Bayesian Spatio-Temporal modelling of crime against women in Uttar Pradesh

Author CIDs: 02094608, 06027920, 06035613, 02075444

02/04/2025

Abstract

This project investigated spatio-temporal patterns of crimes against women, specifically rape and dowry deaths, in Uttar Pradesh, India, from 2001 to 2014. Bayesian hierarchical models (BHMs) using Besag–York–Mollié 2 (BYM2) priors and Type I interactions were applied to district-level data from the National Crime Records Bureau. Integrated Nested Laplace Approximation (INLA) facilitated efficient model fitting. Distinct spatial and temporal patterns were revealed: rape risk was the highest in the west and south, around Aligarh and Chitrakoot, respectively; while dowry deaths clustered in the southwest around Etah and Mainpuri. Rape incidents notably increased post-2012, likely influenced by enhanced reporting following the Delhi gang rape incident, whereas dowry death rates remained stable. Incorporating spatio-temporal interactions improved model performance, capturing localised district-specific trends. These differences reveal latent underpinning factors, possibly socio-cultural, suggesting targeted anti-dowry measures in hotspots and broader rape interventions statewide. Sensitivity analyses indicated that the models were robust to hyperprior specifications. Methodologically, we validated the research by Vicente et al. (2018 and 2020), highlighting the importance of Bayesian spatio-temporal analyses. Future research building on these findings can facilitate policy improvements, effective resource allocation and interventions addressing crimes against women in India.

Introduction

Violence against women remains a major public health and human rights issue in India. Rape and dowry deaths are among the gravest forms, reflecting deep-rooted gender inequality. Rape in particular is underreported due to social stigma and fear of retaliation (Shanmugam 2013). The dowry system, whereby a bride's family provides payment to the groom's family, can lead to dowry deaths when brides are murdered or driven to suicide through harassment and violence (Rastogi and Therly 2006). We look to Uttar Pradesh, India's most populous state, as a reference point for analysis.

Vicente et al (Vicente et al. 2018) first applied Bayesian Hierarchical Models (BHM) to the rape data, hypothesising a temporal effect from the widely publicised 2012 Delhi gang rape incident (The Hindu 2013). They then expanded their methods to dowry deaths (Vicente et al. 2020). Here, we build on that work by adopting Besag–York–Mollié 2 (BYM2) priors (Besag, York, and Mollié 1991), partitioning spatially structured and unstructured variation to validate and extend Vicente et al.'s conclusions. Employing BHMs allows us to explicitly model spatial and temporal correlation and obtain more precise risk estimates via "strength borrowing" across spatial and temporal structure. This approach is particularly valuable from a policy perspective, as it produces reliable district-level estimates even where crime reporting is sparse, facilitating more accurate resource allocation and potential targeted interventions.

Our spatiotemporal analyses of both crimes focus on the following questions:
1. Can we replicate Vicente et al.'s findings of distinct spatial patterns for rape and dowry deaths using BYM2 and Type I interactions?
2. How have the temporal trends for rape and dowry deaths evolved over the 14-year period, and what insights arise when examining the patterns of the two crimes in parallel?
3. Is there evidence of space-time interaction in the occurrence of these crimes, suggesting localised, rather than state-wide, spreading patterns?

By leveraging BHM techniques in this manner, we aim to produce more robust spatiotemporal insights that can guide targeted interventions and strengthen the literature on gender-based violence in India.

Methods

Data on rape and dowry deaths were sourced from the National Crime Records Bureau, covering 70 districts in Uttar Pradesh, North India, from 2001 to 2014 (National Crime Records Bureau 2017). Standardised morbidity ratios (SMRs) for rape and dowry deaths were provided in table form. A consistency check confirmed that the expected values in SMR calculations were correctly derived from aggregated statewide data, ensuring a clear internal baseline for comparisons across districts. Unsmoothed SMR data by district and year were mapped and graphed to visualise crude baseline spatial and temporal patterns as a starting point for our analysis.

BHMs were then employed to explicitly model spatial and temporal autocorrelation. Each model was fitted separately for each crime, using Integrated Nested Laplace Approximation (INLA) to leverage its computational simplifications (Rue, Martino, and Chopin 2009).

Three models were fitted: spatial (S), spatio-temporal (ST), and spatio-temporal with Interaction Type I (ST+I), specified as follows:

S Model

$$O_i \sim \text{Poisson}(\rho_i E_i)$$

$$\log(\rho_i) = b_0 + b_i$$

$$\mathbf{b} = \frac{1}{\sqrt{\tau_b}}\left(\sqrt{1-\phi}\,\mathbf{v}^* + \sqrt{\phi}\,\mathbf{u}^*\right)$$

ST Model

$$O_{it} \sim \text{Poisson}(\rho_{it} E_{it})$$

$$\log(\rho_{it}) = b_0 + b_i + \gamma_t$$

$$\mathbf{b} = \frac{1}{\sqrt{\tau_b}}\left(\sqrt{1-\phi}\,\mathbf{v}^* + \sqrt{\phi}\,\mathbf{u}^*\right)$$

$$\gamma_t \sim \text{RW}(1)$$

ST+I Model

$$O_{it} \sim \text{Poisson}(\rho_{it} E_{it})$$

$$\log(\rho_{it}) = b_0 + b_i + \gamma_t + \delta_{it}$$

$$\mathbf{b} = \frac{1}{\sqrt{\tau_b}}\left(\sqrt{1-\phi}\,\mathbf{v}^* + \sqrt{\phi}\,\mathbf{u}^*\right)$$

$$\gamma_t \sim \text{RW}(1)$$

$$\delta_{it} \sim \mathcal{N}\left(0, \sigma_\delta^2\right)$$

Where, for area i and time t, O and E denote the number of observed and expected cases, respectively, ρ is the relative risk of crime, and b_0 is the intercept term representing baseline log-risk. b_i is the spatial effect under BYM2 specification, decomposed into structured (u_i) and unstructured (v_i) components with a mixing parameter ϕ and overall precision τ_b. γ_t denotes the temporal component modelled as a first-random walk ($RW1$), and δ_{it} corresponds to a Type I interaction.

