}, 10^2 I_P)$. For a, we consider a discrete uniform prior as $a \sim \text{DU}(0.1, 0.2, \ldots, 2.0)$. For

Σ_I and Σ_B, we take inverse-Wishart conjugate prior as $\Sigma_I \sim \mathrm{IW}(0.01, 0.01 I_P)$ and $\Sigma_B \sim \mathrm{IW}(0.01, 0.01 I_L)$. There are no established conjugate prior distributions available for the Matérn correlation parameters ρ, ν, and γ. We take their prior as $\rho \sim \mathrm{U}(0, \|\mathcal{D}\|)$, $\log(\nu) \sim \mathrm{N}(-1.2, 1^2)$, and $\gamma \sim \mathrm{U}(0, 1)$, where $\|\mathcal{D}\|$ denotes the maximum distance between two points in the spatial domain in degrees.

Denoting $\boldsymbol{Y}$ as $[\boldsymbol{Y}_1', \boldsymbol{Y}_2', \ldots, \boldsymbol{Y}_T']'$, the combining the likelihood of it to the prior distribution of the parameters, the full posterior distribution of all the parameters is given as

$$
\begin{aligned}
&f\big(\boldsymbol{\beta}, \mu_{\boldsymbol{\beta}}, \boldsymbol{\lambda}, \{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T, a, \Sigma_I, \Sigma_B, \rho, \nu, \gamma | \mathbf{Y}\big) \\
&\propto f\big(\boldsymbol{Y} | \boldsymbol{\beta}, \mu_{\boldsymbol{\beta}}, \boldsymbol{\lambda}, \{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T, a, \Sigma_I, \Sigma_B, \rho, \nu, \gamma\big) \\
&\qquad \times f\big(\boldsymbol{\beta}, \mu_{\boldsymbol{\beta}}, \boldsymbol{\lambda}, \{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T, a, \Sigma_I, \Sigma_B, \rho, \nu, \gamma\big) \\
&= \Big[\Pi_{t=1}^T f\big(\mathbf{Y}_t | \boldsymbol{\beta}, \mu_{\boldsymbol{\beta}}, \boldsymbol{\lambda}, \{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T, a, \Sigma_I, \Sigma_B, \rho, \nu, \gamma\big)\Big] \\
&\qquad \times f\big(\boldsymbol{\beta}, \mu_{\boldsymbol{\beta}}, \boldsymbol{\lambda}, \{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T, a, \Sigma_I, \Sigma_B, \rho, \nu, \gamma\big) \\
&= \Big[\Pi_{t=1}^T f\big(\mathbf{Y}_t | \boldsymbol{\beta}, \boldsymbol{\lambda}, z_t, \sigma_t^2, a, \Sigma_I, \rho, \nu, \gamma\big)\Big] \\
&\qquad \times f\big(\boldsymbol{\beta} | \mu_{\boldsymbol{\beta}}, \boldsymbol{\lambda}, \{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T, a, \Sigma_I, \Sigma_B, \rho, \nu, \gamma\big) \\
&\qquad\qquad \times f(\mu_{\boldsymbol{\beta}}, \boldsymbol{\lambda}, \{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T, a, \Sigma_I, \Sigma_B, \rho, \nu, \gamma) \\
&= \Big[\Pi_{t=1}^T f\big(\mathbf{Y}_t | \boldsymbol{\beta}, \boldsymbol{\lambda}, z_t, \sigma_t^2, a, \Sigma_I, \rho, \nu, \gamma\big)\Big] \\
&\qquad \times f\big(\boldsymbol{\beta} | \mu_{\boldsymbol{\beta}}, \Sigma_I, \Sigma_B, \rho, \nu, \gamma\big) \times f(\{z_t\}_{t=1}^T, \{\sigma_t^2\}_{t=1}^T | a) \times f(a) \\
&\qquad\qquad \times f(\boldsymbol{\lambda}) \times f(\mu_{\boldsymbol{\beta}}) \times f(\Sigma_I) \times f(\Sigma_B) \times f(\rho) \times f(\nu) \times f(\gamma) \\
&= \Big[\Pi_{t=1}^T f\big(\mathbf{Y}_t | \boldsymbol{\beta}, \boldsymbol{\lambda}, z_t, \sigma_t^2, a, \Sigma_I, \rho, \nu, \gamma\big)\Big] \\
&\qquad \times f\big(\boldsymbol{\beta} | \mu_{\boldsymbol{\beta}}, \Sigma_I, \Sigma_B, \rho, \nu, \gamma\big) \times \Pi_{t=1}^T \Big[f(z_t | \sigma_t) \times f(\sigma_t^2 | a)\Big] \\
&\qquad\qquad \times f(a) \times f(\boldsymbol{\lambda}) \times f(\mu_{\boldsymbol{\beta}}) \times f(\Sigma_I) \times f(\Sigma_B) \times f(\rho) \times f(\nu) \times f(\gamma)
\end{aligned}
$$

