Exploratory analysis of accesses to support centers for gender-based violence in Apulia

Data

The dataset employed regards the counts of accesses to gender-based violence support centers in the Apulia region by residence municipality of the women victims of violence in 2021-2023. `R` codes to generate the dataset are in the R script posted here which this report is based on.

Here, we only take into account the violence reports which support centers actually take charge of, at the risk of underestimating the counts of gender-based violence cases. This choice is driven by the need of avoiding duplicated records, since e.g. it may happen that a support center redirects a victim to another support center.

In order to avoid singletons, i.e. municipalities with no neighbours, in the spatial structure of the dataset, the Tremiti Islands need to be removed from the list of municipalities included (0 accesses recorded so far).

Therefore, the municipality-level dataset in scope consists of 256 observations.

We can only take into account the accesses to support centers for which the origin municipality of victims is reported; therefore the total count of accesses in scope is 1477, 1516 and 1822 for the three reference years respectively:

```
dd %>% sf::st_drop_geometry() %>%
  dplyr::group_by(.data$Year) %>%
  dplyr::summarise(Tot_accesses = sum(.data$N_ACC))
```

```
## # A tibble: 3 x 2
##    Year Tot_accesses
##   <dbl>        <dbl>
## 1     1         1477
## 2     2         1516
## 3     3         1822
```

Here, we plot the log-access rate per residence municipality, i.e. the logarithm of the ratio between access counts and female population. Blank areas correspond to municipalities from which zero women accessed support centers (82 municipalities).

Covariates

Our target is explaining the number of accesses to support centers, y, defined at the municipality level, on the basis of a set of candidate known variables. y is modelled with simple Poisson regression.

We have at disposal a number of candidate explanatory variables, which include the distance of a municipality from the closest support center and a set of variables measuring social vulnerability under different dimensions; these latter covariates are provided by the ISTAT and are among the components of the Municipality Frailty Index (MFI). A more detailed description of MFI components is in this excel metadata file.

All covariates are quantitative variables and to ease model interpretation are scaled to have null mean and unit variance.

- TEP_th, i.e. the distance of each municipality from the closest municipality hosting a support center. Distance is measured by road travel time in minutes (acronym TEP stays for Tempo Effettivo di

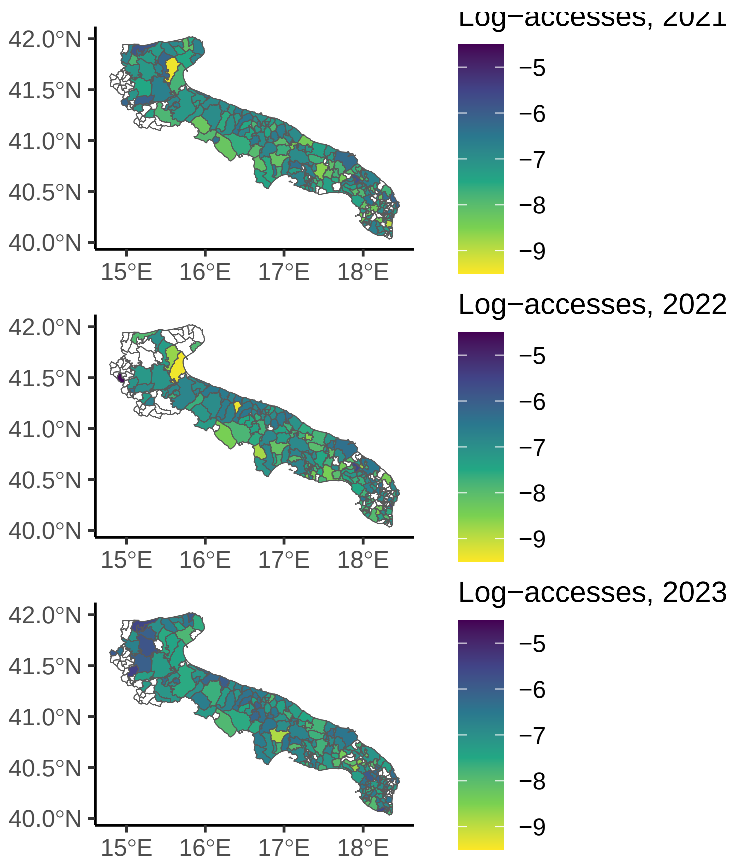


Figure 1: Log-access rate

Percorrenza, i.e. Actual Travel Time). Since to the best of our knowledge the list of active support centers changed between 2022 and 2023, we employ the list of centers active until 2022 for 2021-2022 data, and the list of centers active in 2023 for 2023 data.

- AES, the distance from the closest infrastructural pole, always measured in travel time. Infrastructural poles are defined as municipalities or clusters of neighbouring municipalities provided with a certain endowment in health, education and transport infrastructure (more details can be found, e.g., here).
- MFI, i.e. the decile of municipality frailty index.
- PDI, i.e. the dependency index, i.e. population either ≤ 20 or ≥ 65 years over population in $[20-64]$ years.
- ELL, i.e. the proportion of people aged $[25-54]$ with low education.
- ERR, i.e. employment rate among people aged $[20-64]$.
- PGR, i.e. population growth rate with respect to 2011.
- UIS, i.e. the ventile of the density of local units of industry and services (where density is defined as the ratio between the counts of industrial units and population).
- ELI, i.e. the ventile of employees in low productivity local units by sector for industry and services.

First, we visualise the correlations among these explanatory variables:

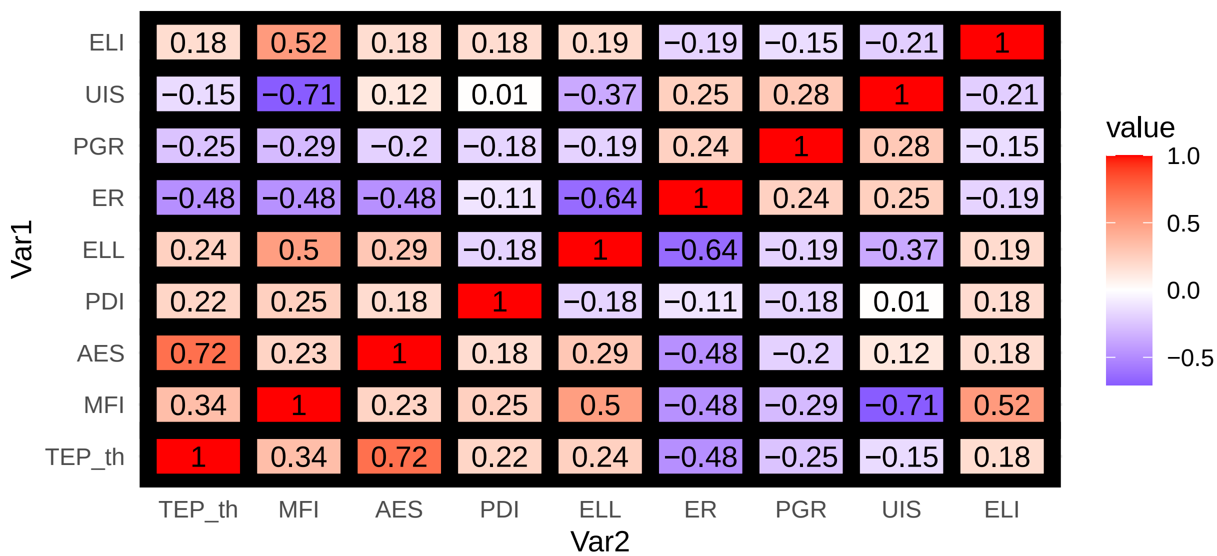


Figure 2: Correlations in explanatory variables

We see the correlation between the two distances is very high (0.72), and so is the correlation between the fragility index decile and the density of productive units.