The BYM2 model was chosen for its decomposition of spatial variability into structured and unstructured components, which are easily controlled through the spatial mixing parameter ϕ (Riebler et al. 2016). This offers clearer interpretability and more flexible spatial smoothing compared to standard BYM or intrinsic conditional autoregressive (ICAR) models, which either lack explicit decomposition (ICAR) or provide limited control over component mixing (standard BYM).

Penalised Complexity (PC) priors were used for all precision and spatial mixing parameters to impose regularisation and avoid overfitting, defined as per Riebler et al's work. All precision priors were defined as $P(\sigma > 0.5) = 0.01$, reflecting a strong prior belief in smooth effects across space and time, motivated by our assumption that risk changes gradually due to continuity of sociocultural and administrative characteristics.

The PC prior for spatial mixing parameter ϕ was defined as $P(\phi < 0.5) = \frac{2}{3}$, reflecting our preference towards unstructured spatial variation. This is driven by the assumption that broader spatial trends, like systemic poverty and organised crime, exert greater influence on crime risk than district-specific factors (such as local crime reporting practices, governance effectiveness, opportunistic crime, or social norms).

To assess the robustness of our hyperprior choices, we conducted sensitivity analyses by implementing three Penalised Complexity (PC) prior configurations: Base_PC, Weaker_Shrinkage, and Stronger_Spatial, each varying in shrinkage strength and spatial structure preference, and compared model performance (WAIC) and hyperparameter summaries for both dowry death and rape models with Type I spatio-temporal interaction (please see Supplementary Table 2 for details).

All ST and ST+I models were assessed using the Watanabe-Akaike Information Criterion (WAIC), chosen for its ability to balance model fit with effective penalisation of model complexity in BHMs. The S model was excluded from WAIC comparison due to being fit on aggregate data.

Results

Between 2001 and 2014, a total of 25,513 rape cases (mean: 2,043 ± 334 per year) and 28,600 dowry deaths (mean: 1,822 ± 691 per year) were reported. Figure 1 shows the unsmoothed temporal trends and spatial variability of both crimes' SMRs.

All three models identified comparable patterns of posterior spatial RRs (Supplementary Figures 2-3). The ST+I model showed a better fit compared to the ST model, as reflected in its lower WAIC for both crimes (6240 vs 7781 for rape; 6318 vs 6654 for dowry deaths). Its results are therefore discussed further. Sensitivity analyses showed strong consistency in WAIC and posterior hyperparameter values across all ST+I model hyperprior specifications, indicating that posterior inferences were primarily driven by the data, and that the models are robust to prior assumptions (Supplementary Table 3).

Figure 2 shows the smoothed spatial posterior relative risks (RRs), as well as the posterior probabilities of exceeded risk (PPs). Both unsmoothed and smoothed estimates of dowry deaths maintained an east-to-west gradient, with high-risk districts identified around Etah and Mainpuri in the southwest. For rape, two high-risk clusters emerged around Aligarh in the west and Chitrakoot in the south. Smoothed RRs ranged wider for rape (0.3–2.3) than dowry deaths (0.4–2.1), corroborated by the lower precision parameter of the spatial component for rape (Table 1).

Figure 3 presents temporal trends estimated by the ST and ST+I models, both of which are similar to the unsmoothed RRs (Figure 1). The wider credible intervals in the ST+I model account for the district-specific variation captured by with the interaction term. Spatio-temporal interactions indicated district-specific temporal deviations, with clearer clustering noted for dowry deaths than for rape (available in Supplementary Figure 4).

Drawing on all the results above, Figure 4 shows the posterior mean RRs of the ST+I model, combining spatial, temporal, and spatio-temporal components, which capture geographical and temporal dynamics. Notable spatio-temporal patterns emerge: while dowry death risks remained relatively stable after 2004, rape risks experienced a rapid resurgence after 2012, with several districts exhibiting RRs of 2.5–4.75 in 2014.

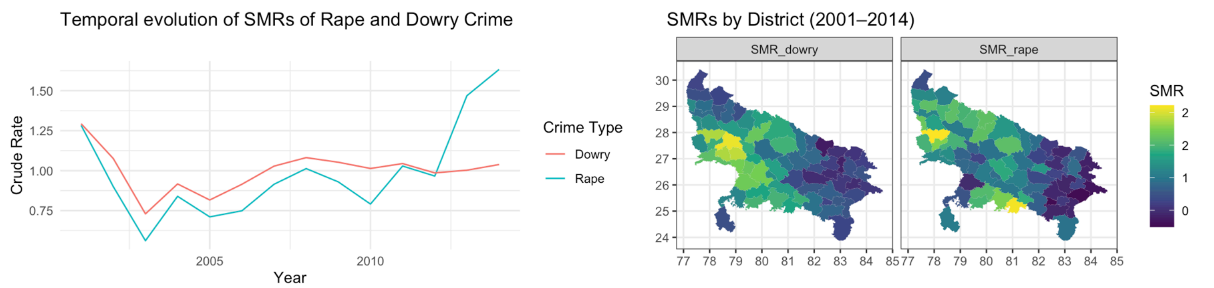


Figure 1: Temporal evolution of standardised SMRs for rape incidents and dowry deaths in Uttar Pradesh (2001-2014) (left) and spatial distribution of SMRs aggregated over the study period (right)