We further specify the distribution of latent variables. The expression of $f(\boldsymbol{\beta}|\cdot)$ is given by (5), distribution of $(Z_t|\sigma_t)$ is given after (7), and distribution of $(\sigma_t^2|a)$ is given after (4).

The full posterior distributions are given below:

$\underline{\beta | rest}$

The posterior density of β is $\beta | rest \sim N_p((\mu_\beta^*, \Sigma_\beta^*))$ where

$$
\Sigma_\beta^* = \Sigma_S \otimes \Sigma_I \otimes \left[\Big(\sum_{t=1}^T \frac{1}{\sigma_t^2} \boldsymbol{x}_t \boldsymbol{x}_t' \Big) + \Sigma_B^{-1} \right]^{-1}
$$

$$\mu_\beta^* = \Sigma_\beta^* \left[\sum_{t=1}^{T} \frac{1}{\sigma_t^2} [(\Sigma_S^{-1} \otimes \Sigma_I^{-1})(\boldsymbol{Y}_t - |z_t| \mathbf{1}_{\boldsymbol{n}} \otimes \boldsymbol{\lambda})] \otimes \boldsymbol{x}_t + (\Sigma_S^{-1} \mathbf{1}_n) \otimes (\Sigma_I^{-1} \boldsymbol{\mu}_\beta) \otimes (\Sigma_B^{-1} \mathbf{1}_L) \right]$$

<u>$\mu_\beta | rest$</u>

Suppose $\mathbf{B}$ denotes the $nL \times P$ matrix with the p-th column $\boldsymbol{B}_p = [\beta_p(s_1)', \beta_p(s_2)', \ldots, \beta_p(s_n)']'$.We consider the prior $N(0, 100^2 I_P)$.The posterior density of μ_β is $\mu_\beta | rest \sim N_p(\bar{\mu}_\beta, \Sigma_{\mu_\beta})$ where

$$\Sigma_{\mu_\beta} = \left[(\mathbf{1}_n' \Sigma_S^{-1} \mathbf{1}_n)(\mathbf{1}_L' \Sigma_B^{-1} \mathbf{1}_L) \Sigma_I^{-1} + 100^{-2} I_P \right]^{-1}$$

$$\bar{\mu}_\beta = \Sigma_{\mu_\beta} \Sigma_I^{-1} B' [(\Sigma_S^{-1} \mathbf{1}_n) \otimes (\Sigma_B^{-1} \mathbf{1}_L)]$$

<u>$\lambda | rest$</u>

We consider the prior $\boldsymbol{\lambda} \sim N(\mathbf{0}, 10^2 I_P)$. The posterior density of $\boldsymbol{\lambda}$ is $\boldsymbol{\lambda} | rest \sim N_p(\boldsymbol{\mu}_{\boldsymbol{\lambda}}^*, \Sigma_\lambda^*)$ where

$$\Sigma_\lambda^* = \left[(\mathbf{1}_n' \Sigma_S^{-1} \mathbf{1}_n) \left(\sum_{t=1}^{T} \frac{z_t^2}{\sigma_t^2} \right) \Sigma_I^{-1} + 10^{-2} I_p \right]^{-1}$$

$$\mu_\lambda^* = \Sigma_\lambda^* \left[\Sigma_I^{-1} \left(\sum_{t=1}^{T} \frac{|z_t|}{\sigma_t^2} [Y_t^* - \mu_t^*]' \right) (\Sigma_S^{-1} \mathbf{1}_n) \right]$$

where Y_t^* and μ_t^* are $n \times P$ matrices with the p-th columns are $[Y_{tp}(s_1), \ldots, Y_{tp}(s_n)]'$ and $[\mu_{tp}(s_1), \ldots, \mu_{tp}(s_n)]'$ respectively.