In the first case, we drop the distance from the nearest infrastructural pole. In the latter we drop `MFI`, which is a combination of all covariates except for `TEP_th`, and is a weakly informative choice.

Nonspatial regression

We regress the counts of accesses y to support centers on the aforementioned explanatory variables. To interpret regression coefficients in an easier way, all covariates are scaled to zero mean an unit variance.

$$y_{it} \mid \eta_{it} \sim \text{Poisson}(E_{it}\, e^{\eta_{it}}) \quad \text{where} \quad \eta_{it} = X_{it}^{\top}\alpha \tag{1}$$

Where X are the covariates defined earlier, α are covariate effects, and E_{it} is the female population aged ≥ 15 in municipality i and year t.

To gain more insight on the role of all explanatory variables we show the posterior summaries of the regression model; the posterior distribution is approximated with the INLA (**INLA?**).

% latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:32:42 2025

Effect	Mean	Sd	0.025quant	0.975quant
Int_2021	-7.343	0.032	-7.405	-7.281
Int_2022	-7.315	0.031	-7.376	-7.254
Int_2023	-7.139	0.030	-7.197	-7.080
TEP_th	-0.256	0.020	-0.296	-0.217
ELI	-0.058	0.018	-0.093	-0.023
PGR	0.033	0.023	-0.012	0.079
UIS	0.013	0.019	-0.024	0.050
ELL	-0.161	0.024	-0.209	-0.114
PDI	-0.064	0.024	-0.112	-0.017
ER	-0.221	0.026	-0.272	-0.169

- `TEP_th_22`: The distance from the closest support center appears to play an important role. The easiest interpretation is that the physical distance represents a barrier to violence reporting. This is quite intuitive if we think of the material dynamics of reporting gender-based violence: one could reasonably expect violent men to prevent their partners to take a long rout and come out to report the violence suffered.
- `ELI`: The (ventile of the distribution of the) share of employees in low productivity economic units is a clear indicator of (relative) economic underdevelopment. The most naive interpretation would be that in underdeveloped areas reporting gender violence is somewhat harder than in developed ones.
- `PGR`: The association with population growth rate does not appear to be significantly different from zero.
- `UIS`: The (ventile of the distribution of the) density of production units does not appear to have a significantly $\neq 0$ association with VAW reporting.
- `ELL`: The association with the proportion of people with low educational level has negative sign and is high in absolute value. The interpretation seems quite easy: cultural development, in general, would encourage reporting violence.
- `PDI`: The association with population dependency index is negative, and its interpretation is ambiguous as well since the proportion of old or very young people can be read as a limiting factor both for VAW incidence and reporting.
- `ER`: The association with employment rate is very strong and bears negative sign for 2021 and 2023 data.

Spatial regression

Exploratory analysis of residuals We plot the log-residuals ε of the nonspatial regression models, defined as $\varepsilon := \ln y_{it} - \ln \hat{y}_{it}$ being $\hat{y}_{it}$ the fitted value.

Residuals may exhibit spatial structure. To assess it, we employ the Moran and Geary tests. Since

Please notice that log-residuals only take finite values across the municipalities whose female citizens have reported at least one case of violence in 2022.

Additionally, this set of municipalities may include some singletons, which we remove to assess the value of the Moran and Geary statistics. Thus, for each year we have defined the indexes set `nonzero_con` as the set of municipalities from which at least one case of gender-based violence has been reported, *and* which have

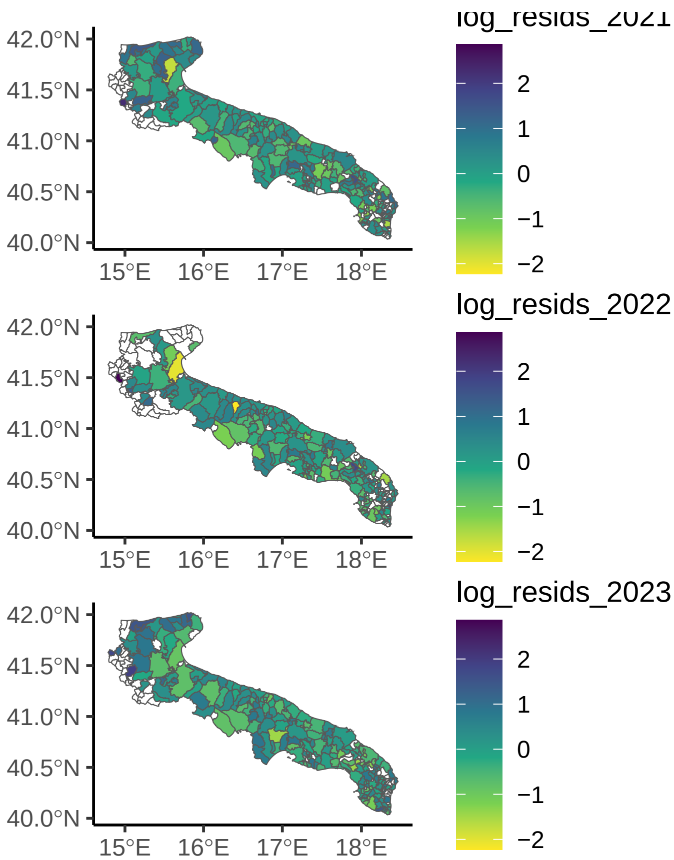


Figure 3: Log-residuals in GLM regression

at least one neighbouring municipality from which at least one case of gender-based violence was reported as well. For brevity, we only show the standardised I values, which under the null hypothesis should be distributed as $N(0, 1)$. The Geary's test is also included for completeness.

% latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:32:45 2025

Year	Test	Statistic_std	p.value
2021	Moran	2.04	0.02
2021	Geary	2.48	0.01
2022	Moran	1.91	0.03
2022	Geary	2.14	0.02
2023	Moran	5.00	0.00
2023	Geary	5.05	0.00

We find evidence for spatial autocorrelation. However, we must stress out that this result does not refer to all the regional territory, but only to a subset of all municipalities.

Based on the autocorrelation evidence, though it has only been assessed for a subset of all municipalities, we try implementing some simple spatial models by adding a conditionally autoregressive latent effect, say z, to the linear predictor

$$\eta_{it} = X_{it}^{\top}\alpha + z_{it} \tag{2}$$

We test a total of four models, all of which have a prior distribution depending on the spatial structure of the underlying graph, in this case the Apulia region.

In the following, the area-specific latent field is denoted as $z_i = (z_{i,2021}\, z_{i,2022}\, z_{i,2023})^{\top}$

We describe the spatial structure starting from municipalities neighbourhood, and introduce the neighbourhood matrix W, whose generic element w_{ij} takes value 1 if municipalities i and j are neighbours and 0 otherwise. For each $i \in [1, n]$, $d_i := \sum_{j=1}^{n} w_{ij}$ is the number of neigbours of i-th municipality. Please notice we have have $n = 256$.

For all models, we define Λ as the precision parameter of the latent effect, and assign it a Wishart prior. Additionally, $\Sigma = \Lambda^{-1}$ is the covariance matrix of z, and its off-diagonal entries describe dependence in z between different years.