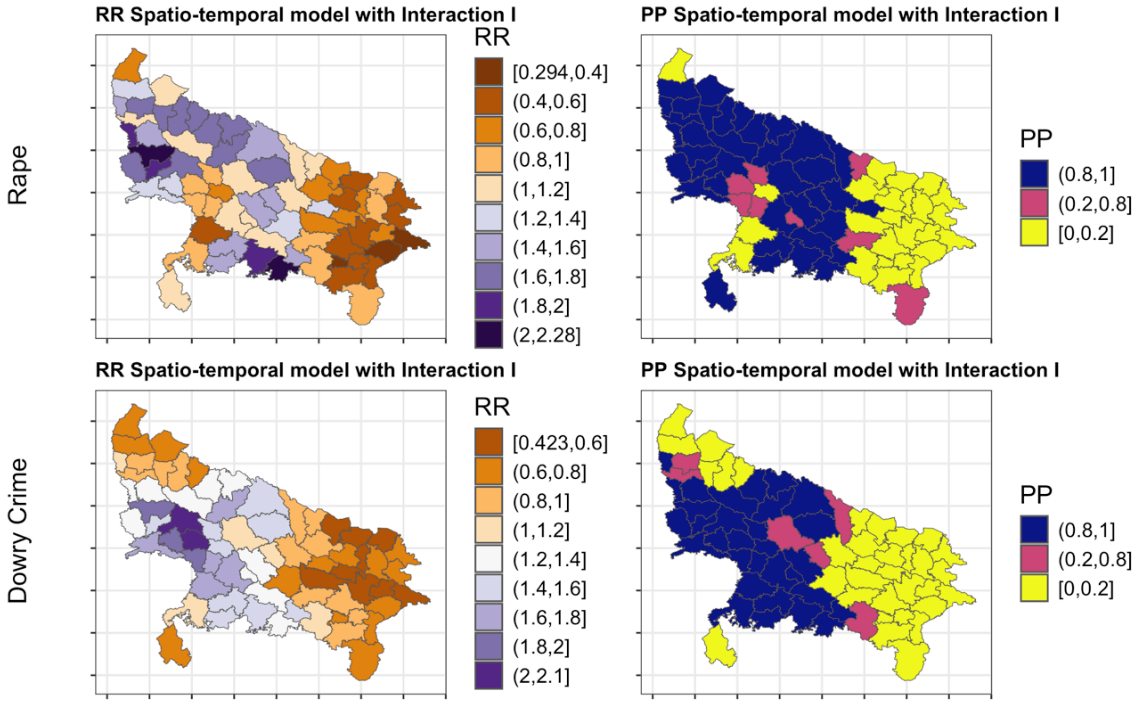


Figure 2: Spatial patterns of posterior mean relative risks (RR) and posterior probabilities (PP) exceeding 1 from spatio-temporal models with Type I interaction for rape incidents (top) and dowry deaths (bottom).

Table 1: Posterior Summaries of Hyperparameters

Parameter	Rape: Mean (95% CrI)	Dowry Deaths: Mean (95% CrI)
Precision for ID_area	5.91 (3.854, 8.538)	11.342 (7.555, 16.159)
Phi for ID_area	0.849 (0.562, 0.984)	0.894 (0.675, 0.99)
Precision for ID_year	15.088 (5.955, 30.72)	52.502 (19.606, 110.823)
Precision for ID_area_year	12.55 (10.765, 14.553)	43.712 (34.535, 54.928)

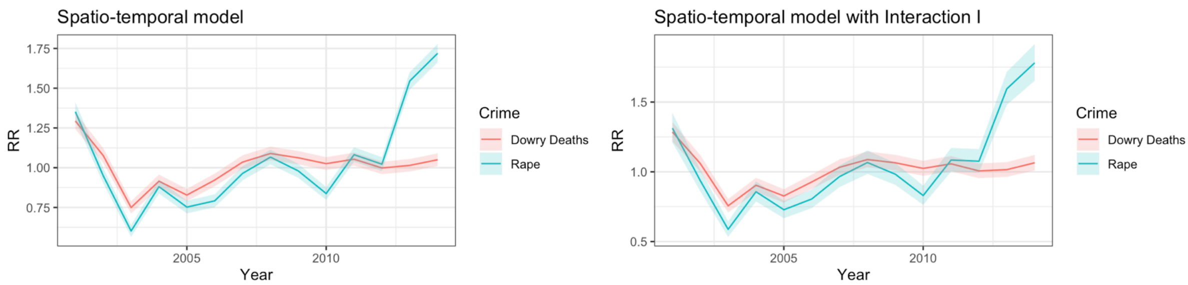


Figure 3: Comparison of temporal trends in relative risks between spatio-temporal models without interaction (left) and with Type I interaction (right) for rape incidents and dowry deaths in Uttar Pradesh (2001-2014).