<u>$|z_t| | rest$</u>

As z_t is not identifiable, we treat $|z_t|$ as a parameter and update within the MCMC steps. Here $|z_t| \sim HN(\sigma_t^2)$ where HN denotes the half-normal density. The posterior density of $|z_t|$ conditioned on rest is given by

$|z_t| \ | rest \sim N_{(0,\infty)}(\mu_z^*, \sigma_z^{*2})$ where,

$$\sigma_z^{*2} = \sigma_t^2 [1 + (\mathbf{1}_n' \Sigma_S^{-1} \mathbf{1}_n)(\boldsymbol{\lambda}' \Sigma_I^{-1} \boldsymbol{\lambda})]^{-1}$$

$$\mu_z^* = [1 + (\mathbf{1}_n' \Sigma_S^{-1} \mathbf{1}_n)(\boldsymbol{\lambda}' \Sigma_I^{-1} \boldsymbol{\lambda})]^{-1} [\mathbf{1}_n' \Sigma_S^{-1} [Y_t^* - \mu_t^*] \Sigma_I^{-1} \boldsymbol{\lambda}]$$

where Y_t^* and μ_t^* are $n \times P$ matrices as defined for calculating the posterior density of λ.

$\underline{\sigma_t^2|rest}$

The posterior density of $\sigma_t^2|rest$ is

$$\sigma_t^2|rest \sim IG\left(\frac{a+nP+1}{2}, \frac{a+(\boldsymbol{Y}_t-\boldsymbol{\mu}_t-|z_t|\mathbf{1}_n\otimes\boldsymbol{\lambda})'\Sigma_S^{-1}\otimes\Sigma_I^{-1}(\boldsymbol{Y}_t-\boldsymbol{\mu}_t-|z_t|\mathbf{1}_n\otimes\boldsymbol{\lambda})+z_t^2}{2}\right)$$

$\underline{a|rest}$

Considering discrete uniform prior of a over $(0.1, 0.2, \ldots, 19.9, 20.0)$, the posterior distribution of a given rest is

$$Pr(a=a^*|rest) \propto \prod_{t=1}^{T} f_{IG}(\sigma_t^2; \frac{a^*}{2}, \frac{a^*}{2})$$

where f_{IG} denotes the inverse gamma density. We draw random sample from the discrete support $0.1, 0.2, \ldots, 19.9, 20.0$ with probabilities proportional to $Pr(a=a^*|rest)$.

$\underline{\Sigma_I|rest}$

The prior of Σ_I is considered as $\Sigma_I \sim IW(0.01, 0.01I_P)$. The posterior density of Σ_I given rest is then $IW(\nu_I^*, \Psi_I^*)$ where

$$\nu_I^* = 0.01 + NT + NL$$

$$\Psi_I^* = 0.01I_P + \sum_{t=1}^{T}(\boldsymbol{Y}_t^*-\boldsymbol{\mu}_t^*-|z_t|\mathbf{1}_n\otimes\boldsymbol{\lambda}')'\Sigma_S^{-1}(\boldsymbol{Y}_t^*-\boldsymbol{\mu}_t^*-|z_t|\mathbf{1}_n\otimes\boldsymbol{\lambda}')$$
$$+ (B-\mathbf{1}_{nL}\otimes\boldsymbol{\mu}_{\boldsymbol{\beta}}')'\Sigma_S^{-1}\otimes\Sigma_B^{-1}(B-\mathbf{1}_{nL}\otimes\boldsymbol{\mu}_{\boldsymbol{\beta}}')$$

where Y_t^* and μ_t^* are $n \times P$ matrices defined in the posterior density of $\boldsymbol{\lambda}$ and B is defined for the posterior density of $\boldsymbol{\mu_B}$.