Spatial models are computed by approximating the marginal posteriors of interest via the Integrated Nested Laplace Approximation (INLA), adopting the novel Variational Bayes Approach (Van Niekerk et al. 2023).

Priors for spatial effects have been defined using the `INLAMSM` `R` package (Palmí-Perales, Gómez-Rubio, and Martinez-Beneito 2021).

ICAR model The Intrinsic CAR model is the simplest formulation among spatial autoregressive models. The conditional distribution of each value $z_i \mid z_{-i}$ is:

$$z_i \mid z_{-i} \sim \mathcal{N}\left(\sum_{j=1}^{n} \frac{w_{ij}}{d_i} z_j\, ,\, \frac{1}{d_i}\Lambda^{-1}\right) \tag{3}$$

And the joint prior distribution is:

$$z \mid \Sigma \sim \mathcal{N}\left(0, \Sigma \otimes (D - W)^{+}\right) \tag{4}$$

Where z is a 256×3 matrix. Since the joint distribution of z is improper, a sum-to-zero constraint, i.e. $\sum_{i=1}^{n} z_{it} = 0$ for $t = 2021, 2022, 2023$ is required for identifiability.

PCAR model The intrinsic autoregressive model is relatively simple to interpret and to implement, while also requiring the minimum number of additional parameters (either the scale or the precision).

The drawback, however, is that we implicitly assume a deterministic spatial autocorrelation coefficient equal to 1. When the autocorrelation is weak, setting an ICAR prior may be a form of misspecification.

A generalisation of this model is the PCAR (proper CAR), which introduces an autocorrelation parameter ρ:

$$z_i \mid z_{-i} \sim \mathcal{N}\left(\sum_{j=1}^{n} \rho \frac{w_{ij}}{d_i} z_j \, , \frac{1}{d_i} \Lambda^{-1}\right) \tag{5}$$

We show the posterior summary for the autocorrelation coefficient, obtained with the data at hand.

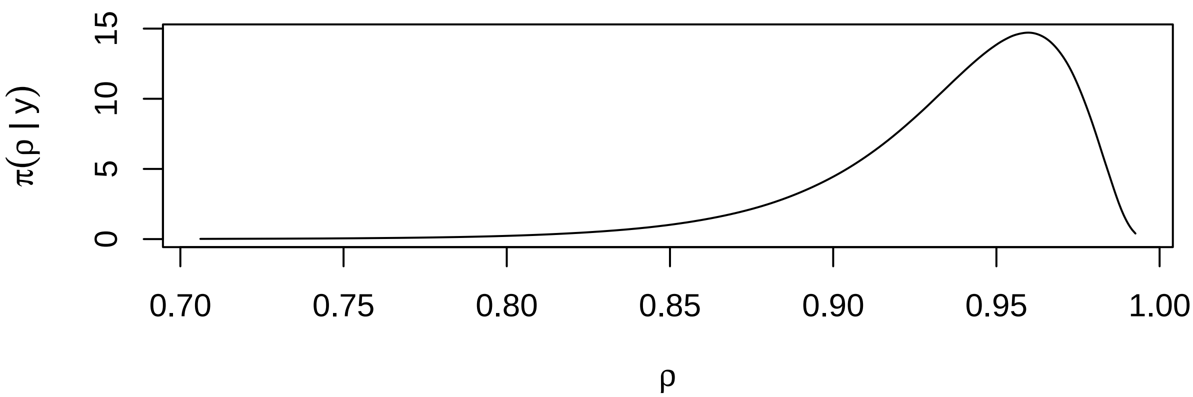


Figure 4: Posterior marginal of the PCAR autocorrelation parameter

```
##       mean         sd quant0.025  quant0.25   quant0.5  quant0.75 quant0.975
## 0.93595762 0.03700831 0.84101638 0.91847248 0.94414543 0.96244829 0.98280419
```

The credible interval for ρ is quite pushed towards unity, denoting the model estimates a strong spatial autocorrelation.

Leroux model As an alternative to take into account both structured and unstructured latent effects, we also test the Leroux autoregressive model (Leroux, Lei, and Breslow 2000); throughout this report, this model will be referred to as the LCAR. In this case, the local prior for z_i is

$$z_i \mid z_{-i} \sim \mathcal{N}\left(\sum_{j=1}^{n} \frac{\xi w_{ij}}{1-\xi+\xi d_i} z_j \, , \Lambda^{-1} \frac{1}{1-\xi+\xi d_i}\right) \tag{6}$$

Where $\xi \in [0, 1]$ is labelled as the mixing parameter. A more interesting representation of the Leroux model is the joint prior

$$z \mid \Lambda, \xi \sim \mathcal{N}(0, [\Lambda \otimes (\xi L + (1-\xi)I)]^{-1})$$

where $L := D - W$ is the graph Laplacian matrix, W is the neighbourhood matrix and D is the corresponding degree matrix. We can clearly see how the mixing parameter allocates variability between two precision components, i.e. the Laplacian matrix for the spatial part and the the identity matrix for the noise.

The drawback of this model is the scarce interpretability with respect to more sophisticated ones like the BYM, but the multivariate framework complicates the definition of the BYM model (and it does not seem to be possible to reparametrise it in such a way to have a sparse precision, hence computations are unfeasible) and hampers the use of PC priors, which would have indeed been a useful tool to prevent overfitting.

BYM model Another popular model to control for both spatial autocorrelation and random noise is the Besag, York and Mollié model. We employ a simplified formulation, with a unique mixing parameter for all three years. Under this model the latent effect is defined as:

$$z = \sqrt{\phi}uM + \sqrt{1-\phi}vM \tag{7}$$

Where:

- $u \sim \mathcal{N}(0, I_p \otimes L_{\text{scaled}})$ is an independent multivariate ICAR process whose precision matrix is scaled in order that the geometric mean of marginal variances (i.e. the diagonal entries of L^+) is one
- $v \sim \mathcal{N}(0, I_p \otimes I_n)$ is a Standard Normal variable
- M is a positive definite matrix such that $M'M = \Lambda^{-1}$. It is not necessarily the Cholesky factor of the scale parameter; in fact, a convenient but not unique way to define it may be $M = D^{-\frac{1}{2}}E^\top$ where E and D are the eigenvector and eigenvalues matrices of Λ (Urdangarin et al. 2024).
- $\phi \in [0, 1]$ is the mixing parameter.

Please notice that in the case of our data, $p = 3$ and $n = 256$. We assign a Uniform prior on ϕ (but the PC-prior would be a more rigorous choice) and the usual Wishart prior to Λ.

The BYM model features a mixing parameter as the LCAR does, but improves its interpretability by allowing to scale the ICAR and IID components.

We use a rather simplified BYM definition here. It could be improved in many directions: either defining a PC prior on the mixing parameter and employing a more elegant parametrisation allowing for sparse precision (Riebler et al. 2016), or relaxing the naive hypothesis of a unique mixing parameter for all the $p = 3$ years implementing an M-model.

```
## Time needed for scaling Laplacian matrix:  1.601  seconds
```

Model assessment We briefly compare the three models in scope through the WAIC (Gelman, Hwang, and Vehtari 2014):

% latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:34:15 2025

Model	WAIC	Eff_params
Null	3476.400	39.218
ICAR	2904.624	184.391
PCAR	2903.529	194.421
Leroux	2901.132	193.919
BYM	2886.894	189.783

As we can see, adding a spatial model is an improving element, not a waste of complexity.

Here we show some posterior summaries for α under the BYM model.