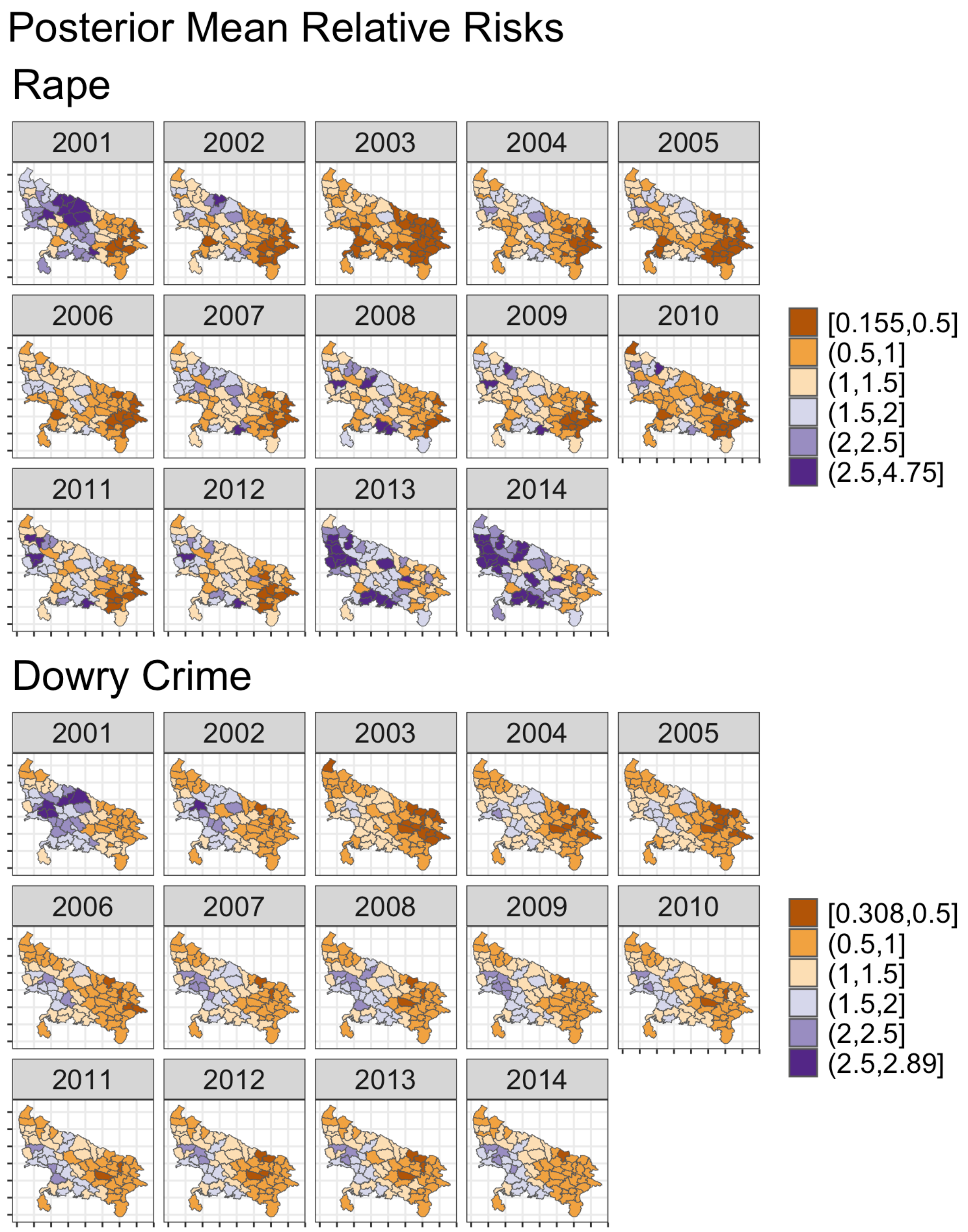


Figure 4: Annual posterior mean relative risks for rape incidents (top) and dowry deaths (bottom) across Uttar Pradesh districts from 2001 to 2014, demonstrating the evolution of spatial patterns over time.

Discussion

The most notable finding is the distinct temporal pattern observed between the two crimes: rape incidence showed a substantial increase especially after 2012, while dowry deaths remained relatively stable (Figure 3). Vicente et al. (2018) reported similar findings for rape, noting its temporal alignment with the highly publicised 2012 Delhi gang rape incident and subsequent 2013 legal reforms that expanded the definition of rape (Vicente et al. 2018) (The Hindu 2013). The rape models also exhibited a markedly lower precision hyperparameter of the temporal component, compared to that of dowry deaths (Table 1), suggesting that temporal effects played a more prominent role in shaping rape risk compared to dowry risk. Whilst our study cannot establish causation, the divergence in temporal trajectories of risk concurs that a sudden event affected rape incidence more than dowry deaths.

The high values of spatial mixing hyperparameter in both models (Table 1) point to the dominant contributions of structured spatial effects, consistent with observed spatial patterns. Rape reporting showed a clear west-to-east gradient which was more marked in the later years (Figure 4), with higher rates in the western districts neighbouring Delhi. This suggests that proximity to Delhi may have influenced reporting patterns, possibly through more effective awareness campaigns or stronger identification with the Delhi incident in western regions.

In contrast, the persistent clustering of dowry deaths around nearby southwestern districts like Etah and Mainpuri did not change with the 2012 incident. It is noted that marital rape was not included in the widening of the definition of rape in early 2013 (Hindustan Times 2013). While merely speculative, this potentially limited reporting by women experiencing chronic domestic abuse that may have ultimately resulted in dowry deaths. The more consistent spatial clustering observed for dowry deaths compared to rape incidents likely reflects different underlying socio-cultural dynamics. Dowry deaths may represent the culmination of prolonged domestic violence patterns tied to district-specific cultural factors, resulting in less heterogeneity compared to rape which may be more 'acute' and thus random in comparison.

We were not able to explain why both crimes showed an initial decline between 2001 to 2003, but infer that there were common statewide, or even national, factors affecting both crimes, in contrast to the post-2012 period.

The superior performance of the ST+I model highlights the importance of at least checking for localised temporal trajectories in crime data analysis. The assumption that all districts follow parallel temporal trends, implicit in non-interaction models (ST), may not capture the complex reality of violent crime.