$\underline{\Sigma_B|rest}$

The prior of Σ_B is considered as $\Sigma_I \sim IW(0.01, 0.01I_P)$. The posterior density of Σ_I given rest is

$$\Sigma_B|rest \sim IW(0.01+nP, 0.01I_L + (B^*-\mathbf{1}_L'\otimes\mathbf{1}_n\otimes\boldsymbol{\mu}_{\boldsymbol{\beta}})'\Sigma_S^{-1}\otimes\Sigma_I^{-1}(B^*-\mathbf{1}_L'\otimes\mathbf{1}_n\otimes\boldsymbol{\mu}_{\boldsymbol{\beta}}))$$

where B^* denotes the $nP \times L$ matrix with the l-th column $B_l^* = [\beta_l^*(s_1)', \ldots, \beta_l^*(s_n)']$ and $\beta_l^*(s_i) = [\beta_{l1}(s_i), \ldots, \beta_{lP}(s_i)]'$.

$\underline{\rho, \nu, \gamma | rest}$

These three parameters do not have any known conjugate priors, which is why we update them using the Metropolis-Hastings algorithm. We discussed the prior for these parameters as $\rho \sim \mathrm{U}(0, \|D\|)$, $\log(\nu) \sim \mathrm{N}(-1.2, 1^2)$, and $\gamma \sim \mathrm{U}(0,1)$. Now, as the support of the parameters are $[0, \infty)$, $[0, \infty)$ and $[0, 1]$, we first transform them accordingly to be a real-valued quantity to felicitate the process of taking a sample from the candidate distribution which we have taken to be a normal distribution with mean being the parameter value of the previous iteration and aptly tuned variance value. We take logarithmic transformation on ρ and ν and logit transformation on the γ parameter. Let, $\rho_{(m)}$, $\nu_{(m)}$, and $\gamma_{(m)}$ be the current value of the the parameter at $(m+1)$-th iteration of the MCMC. Let the transformed values are denoted as $\rho^*_{(m)} = \log[(\rho)_{(m)}]$, $\nu^*_{(m)} = \log[(\rho)_{(m)}]$, and $\gamma^*_{(m)} = \mathrm{logit}[\gamma_{(m)}]$, where $\mathrm{logit(p)} = \log[p/(1-p)]$. We draw candidate samples as $\rho^*_{(c)} \sim \mathrm{N}(\rho^*_{(m)}, s^2_\rho)$, $\nu^*_{(c)} \sim \mathrm{N}(\nu^*_{(m)}, s^2_\nu)$, and $\gamma^*_{(c)} \sim \mathrm{N}(\gamma^*_{(m)}, s^2_\gamma)$. Now, to get the candidate values of the parameters in their actual support, we employ the corresponding inverse function, given as $\rho_{(c)} = \exp[\rho^*_{(c)}]$, $\nu_{(c)} = \exp[\nu^*_{(c)}]$, and $\gamma_{(c)} = \mathrm{expit}[\gamma^*_{(c)}]$, where 'expit' function is the inverse of the 'logit' function, with $\mathrm{expit}(p) = \exp(p)/[1 + \exp(p)]$.

Now, with all the candidate values of the three parameters $\rho_{(c)}, \nu_{(c)}$, and $\gamma_{(c)}$, let the candidate value of the spatial correlation matrix is given as $\Sigma_S^{(c)}$. Also, say $\Sigma_S^{(m)}$ the value of the same for the parameter values from the previous iteration, i.e., $\rho_{(m)}, \nu_{(m}$, and $\gamma_{(m)}$. Then, we accept the proposed values of the thress parameters for the current iteration with probability $p = \min(R, 1)$, where