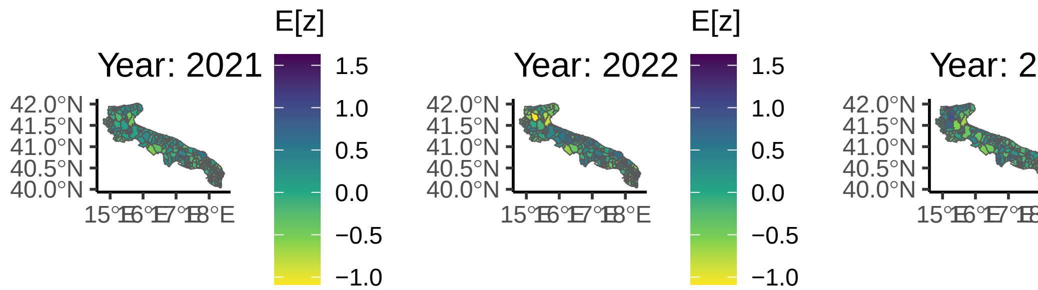
% latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:34:16 2025

Estimations of α differ slightly from the nonspatial model and credibility intervals are generally wider due to increased uncertainty.

- `TEP_th_22` The effect of the distance from the closest support center remains similar in mean and the interpretation is not altered.

Effect	Mean	Sd	Q_0.025	Q_0.975
Int_2021	-7.411	0.044	-7.498	-7.327
Int_2022	-7.540	0.051	-7.643	-7.442
Int_2023	-7.184	0.042	-7.268	-7.102
TEP_th	-0.227	0.043	-0.313	-0.142
ELI	-0.026	0.038	-0.100	0.049
PGR	0.070	0.043	-0.014	0.155
UIS	0.017	0.040	-0.061	0.095
ELL	-0.233	0.050	-0.331	-0.136
PDI	-0.067	0.045	-0.156	0.021
ER	-0.273	0.057	-0.384	-0.162

- `ELI`: The effect of the incidence of low-productivity economic units is utterly negligible
- `PGR`: The association with population growth rate can only be considered barely significant for 2022 data
- `UIS`: The association with the density of productive units is negligible for 2021 and 2022 data, and can be considered slightly significant for 2023, bearing positive sign.
- `ELL`: The association with the incidence of low education levels, is even higher in mean than under the nonspatial model. We interpret this result as a strong *potential* impact of education on the chance that gender violence is reported
- `PDI`: The effect of structural dependency index is utterly negligible, as for the GLM.
- `ER`: The effect associated with employment rate is increased for 2021 data, more than doubled for 2022 data, and slightly increased for 2023 data. How to interpret this finding? Employment rate is clearly an indicator of economic development, hence the easiest interpretation is that `-` as it was with `ELI` under the nonspatial model `-` in more developed areas there is a higher chance that gender violence is reported.



We show the expected latent effects
As we can see, the behaviour of 2023 data is quite puzzling, since high values of the latent effect are observed in areas such as the Northern province of Foggia, which was previously characterised by low accesses.

We additionally show the fitted values of the counts of accesses, rounded to the closest integer.

We show the posterior median of the marginal variances of z, i.e. the diagonal entries of Σ:

For more insight on the dependence structure between the three years, posterior summaries for correlations are shown in the following. In terms of point estimates, the only high correlation appears that of 2021-2022. % latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:34:18 2025

Years	Mean	Sd	Q0.025	Median	Q0.975
2021-22	0.714	0.086	0.518	0.724	0.853
2021-23	0.335	0.129	0.064	0.341	0.567
2022-23	0.207	0.127	-0.049	0.210	0.445

Lastly, we have a look at the mixing parameter under both the Leroux and BYM models.

```
##       mean        sd quant0.025  quant0.25   quant0.5  quant0.75 quant0.975
##  0.7469941  0.1104227  0.4934809  0.6792042  0.7621042  0.8300082  0.9163707
```

As we see, the Leroux mixing parameter is not particularly high, but interpreting it properly is hampered by the difficulty in scaling the precision matrix, which is a drawback of non-intrinsic models.

```
##       mean        sd quant0.025  quant0.25   quant0.5  quant0.75 quant0.975
## 0.87332597 0.07475783 0.68104577 0.83713897 0.89052852 0.92752281 0.96771786
```

M-models Until now, we have been assuming the hyperparameter controlling for the strength spatial association is constant across the three years. We try to relax this hypothesis by allowing for year-specific