From a policy perspective, our findings suggest different intervention approaches may be needed: targeted enforcement of anti-dowry laws in *specific* districts, versus broader educational and legal campaigns addressing rape *across the state* given its increased variability. The district-specific deviations identified for rape could *then* inform to what degree localised prevention strategies should build on statewide efforts. Further research into the spatial Delhi 'spillover' effect may bring interesting insights in helping to develop these interventions.

This study was limited by the absence of district-level covariates constraining interpretation and the study period ending in 2014. Future research should incorporate socioeconomic variables, for example internet coverage, explore different interaction models (II-IV), and conduct qualitative studies to better understand district-level variations in crime reporting and responses.

In conclusion, this spatio-temporal analysis demonstrates the value of advanced Bayesian modelling techniques for understanding the complex dynamics of gender-based violence. We validated Vicente et al.'s findings of rape and dowry deaths' spatiotemporal patterns and demonstrated that they were robust to hyperprior variation (Vicente et al. 2018) (Vicente et al. 2020). We detailed the temporal trends of both crimes and showed the value of adding an interaction term in unveiling how different districts varied individually with time, while taking smoothing into account. The distinct spatial patterns, temporal trends, and interaction effects observed for rape and dowry deaths underscore the importance of these models to aid crime-specific, geographically targeted interventions to address violence against women in Uttar Pradesh.

Supplementary material

Please note: the analysis code can be found in a separate file: Group09_SM.html and Group09_SM.Rmd.

Supplementary Table 1: Yearly summary of crude crime rates (per 100,000 women)

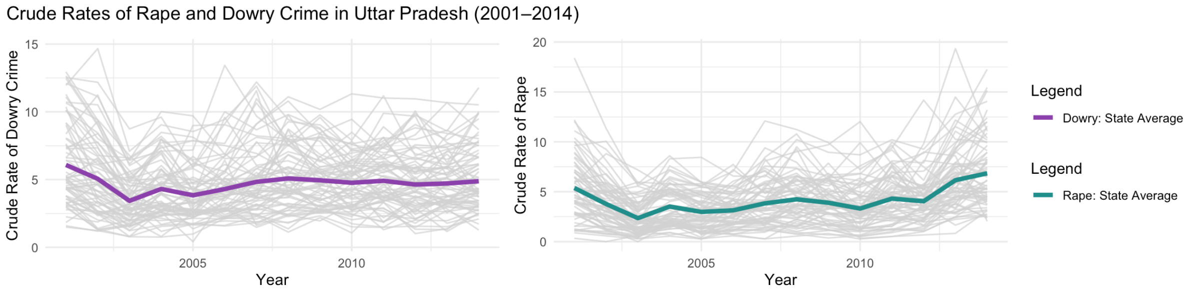
Year	Rape	Dowry	Higest Rape	Lowest Rape	Highest Dowry	Lowest Dowry
2001	5.4	6.1	Pilibhit (18.4)	Sant Ravidas Nagar Bhadohi (0.3)	Kheri (12.9)	Sant Ravidas Nagar Bhadohi (1.5)
2002	3.8	5.1	Pilibhit (11.3)	Sant Ravidas Nagar Bhadohi (0)	Etah (14.7)	Sant Ravidas Nagar Bhadohi (1.2)
2003	2.4	3.4	Hathras (7.2)	Sant Kabir Nagar (0)	Mainpuri (8.7)	Sant Kabir Nagar (0.8)
2004	3.5	4.3	Sitapur (8.6)	Ghazipur (0.6)	Firozabad (10)	Ghazipur (0.8)
2005	3.0	3.8	Aligarh (8.4)	Ghazipur (0.3)	Firozabad (9.7)	Ghazipur (0.4)
2006	3.1	4.3	Gautam Buddha Nagar (7.7)	Sant Ravidas Nagar Bhadohi (0.6)	Kanpur Dehat (13.4)	Sant Ravidas Nagar Bhadohi (1.6)
2007	3.8	4.8	Chitrakoot (12.1)	Deoria (0.3)	Mainpuri (12.2)	Deoria (1.5)
2008	4.2	5.1	Chitrakoot (11.2)	Sant Ravidas Nagar Bhadohi (0.6)	Aligarh (11.1)	Sant Ravidas Nagar Bhadohi (1.1)
2009	3.9	4.9	Chitrakoot (9.9)	Ghazipur (0.4)	Aligarh (10.2)	Ghazipur (1.9)
2010	3.3	4.8	Rampur (12)	Sant Ravidas Nagar Bhadohi (0.3)	Aligarh (11.3)	Sant Ravidas Nagar Bhadohi (1.1)
2011	4.3	4.9	Aligarh (10.2)	Sant Ravidas Nagar Bhadohi (0.5)	Etawah (11.2)	Sant Ravidas Nagar Bhadohi (1.3)
2012	4.1	4.6	Chitrakoot (14.2)	Mirzapur (0.7)	Aligarh (11)	Mirzapur (1)
2013	6.2	4.7	Hamirpur (19.3)	Mirzapur (0.8)	Firozabad (10.7)	Mirzapur (1.2)
2014	6.8	4.9	Banda (17.3)	Shrawasti (2.1)	Etah (11.8)	Shrawasti (1.3)