$$R = \frac{\Pi_{t=1}^T N_{nP}\left(\boldsymbol{Y}_t; \boldsymbol{\mu}_t + |z_t|\mathbf{1}_n \otimes \boldsymbol{\lambda}, \sigma_t^2 \Sigma_S^{(m)} \otimes \Sigma_I\right)}{\Pi_{t=1}^T N_{nP}\left(\boldsymbol{Y}_t; \boldsymbol{\mu}_t + |z_t|\mathbf{1}_n \otimes \boldsymbol{\lambda}, \sigma_t^2 \Sigma_S^{(c)} \otimes \Sigma_I\right)} \times \frac{f(\nu^*_{(c)})}{f(\nu^*_{(c)})} \times \frac{\rho_{(c)}(D - \rho_{(c)})}{\rho_{(m)}(D - \rho_{(m)})},$$

where $f(\nu^*)$ is the prior distribution of $\nu^* = \log(\nu)$, and D is as defined before. Also, as it is with the Metropolis-Hastings algorithm, if the proposed values are rejected, we keep the values of the parameters from the previous iteration for this iteration, too.

5 Data application

Contributors: Pratyusha, Swarnajit

We utilized the dataset introduced in Section 2.1 to drive the conducted model fitting. The data comprises 10 different indices regarding extreme precipitation during 1901-2022, focusing on 34 different regions across the mainland of India. As outlined in Section 3 we employed a Skew-t process on the dataset using the CLIMDEX indices as response variables. This analysis was done separately across all 34 regions. We are interested to study the parameter $\Delta_p(s) = [\mu_{122,p}(s) - \mu_{1,p}(s)]/5$, the average change over every 25 years.

As per the fitting of the proposed MSTP model, the estimates of the posterior mean of $\Delta_p(s)$ for every p and s are provided in fig. 4 and 5.

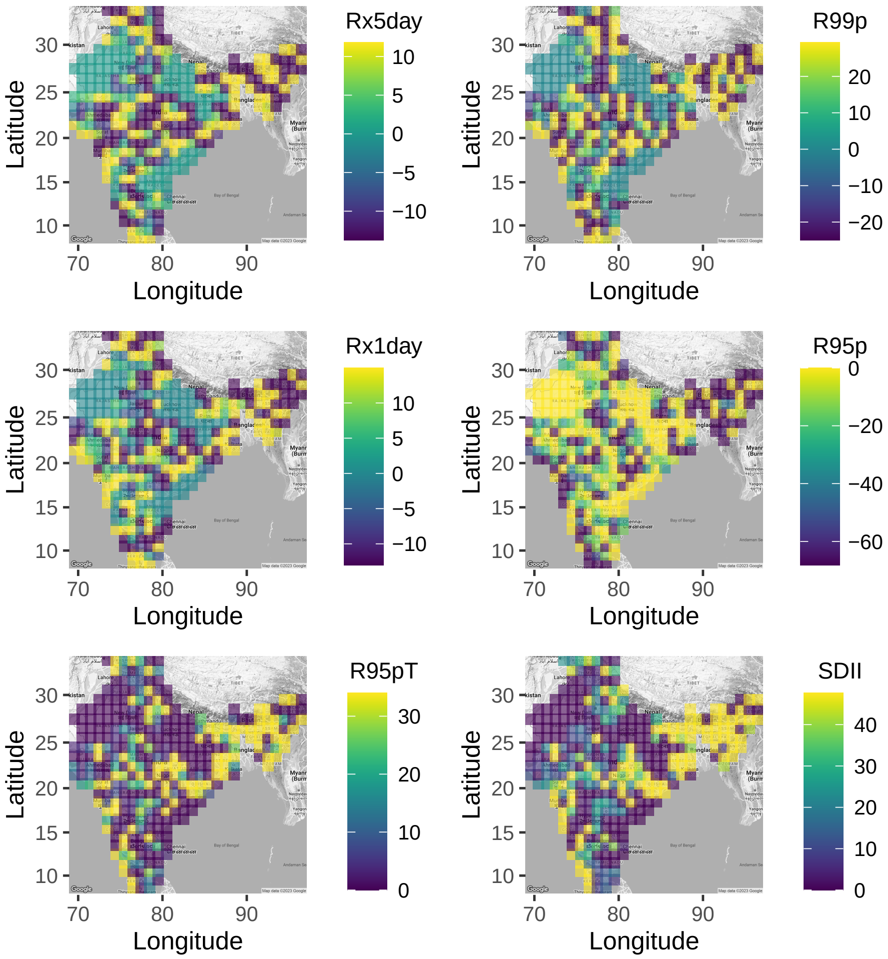

Figure 4: Spatial maps of the posterior mean change per 25 years for the first 6 precipitation indexes.