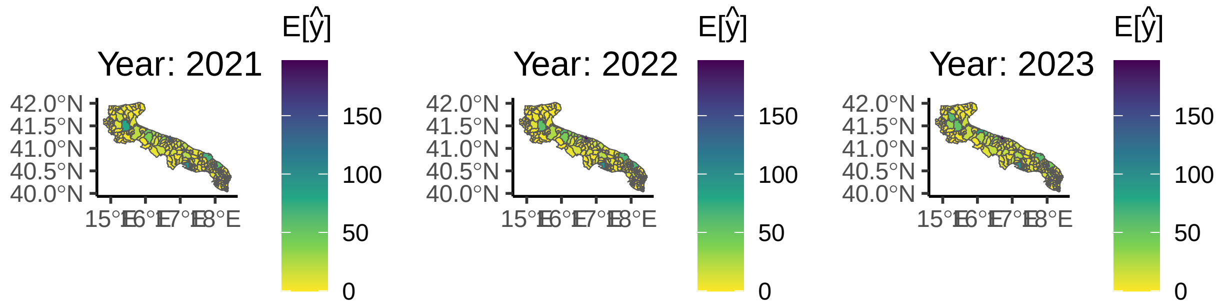


Figure 5: Fitted access counts, BYM model

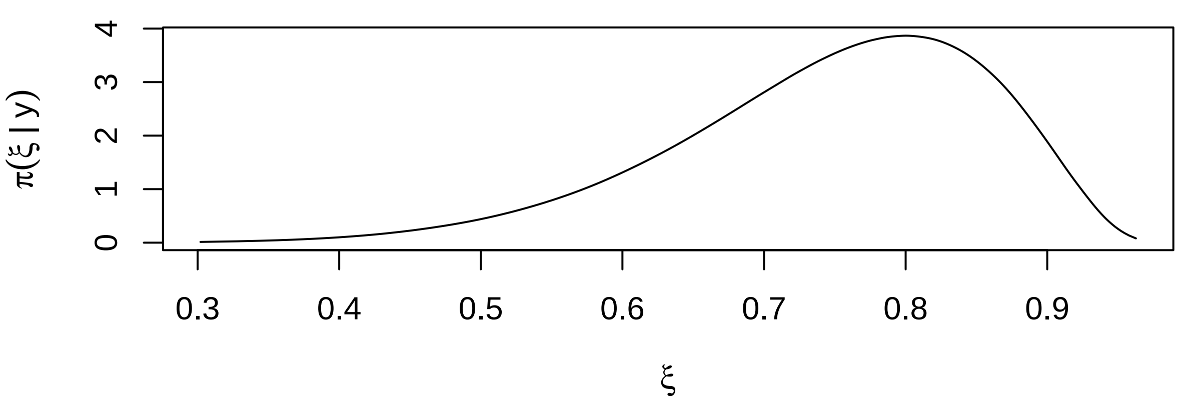


Figure 6: Posterior marginal of the LCAR mixing parameter

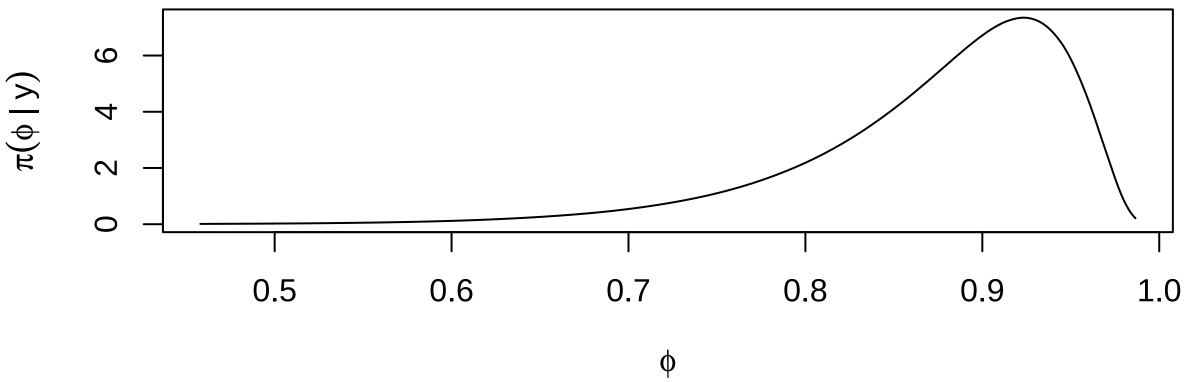


Figure 7: Posterior marginal of the BYM mixing parameter

hyperparameters, which can be achieved by using a more flexible class of multivariate models, i.e. the M-models (Botella-Rocamora, Martinez-Beneito, and Banerjee 2015).

Consider the scale parameter Σ, i.e. the marginal variance of the latent effect z, and let us define the definite-positive matrix M, such that $M\top M = \Sigma$.

Starting with the PCAR model, if we allow the autocorrelation parameter ρ to change over time, we introduce the matrix-valued parameter $R := \text{diag}(\rho_{2021}, \rho_{2022}, \rho_{2023})$ and the model becomes thus:

$$z \mid \Sigma, R \sim \mathcal{N}\left(0, (M^{\top} \otimes I_n)[(I_p \otimes D) - (R \otimes W)]^{-1}\right) \tag{8}$$

Where $n = 256$ and $p = 3$, and the remainder of notation is equal to previous sections. Since the matrix M has dimension p^2 and is not symmetric, modelling it directly would become burdensome. We therefore opt to make directly inference on Σ. In line with (Vicente et al. 2023), we employ the Bartlett factorisation to Σ, which briefly means that we set $\Sigma := AA^{\top}$, where A is a lower-triangular matrix whose off-diagonal elements are assigned a Normal prior and the squares of its diagonal are assigned a χ^2 distribution. This is equivalent to the assumption that Σ follows a Wishart prior; for a proof see e.g. (Kabe 1964).

For the time being, we employ the `bigDM` R package(Vicente et al. 2023), which implements the M-model formulation of both the PCAR and LCAR models. Autoregressive parameters are assumed to be independent and follow a Uniform prior.

Likewise, we test the LCAR model, which under the M-model formulation features a matrix-valued mixing parameter $\Xi := \text{diag}(\xi_{2021}, \xi_{2022}, \xi_{2023}$

$$z \mid \Sigma, R \sim \mathcal{N}\left(0, (M^{\top} \otimes I_n)[\Xi \otimes (D - W) + (I_p - \Xi) \otimes I_n]^{-1}\right) \tag{9}$$

For the sake of comparison, we also test the ICAR model with Bartlett decomposition of the scale parameter - which is equivalent to the ICAR discussed in the previous section, with the only difference that now it is the scale parameter and not the precision which is assumed to follow a Wishart distribution.

Though it is a simple model, `INLA` seems to have some crashes. It will need further investigation.

Lastly, we propose the M-model extension of the BYM.

If we define the matrix-valued mixing parameter$\Phi := \text{diag}(\phi_1, \phi_2, \dots \phi_p)$ and $\bar{\Phi} := \text{diag}(1-\phi_1, 1-\phi_2, \dots 1-\phi_p) = (I_p - \Phi)$, with $p = 3$, the BYM field $\mathbf{Y}$ is given by:

$$Z = U\Phi^{\frac{1}{2}}M + V\bar{\Phi}^{\frac{1}{2}}M \tag{10}$$

Where

- $U = (U_1, U_2, \dots U_p)$ is a multivariate and independent ICAR field, such that $U_j \sim \mathcal{N}_n(0, L^+)$ for $j = 1, 2 \dots p$, and L^+ is the pseudoinverse of the scaled graph Laplacian matrix
- $V = (V_1, V_2, \dots V_p)$ is a collection of *iid* random Normal vectors, such that $V_j \sim \mathcal{N}_n(0, I_n)$ for $j = 1, 2 \dots p$.
- M is a generic $p \times p$ full-rank matrix such that $M^\top M = \Sigma$, being Σ is the scale parameter. For instance, M can be defined as $D^{1/2}E^\top$, where D is the diagonal matrix of the eigenvalues of Σ and E is the corresponding eigenvectors matrix.

```
Mmodel.bym2.bartlett.sparse <-
 function (cmd = c("graph", "Q", "mu", "initial", "log.norm.const",
                   "log.prior", "quit"), theta = NULL) {

   envir <- parent.env(environment())
   if(!exists("cache.done", envir=envir)){
     starttime.scale <- Sys.time()
     #' Laplacian matrix scaling: only needs being done once
     L_unscaled <- Matrix::Diagonal(n=nrow(W), x=rowSums(as.matrix(W))) -  W
     constr <- INLA:::inla.bym.constr.internal(L_unscaled, adjust.for.con.comp = T)
     scaleQ <- INLA:::inla.scale.model.internal(
       L_unscaled, constr = list(A = constr$constr$A, e = constr$constr$e))
     L <- scaleQ$Q
     endtime.scale <- Sys.time()
     cat("Time needed for scaling Laplacian matrix: ",
         round(difftime(endtime.scale, starttime.scale), 3), " seconds \n")
     assign("L", L, envir = envir)
     assign("cache.done", TRUE, envir = envir)
   }
   interpret.theta <- function() {
     #' Same as the dense version, plus the M-factorisation of the scale parameter
     #' like in the M-models
     alpha <-  1/(1+exp(-theta[as.integer(1:J)]))

     diag.N <- sapply(theta[J+1:J], function(x) {exp(x)})
     no.diag.N <- theta[2*J+1:(J*(J-1)/2)]

     N <- diag(diag.N)
     N[lower.tri(N)] <- no.diag.N

     Sigma <- N %*% t(N)
     e <- eigen(Sigma)
     #M <- t(e$vectors %*% diag(sqrt(e$values)))
     invM.t <- diag(1/sqrt(e$values)) %*% t(e$vectors)
     Lambda <- t(invM.t) %*% invM.t
     return(list(alpha = alpha, Lambda = Lambda, invM.t = invM.t))
   }
   graph <- function() {
     QQ <- Q()
```

```r
    G <- (QQ != 0) * 1
    return(G)
  }

  Q <- function() {
    param <- interpret.theta()

    In <- Matrix::Diagonal(nrow(W), 1)

    q11 <- t(param$invM.t) %*% Matrix::Diagonal(x = 1/(1 - param$alpha), n = J) %*% param$invM.t
    q12 <- t(param$invM.t) %*% Matrix::Diagonal(x = sqrt(param$alpha)/(1 - param$alpha), n = J)
    q22 <- Matrix::Diagonal(x = param$alpha/(1 - param$alpha), n = J)

    Q11 <- kronecker(q11, In)
    Q12 <- - kronecker(q12, In)
    Q21 <- Matrix::t(Q12)
    Q22 <- kronecker(q22, In) + kronecker(Matrix::Diagonal(x=1, n=J), L)
    #' Structure matrix. Problem: it is of size 2n * 2n
    Q <- rbind(cbind(Q11, Q12), cbind(Q21, Q22))
    return(Q)
  }
  mu <- function() {
    return(numeric(0))
  }
  log.norm.const <- function() {
    val <- numeric(0)
    return(val)
  }
  log.prior <- function() {
    #' As in the dense version
    param <- interpret.theta()
    val <-  sum(-theta[1:J] - 2*log(1+exp(-theta[1:J])))
    # n^2_jj ~ chisq(J-j+1) (J degrees of freedom)
    val <- val + J*log(2) + 2*sum(theta[J+1:J]) + sum(dchisq(exp(2*theta[J+1:J]), df=J-1:J+1, log=TRUE
    # n_ji ~ N(0,1)
    val <- val + sum(dnorm(theta[(2*J)+1:(J*(J-1)/2)], mean=0, sd=1, log=TRUE))
    return(val)
  }
  initial <- function(){
    if(!exists("initial.values", envir= envir )){
      return(c(rep(0, J*(J+3)/2)))
    } else {
      return(initial.values)
    }
  }
  quit <- function() {
    return(invisible())
  }
  if (as.integer(R.version$major) > 3) {
    if (!length(theta))  theta <- initial()
  }
  else {
    if (is.null(theta))  theta <- initial()
```

```r
    }
    val <- do.call(match.arg(cmd), args = list())
    return(val)
  }

inla.BYM.sparse <- function(...) INLA::inla.rgeneric.define(Mmodel.bym2.bartlett.sparse, ...)
constr.BYM <- list(
        A = cbind(Matrix::Matrix(0, nrow = nrow(A.constr), ncol = ncol(A.constr)), A.constr),
        e=c(0,0,0) )

cav_MMBYM_bigDM <- inla(
  N_ACC ~ 0 +TEP_th + ELI + PGR + UIS + ELL + PDI + ER+
    f(Year, model = "iid", hyper = list(prec = list(initial = 1e-3, fixed = TRUE)))  +
    f(ID, model = inla.BYM.sparse(J = 3, W = W_con),
      extraconstr = constr.BYM),
  offset = log(nn), family = "poisson", data =dd,
  num.threads = 1, control.compute = list(internal.opt = F, cpo = T, waic = T, config = T),
  verbose = T)
```

```
## Time needed for scaling Laplacian matrix:  0.277  seconds
```