Supplementary Table 2: Hyperprior Settings for Sensitivity Analysis

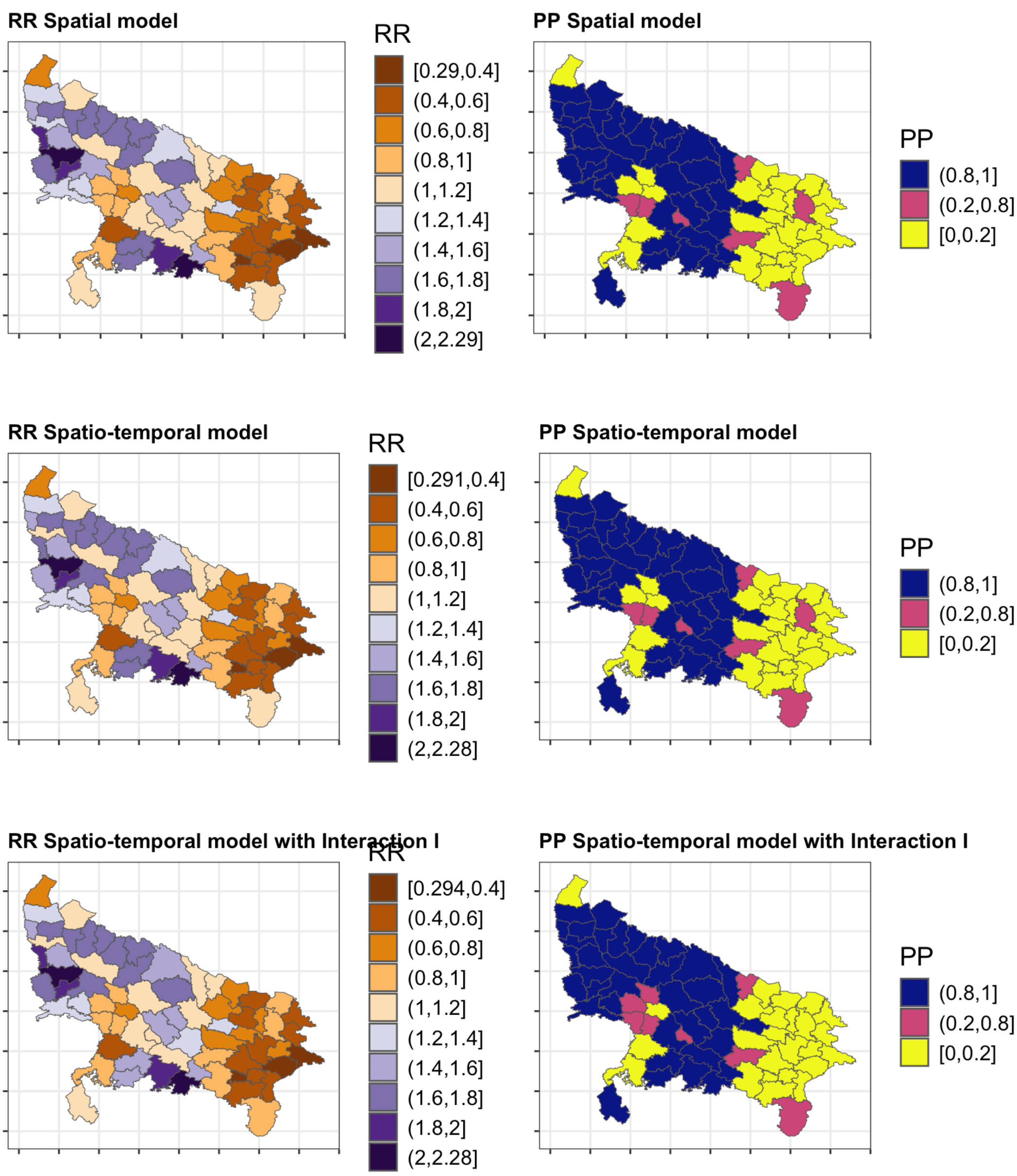
Model	U	Alpha	Phi	Description
Base PC	0.31	0.01	2/3	Original PC prior
Weaker Shrinkage	0.50	0.01	0.5	Weaker shrinkage: allows more variance, favors less spatial structure
Stronger Spatial	0.30	0.05	0.8	Stronger spatial structure: tighter σ, favors more spatial structure

Supplementary Table 3: Sensitivity Analysis Results for Spatio-temporal Models with Interaction I