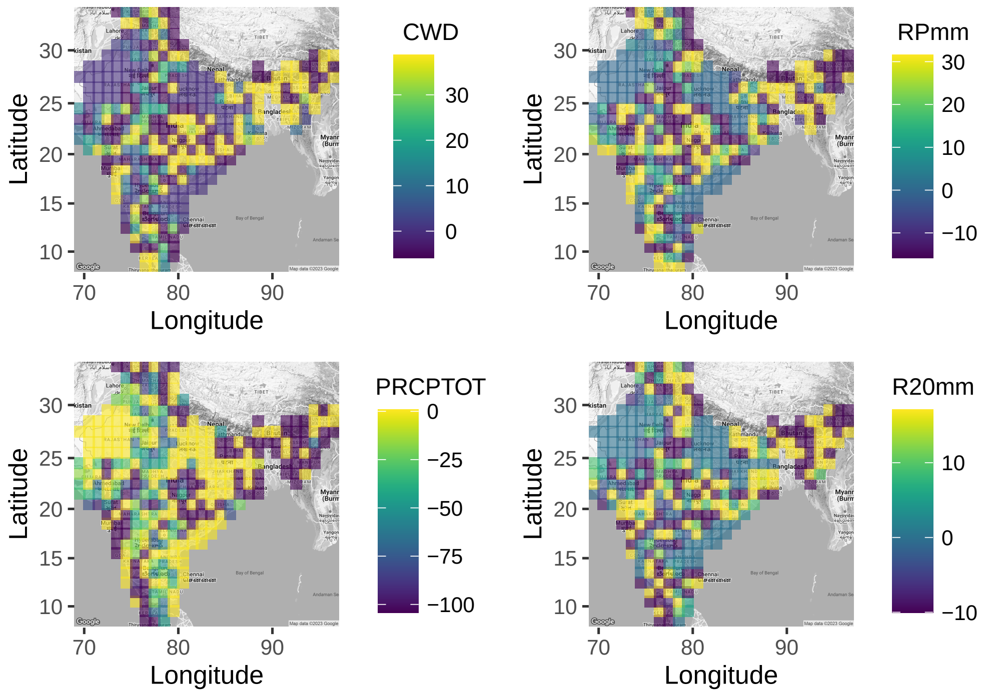

Figure 5: Spatial maps of the posterior mean change per 25 years for the last 4 precipitation indexes.

A Appendix

A.1 Skew-t distribution

A random variable Y_t is said to follow a univariate skew-t distribution with parameters (μ, λ, a, b) if Y_t is given as $Y_t = \mu + \sigma\lambda\,|Z| + \sigma\epsilon$, where $\epsilon \sim \mathrm{N}(0, b), Z \sim \mathrm{N}(0, 1), \sigma \sim \text{Inerse-Gamma}(a/2, a/2)$. Here, μ is the location parameter, λ is the skewness parameter, a is the degrees of freedom parameter, and b is the scale parameter. The distribution is implemented in the `skewt` (King and Anderson, 2021) package in `R`.

A.2 Measures for Extremal Dependence

The $\mathcal{X}$-measure (chi-measure), also known as the $\mathcal{X}$-coefficient, is a statistical measure used to quantify extremal dependence between two random variables. Extremal dependence refers to the dependence structure that becomes evident only when the variables are in the tail regions of their respective distributions. In other words, it focuses on the extreme values of the variables.