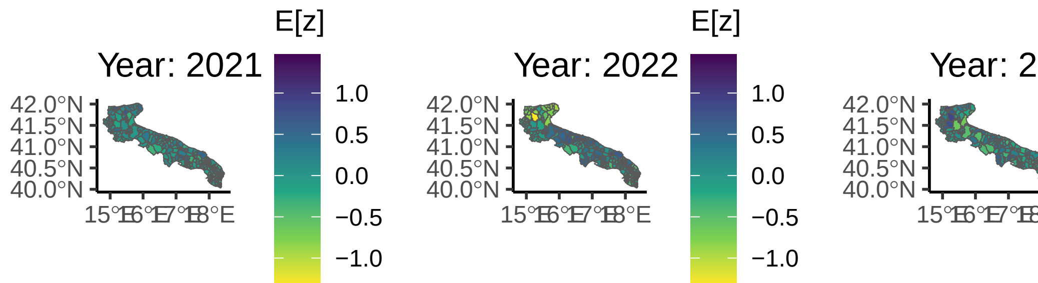
Model comparison follows: % latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:36:20 2025

Model	WAIC	Eff_params
Null	3476.400	39.218
ICAR	2957.145	207.601
PCAR	2944.676	213.661
Leroux	2949.196	210.703
BYM	2937.613	204.358

Here we show some posterior summaries for α under the BYM model.

% latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:36:20 2025

Effect	Mean	Sd	Q_0.025	Q_0.975
Int_2021	-7.375	0.047	-7.467	-7.284
Int_2022	-7.562	0.057	-7.675	-7.450
Int_2023	-7.154	0.047	-7.247	-7.063
TEP_th	-0.237	0.035	-0.305	-0.168
ELI	-0.038	0.031	-0.099	0.024
PGR	0.073	0.036	0.002	0.144
UIS	0.010	0.031	-0.052	0.071
ELL	-0.216	0.041	-0.297	-0.135
PDI	-0.040	0.039	-0.116	0.035
ER	-0.284	0.046	-0.374	-0.194



We show the expected latent effects
As we can see, the behaviour of 2023 data is quite puzzling, since high values of the latent effect are observed in areas such as the Northern province of Foggia, which was previously characterised by low accesses.

We additionally show the fitted values of the counts of accesses, rounded to the closest integer.

Lastly, we have a look at the autoregressive (PCAR) and mixing (LCAR, BYM) parameters.

```
rho_marg_PMMCAR <- lapply(c(1,2,3), function(x){
  inla.tmarginal(fun = function(y) {1/(1 + exp(-y))},
                 marginal = cav_PMMCAR_bigDM$marginals.hyperpar[[x]])
})



xi_marg_LMMCAR <- lapply(c(1,2,3), function(x){
  inla.tmarginal(fun = function(y) {1/(1 + exp(-y))},
                 marginal = cav_LMMCAR_bigDM$marginals.hyperpar[[x]])
})

phi_marg_MMBYM <- lapply(c(1,2,3), function(x){
  inla.tmarginal(fun = function(y) {1/(1 + exp(-y))},
                 marginal = cav_MMBYM_bigDM$marginals.hyperpar[[x]])
})

# PCAR autoregressive
sapply(c(1,2,3), function(x){
  inla.zmarginal(rho_marg_PMMCAR[[x]], silent = T)
})
```