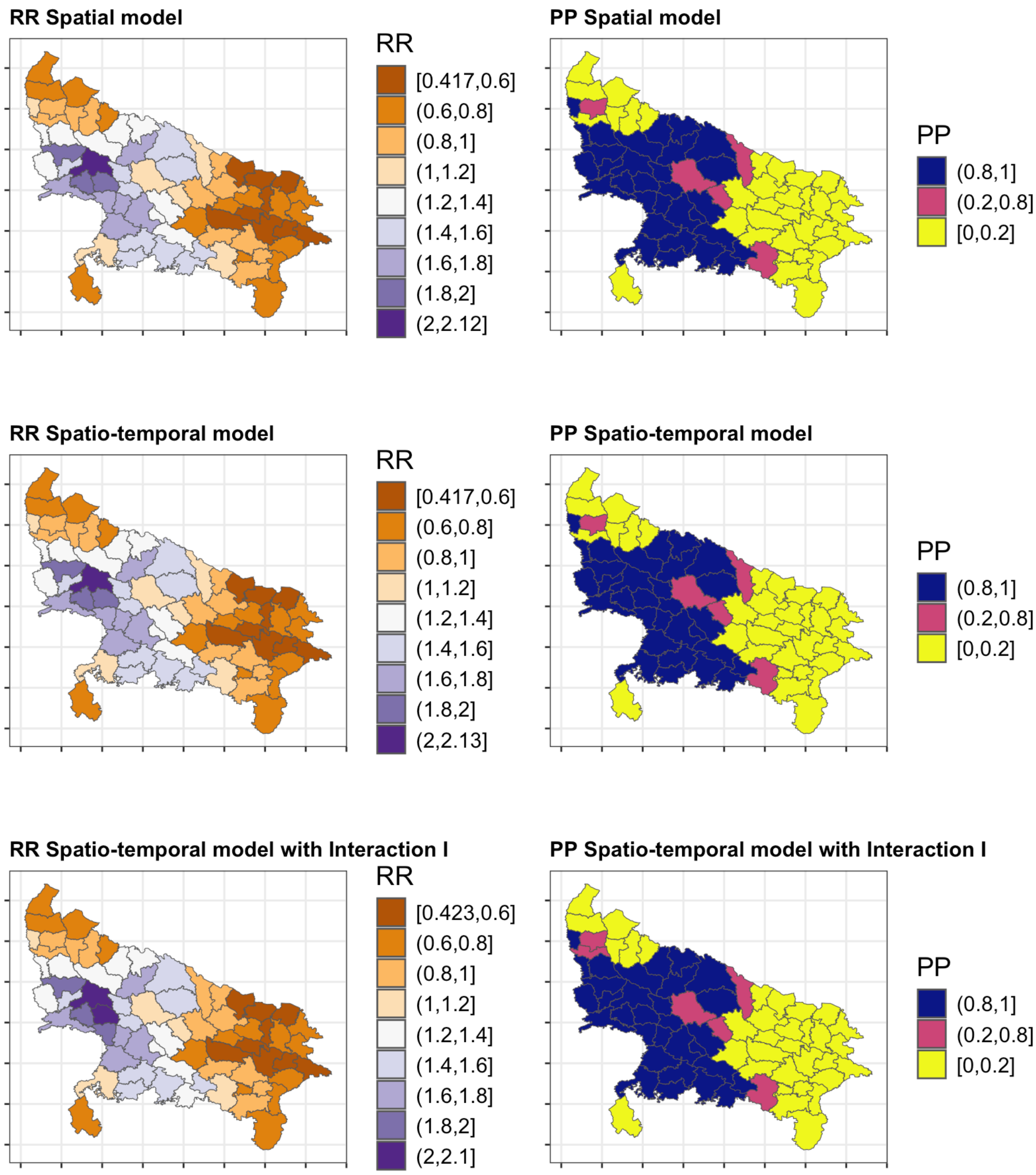
Crime	Model	DIC	WAIC	Prec_ID_area	Phi_ID_area	Prec_ID_year	Prec_ID_area_year
dowry	Base PC	6328.356	6318.483	11.341832	0.8940197	52.50050	43.71170
dowry	Weaker Shrinkage	6328.314	6318.680	11.443379	0.8987401	53.86492	43.78907
dowry	Stronger Spatial	6328.437	6318.467	11.417036	0.8870633	51.71217	43.66984
rape	Base PC	6303.217	6240.220	5.913548	0.8490889	15.09256	12.55015
rape	Weaker Shrinkage	6303.171	6240.441	5.956673	0.8558258	15.67756	12.56806
rape	Stronger Spatial	6303.328	6240.245	5.909798	0.8370025	14.72932	12.53894



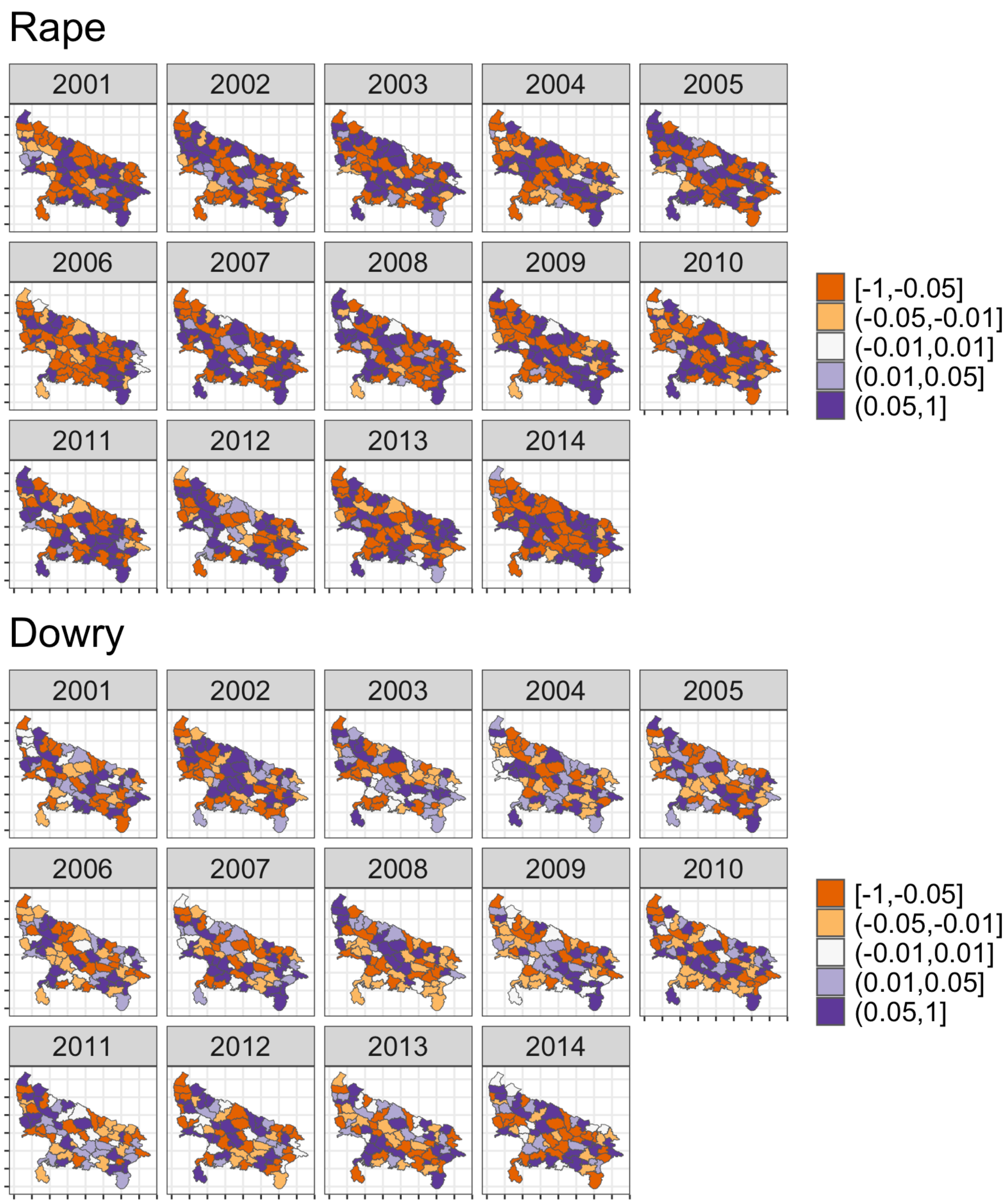
Supplementary Figure 1: Crude rates of rape incidents and dowry deaths in Uttar Pradesh (2001-2014), showing individual district trends (gray lines) and state average trends (colored lines).