The $\mathcal{X}$-measure is defined as the probability that two variables simultaneously exceed their respective high thresholds, given that they have already exceeded them individually. Mathematically, for two random variables X and Y, the $\mathcal{X}$-measure is calculated as

$$\mathcal{X} = \Pr(X > x, Y > y \mid X > x \text{ or } Y > y)$$

where x and y are the thresholds that define what is considered an 'extreme' value for each variable. The $\mathcal{X}$-measure ranges from 0 to 1. For $\mathcal{X} = 0$, we say that there is no extremal dependence. The occurrence of extreme values in one variable does not affect the occurrence of extreme values in the other variable. For $0 < \mathcal{X} < 1$, we say there is partial extremal dependence. The occurrence of extreme values in one variable influences the occurrence of extreme values in the other variable, but not in a perfectly coordinated manner. And for $\mathcal{X} = 1$, we say there is perfect extremal dependence. The occurrence of extreme values in one variable guarantees the occurrence of extreme values in the other variable.

A.3 Kronecker product of two matrices

The symbol '$A \otimes B$' typically represents the Kronecker product (also known as the tensor product) of two matrices, A and B. The Kronecker product is a way of combining two matrices to create a larger block matrix. Mathematically, if we have two matrices A with dimension $(m \times n)$ and B with dimension $(p \times q)$, then the Kronecker product $A \otimes B$ results in a new matrix with dimension $(mp \times nq)$. It is defined as follows

$$A \otimes B = \begin{pmatrix} a_{11} * B & a_{11} * B & \cdots & a_{1n} * B \\ a_{21} * B & a_{22} * B & \cdots & a_{2n} * B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} * B & a_{m2} * B & \cdots & a_{mn} * B \end{pmatrix} \tag{9}$$

Here, a_{ij} represents the (i, j)-th element of matrix A, and $*$ denotes the element-wise multiplication (scalar multiplication) between a matrix and a scalar. The resulting matrix $A \otimes B$ will have blocks of matrix B, scaled by the elements of matrix A.

A.4 Inverse Wishart Distribution

- Support: Real-values matrix-space
- Notation: $X \sim \mathcal{W}^{-1}(\mathbf{\Psi}, \nu)$
- PDF:

$$f_X(\boldsymbol{X}; \mathbf{\Psi}, \nu) = \frac{|\mathbf{\Psi}|^{\nu/2}}{2^{\nu p/2}\Gamma_p(\nu/2)} |\boldsymbol{X}|^{-(\nu+p+1)/2} \exp\Big(-\frac{1}{2}\mathrm{tr}(\mathbf{\Psi}\boldsymbol{X}^{-1})\Big)$$

References

Azzalini, A. (2013). *The skew-normal and related families*, volume 3. Cambridge University Press.

Banerjee, S. and Gelfand, A. (2002). Prediction, interpolation and regression for spatially misaligned data. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 227–245.

Cooley, D., Naveau, P., and Poncet, P. (2006). Variograms for spatial max-stable random fields. In *Dependence in probability and statistics*, pages 373–390. Springer.

Davison, A. C., Huser, R., and Thibaud, E. (2013). Geostatistics of dependent and asymptotically independent extremes. *Mathematical Geosciences*, 45:511–529.

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., and Caesar, J. (2013). Global land-based datasets for monitoring climatic extremes. *Bulletin of the American Meteorological Society*, 94(7):997–1006.

Gelfand, A. E., Diggle, P., Guttorp, P., and Fuentes, M. (2010). *Handbook of spatial statistics*. CRC press.

Guhathakurta, P. and Rajeevan, M. (2008). Trends in the rainfall pattern over india. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 28(11):1453–1469.

IPCC, C. C. et al. (2007). The physical science basis. contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change. *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, 996(2007):113–119.

Kendall, M. G. (1955). Rank correlation methods.

King, R. and Anderson, E. (2021). *skewt: The Skewed Student-t Distribution*. R package version 1.0.

Morris, S. A., Reich, B. J., Thibaud, E., and Cooley, D. (2017). A space-time skew-t model for threshold exceedances. *Biometrics*, 73(3):749–758.

Padoan, S. A. (2011). Multivariate extreme models based on underlying skew-t and skew-normal distributions. *Journal of Multivariate Analysis*, 102(5):977–991.

Sen, P. K. (1968). Estimates of the regression coefficient based on kendall's tau. *Journal of the American statistical association*, 63(324):1379–1389.

Sibuya, M. et al. (1960). Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics*, 11(2):195–210.