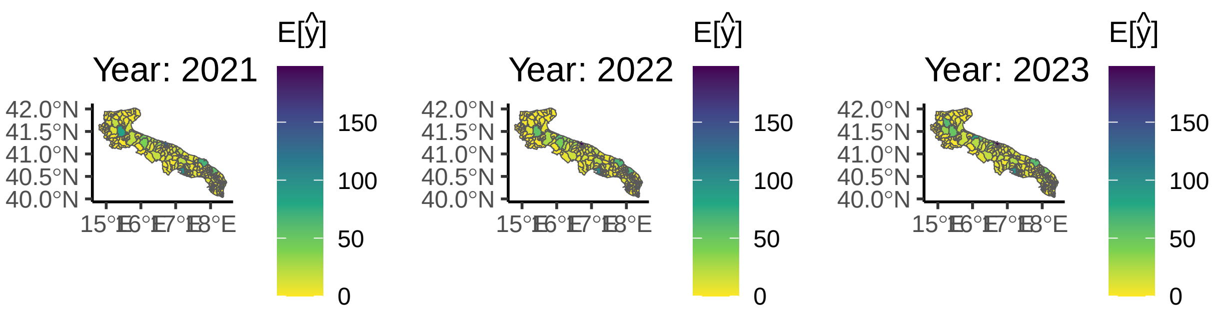


Figure 8: Fitted access counts, BYM M-model

```
##              [,1]        [,2]        [,3]
## mean        0.9591674   0.8828074   0.6258671
## sd          0.002498022 0.007397905 0.01574413
## quant0.025 0.9538409   0.8675676   0.5915828
## quant0.25  0.9575581   0.8779251   0.6157802
## quant0.5   0.9593294   0.8830535   0.6273316
## quant0.75  0.9609213   0.8879282   0.637183
## quant0.975 0.9636336   0.8966096   0.6528276
```

```r
sapply(c(1,2,3), function(x){
  inla.zmarginal(xi_marg_LMMCAR[[x]], silent = T)
})
```

```
##              [,1]      [,2]      [,3]
## mean        0.6259352 0.8254702 0.6601466
## sd          0.1713084 0.1161335 0.156204
## quant0.025 0.2659272 0.5241647 0.3215496
## quant0.25  0.5066661 0.7685793 0.5553054
## quant0.5   0.6418016 0.8547562 0.6772017
## quant0.75  0.7595791 0.9113251 0.7810158
## quant0.975 0.9037107 0.9666613 0.9082578
```

```r
sapply(c(1,2,3), function(x){
  inla.zmarginal(phi_marg_MMBYM[[x]], silent = T)
})
```

```
##              [,1]      [,2]      [,3]
## mean        0.658369  0.7766761 0.4866354
## sd          0.2391751 0.1386787 0.1768709
## quant0.025 0.1550905 0.4529655 0.1842684
## quant0.25  0.4822009 0.6905949 0.3495496
## quant0.5   0.7007158 0.8009372 0.4731403
## quant0.75  0.8648757 0.8867567 0.6157849
## quant0.975 0.9793301 0.9680029 0.8385767
```

Appendix: sparse parametrisation of the BYM Recall the BYM as defined in equation 10; the prior variance of Z is:

$$\mathrm{Var}\left[\mathrm{vec}(Z) \mid \Sigma, \Phi\right] = \left(M^\top \otimes I_n\right) \mathrm{diag}\left(S_1, \dots, S_p\right)\left(M \otimes I_n\right)$$

Where each diagonal block S_j is the variance of $U_j\Phi + V_j\bar{\Phi}$, i.e. $S_j = \phi_j L^+ + (1-\phi_j)I_n$.

Extending the approach of (Riebler et al. 2016) to the multivariate case, we know that:

$$\mathbb{E}[\mathrm{vec}(Z) \mid U, \Phi, \Sigma] = \mathrm{vec}(U\Phi^{\frac{1}{2}}M) = [(M^\top\Phi^{\frac{1}{2}}) \otimes I_n]\mathrm{vec}(U)$$

and similarly

$$\mathrm{Var}[\mathrm{vec}(Z) \mid U, \Phi, \Sigma] = \left[(M^\top\bar{\Phi}^{\frac{1}{2}}) \otimes I_n\right] \mathbb{E}\left[\mathrm{vec}(V)\mathrm{vec}(V)^\top\right] \left[(\bar{\Phi}^{\frac{1}{2}}M) \otimes I_n\right] = \left(M^\top\bar{\Phi}M\right) \otimes I_n$$

The distribution of $Z \mid U, \Sigma, \Phi$ then reads:

$$
\begin{aligned}
-2 \ln \pi \left(\operatorname{vec}(Z) \mid U, \Sigma, \phi\right) \operatorname{vec}(Z)^\top & \left[\left(M^{-1} \bar{\Phi}^{-1} {M^{-1}}^\top\right) \otimes I_n\right] \operatorname{vec}(Z) \\
-2\operatorname{vec}(Z)^\top & \left[\left(M^{-1} \bar{\Phi}^{-1} {M^{-1}}^\top\right) \otimes I_n\right] \left[\left(M^\top \Phi^{\frac{1}{2}}\right) \otimes I_n\right] \operatorname{vec}(U) \\
+\operatorname{vec}(U)^\top \left[\left(\Phi^{\frac{1}{2}} M\right) \otimes I_n\right] & \left[\left(M^{-1} \bar{\Phi}^{-1} {M^{-1}}^\top\right) \otimes I_n\right] \left[(M^\top \Phi^{\frac{1}{2}}) \otimes I_n\right] \operatorname{vec}(U) \\
= C + \operatorname{vec}(Z)^\top & \left[\left(M^{-1} \bar{\Phi}^{-1} {M^{-1}}^\top\right) \otimes I_n\right] \operatorname{vec}(Z) \\
-2\operatorname{vec}(Z)^\top & \left[\left(M^{-1} \bar{\Phi}^{-1} \Phi^{\frac{1}{2}} M^\top\right) \otimes I_n\right] \operatorname{vec}(U) \\
+\operatorname{vec}(U)^\top & \left[\left(\Phi \bar{\Phi}^{-1}\right) \otimes I_n\right] \operatorname{vec}(U)
\end{aligned}
$$

Now, for brevity let us define the following $p \times p$ matrices:

$$
q_{11} := M^{-1} \bar{\Phi}^{-1} {M^{-1}}^\top; \quad q_{12} := M^{-1} \bar{\Phi}^{-1} \Phi^{\frac{1}{2}}; \quad q_{22} := \Phi \bar{\Phi}^{-1}
$$

Hence

$$
\begin{aligned}
-2 \ln \pi \left(\operatorname{vec}(Z) \mid U, \Sigma, \Phi\right) = C + \operatorname{vec}(Z)^\top & \left(q_{11} \otimes I_n\right) \operatorname{vec}(Z) \\
-2\operatorname{vec}(Z)^\top & \left(q_{12} \otimes I_n\right) \operatorname{vec}(U) \\
+\operatorname{vec}(U)^\top & \left(q_{22} \otimes I_n\right) \operatorname{vec}(U)
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
-2 \ln \pi \left(\operatorname{vec}(Z), \operatorname{vec}(U) \mid \Sigma, \phi\right) = C + \operatorname{vec}(Z)^\top & \left(q_{11} \otimes I_n\right) \operatorname{vec}(Z) \\
-2\operatorname{vec}(Z)^\top & \left(q_{12} \otimes I_n\right) \operatorname{vec}(U) \\
+\operatorname{vec}(U)^\top & \left(q_{12} \otimes I_n + I_p \otimes L\right) \operatorname{vec}(U)
\end{aligned}
$$

Hence, with some straightforward algebra, it can be concluded that:

$$
\begin{pmatrix} \operatorname{vec}(Z) \\ \operatorname{vec}(U) \end{pmatrix} \sim \mathcal{N}_{2np} \left(0, \begin{pmatrix} q_{11} \otimes I_n & -q_{12} \otimes I_n \\ -q_{12}^\top \otimes I_n & q_{22} \otimes I_n + I_p \otimes L \end{pmatrix}^{-1} \right) \tag{11}
$$

Which generalises to the multivariate case the sparse precision derived by (Riebler et al. 2016).

Further development: pogit model for underreporting

The observed counts of accesses to AVCs clearly imply a severe under-reporting of gender-based violence, regarding the vast majority of incidents.

If this does not provide a valid reason to model the under-reporting process explicitly - as there are not any data to assess the spatial distribution of true VAW incindents counts at *this* level of granularity, it is still needful to consider that some risk factors with a strong negative association have same-sign effects on accesses counts.

More specifically, employment rate and low-education incidence have a very high negative correlation, as it is reasonable to expect. Notwithstanding this, both appear to have a negative impact on access counts. A possible explanation for this finding is that these two variables impact on two different processes, which we do not observe, but whose interaction yields to the observed response variable, namely accesses counts.