Supplementary Figure 2: Comparison of spatial patterns in relative risks (RR) and posterior probabilities (PP) across three models for rape incidents: spatial-only model (top), spatio-temporal model without interaction (middle), and spatio-temporal model with Type I interaction (bottom).



Supplementary Figure 3: Comparison of spatial patterns in relative risks (RR) and posterior probabilities (PP) across three models for dowry deaths: spatial-only model (top), spatio-temporal model without interaction (middle), and spatio-temporal model with Type I interaction (bottom).



Supplementary Figure 4: Space-time interaction effects for rape incidents (top) and dowry deaths (bottom) in Uttar Pradesh districts from 2001 to 2014, highlighting areas with unusual local temporal trends.

References

Besag, J., J. C. York, and A. Mollié. 1991. "Bayesian Image Restoration, with Two Applications in Spatial Statistics (with Discussion)." *Annals of the Institute of Statistical Mathematics* 43: 1–59.

Hindustan Times. 2013. "Lok Sabha Passes Anti-Rape Bill." https://web.archive.org/web/20130319233614/http://www.hindustantimes.com/India-news/NewDelhi/Lok-Sabha-passes-anti-rape-bill/Article1-1028961.aspx (https://web.archive.org/web/20130319233614/http://www.hindustantimes.com/India-news/NewDelhi/Lok-Sabha-passes-anti-rape-bill/Article1-1028961.aspx).

National Crime Records Bureau. 2017. "Crime in India 2017." New Delhi: Ministry of Home Affairs, Government of India.

Rastogi, Mudita, and Paul Therly. 2006. "Dowry and Its Link to Violence Against Women in India: Feminist Psychological Perspectives." *Trauma, Violence, & Abuse* 7 (1): 66–77. https://doi.org/10.1177/1524838005283927 (https://doi.org/10.1177/1524838005283927).

Riebler, Andrea, Sigrunn H Sørbye, Daniel Simpson, and Håvard Rue. 2016. "An Intuitive Bayesian Spatial Model for Disease Mapping That Accounts for Scaling." *Statistical Methods in Medical Research* 25 (4): 1145–65. https://doi.org/10.1177/0962280216660421 (https://doi.org/10.1177/0962280216660421).

Rue, H., S. Martino, and N. Chopin. 2009. "Approximate Bayesian Inference for Latent Gaussian Model by Using Integrated Nested Laplace Approximations (with Discussion)." *J. R. Statist. Soc. B* 71: 319–92.

Shanmugam, Ramalingam. 2013. "Informatics about Fear to Report Rapes Using Bumped-up Poisson Model." *Current Research in Biostatistics* 3 (November): 17–29. https://doi.org/10.3844/amjbsp.2013.17.29 (https://doi.org/10.3844/amjbsp.2013.17.29).

The Hindu. 2013. "Delhi Gangrape: Chronology of Events." https://www.thehindu.com/news/national/delhi-gang-rape-chronology-of-events/article11862316.ece (https://www.thehindu.com/news/national/delhi-gang-rape-chronology-of-events/article11862316.ece).

Vicente, G., T. Goicoa, P. Fernandez-Rasines, and M. D. Ugarte. 2020. "Crime Against Women in India: Unveiling Spatial Patterns and Temporal Trends of Dowry Deaths in the Districts of Uttar Pradesh." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 183 (2): 655–79. https://www.researchgate.net/publication/362156364_Vicente_G_Goicoa_T_Fernandez-Rasines_P_Ugarte_M_D_2020_Crime_against_women_in_India_unveiling_spatial_patterns_and_temporal_trends_of_dowry_deaths_in_the_districts_of_Uttar_P (https://www.researchgate.net/publication/362156364_Vicente_G_Goicoa_T_Fernandez-Rasines_P_Ugarte_M_D_2020_Crime_against_women_in_India_unveiling_spatial_patterns_and_temporal_trends_of_dowry_deaths_in_the_districts_of_Uttar_P

Vicente, G., T. Goicoa, A. Puranik, and M. D. Ugarte. 2018. “Small Area Estimation of Gender-Based Violence: Rape Incidence Risks in Uttar Pradesh, India.” *Statistics and Applications* 16 (1): 71–90. https://www.researchgate.net/publication/326669154 (https://www.researchgate.net/publication/326669154).