We therefore attempt at implementing a Poisson - logistic model (Stoner, Economou, and Silva 2019) to explicitly account for under-reporting or, to be more precise, to model the propensity to report violence by turning to AVCs or not.

Given the scarce information at our disposal, we can only guess which variable enter either of the two predictor components, i.e. the VAW reporting and the actual occurrence Specifically, we assign distance from closest AVC and low education incidence to the former component and all other variables plus the spatial effect to the latter.

The model is thus defined:

$$y_{i,t} \sim \text{Poisson}(E_{it}\, \lambda_{it}\, \pi_{it}) \quad \forall i \in [1, 256], t \in \{2021, 2022, 2023\} \tag{12}$$

Where $E_{i,t}$ is the offset, λ_{it} is the expected count of incidents π_{it} is the expected reporting rate. We will denote α as the $k+1$ covariate effects entering λ_{it} and β the $m+1$ covariate effects entering π_{it} for two integers k and m.

We would thus have, for each year t and municipality i;

$$\lambda_{it} = e^{\alpha_{0t} + \alpha_1 X_{i1,t} + ... + \alpha_k X_{ik,t} + z_{it}}$$

And

$$\pi_{it} = \left[1 + e^{-(\beta_0 + \beta_1 X_{k+1,t} + ... + \beta_m X_{i(k+m),t})}\right]^{-1}$$

Hence the predictor is not linear anymore.

We approximate the predictor through first-order Taylor expansion, which is made possible by the `inlabru` `R` package (Lindgren et al. 2024). To define the nonlinear predictor component, we follow (Wøllo 2022).

Here, we show the `R` code we propose to achieve this. Due to the computational burden, for the time being we are not testing M-models.

```
logexpit <- function(x, beta){
  pred <- beta[[1]] +
    rowSums(do.call(cbind, lapply(c(1:length(x)), function(n){
      beta[[n+1]]*x[[n]]
      })))
  return(-log(1+exp(-pred)))
}

cmp_spatial <- function(spatial_expr, st_expr = NULL, t_expr = NULL) {

  spatial <- function(id, ...) INLA::f(id, ...)
  spatiotemporal <- function(id, ...) INLA::f(id, ...)
  intercept <- function(id, ...) INLA::f(id, ...)

  f2 <- substitute(spatial_expr)
  f3 <- substitute(t_expr)
  f4 <- substitute(st_expr)
  f1 <- bquote(
    ~ 0 +
      beta_0(main = 1,  model = "linear",
                  mean.linear = -2.2,
                  prec.linear = 1e+2) +
      myoffset(log(nn), model = "offset") +
      alpha_ELI(ELI) +
      alpha_PGR(PGR) +
      alpha_UIS(UIS) +
      alpha_PDI(PDI) +
      alpha_ER(ER) +
      beta_TEP(main = 1, model = "linear",
```

```r
               mean.linear = 0,
               prec.linear = 1e-3) +
      beta_ELL(main = 1, model = "linear",
               mean.linear = 0,
               prec.linear = 1e-3) )
  ff <- f1[[2]]
  if (!is.null(f2)) ff <- as.call(c(quote(`+`), ff, f2))
  if (!is.null(f3)) ff <- as.call(c(quote(`+`), ff, f3))
  if (!is.null(f4)) ff <- as.call(c(quote(`+`), ff, f4))
  final_formula <- as.formula(bquote(~ .(ff)))
  return(final_formula)
}

cmp_nosp <- cmp_spatial(
  spatial_expr = NULL,
  intercept(1, model = "iid", hyper = list(prec = list(initial = 1e-3, fixed = TRUE))))

cmp_IMCAR <- cmp_spatial(
  spatial(ID, model = inla.IMCAR.model(k = 3, W = W_con), extraconstr = list(
      A = A_constr, e = c(rep(0, nrow(A_constr))))),
  intercept(Year, model = "iid", hyper = list(prec = list(initial = 1e-3, fixed = TRUE))))

cmp_PMCAR <- cmp_spatial(
  spatial(ID, model = inla.MCAR.model(k = 3, W = W_con, alpha.min = 0, alpha.max = 1)),
  intercept(1, model = "iid", hyper = list(prec = list(initial = 1e-3, fixed = TRUE))))

cmp_MLCAR <- cmp_spatial(
  spatial(ID, model = inla.MLCAR.model(k = 3, W = W_con)),
  intercept(1, model = "iid", hyper = list(prec = list(initial = 1e-3, fixed = TRUE))))

cmp_BYM <- cmp_spatial(
  spatial(ID, model = inla.MBYM.dense(k = 3, W = W_con,  PC = FALSE),
          extraconstr = list(A = A_constr, e = rep(0, 3))),
  intercept(Year, model = "iid", hyper = list(prec = list(initial = 1e-3, fixed = TRUE))))

library(inlabru)

cav_bru <- function(model_cmp, data = dd){

  terms <- c("0", "myoffset",
             "alpha_ELI",  "alpha_PGR",  "alpha_UIS",
             "alpha_PDI", "alpha_ER",
             paste0("logexpit(x = list(TEP_th, ELL),",
                    "beta = list(beta_0, beta_TEP, beta_ELL))"))

  if (any(grepl("spatial", as.character(model_cmp)))) terms <- c(terms, "spatial")
  if (any(grepl("intercept", as.character(model_cmp)))) terms <- c(terms, "intercept")
  if (any(grepl("spatiotemporal", as.character(model_cmp)))) terms <- c(terms, "spatiotemporal")

  formula <- as.formula(paste("N_ACC ~", paste(terms, collapse = " + ")))

  res <- inlabru::bru(
    components = model_cmp,
```

```
    lik = inlabru::like(family = "poisson", formula = formula, data = data),
    options = list(verbose = T, num.threads = 1,
                   control.compute = list(
                     waic = T, cpo = T, dic = T, internal.opt = F),
                   control.predictor = list(compute =  T)))
  return(res)
}
```

```
## Time needed for scaling Laplacian matrix:  1.174  seconds
```

% latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:41:46 2025

Model	WAIC	Eff_params
Null	3476.40	39.22
ICAR	2904.62	184.39
PCAR	2903.53	194.42
Leroux	2901.13	193.92
BYM	2886.89	189.78

Here we show some posterior summaries for covariate effects, using the BYM model. % latex table generated in R 4.5.0 by xtable 1.8-4 package % Tue Jun 10 17:41:46 2025

Effect	Mean	Sd	Q_0.05	Q_0.975
Int_2021	-4.98	0.10	-5.18	-4.79
Int_2022	-5.11	0.10	-5.31	-4.91
Int_2023	-4.75	0.10	-4.95	-4.56
beta_0	-2.34	0.10	-2.53	-2.14
alpha_ELI	-0.03	0.04	-0.10	0.05
alpha_PGR	0.07	0.04	-0.01	0.16
alpha_UIS	0.02	0.04	-0.06	0.10
alpha_PDI	-0.07	0.05	-0.16	0.02
alpha_ER	-0.27	0.06	-0.39	-0.16
beta_TEP	-0.25	0.05	-0.34	-0.16
beta_ELL	-0.26	0.06	-0.37	-0.15

Model interpretation is equivalent to the linear model.

Botella-Rocamora, Paloma, Miguel A Martinez-Beneito, and Sudipto Banerjee. 2015. "A Unifying Modeling Framework for Highly Multivariate Disease Mapping." *Statistics in Medicine* 34 (9): 1548–59.

Gelman, Andrew, Jessica Hwang, and Aki Vehtari. 2014. "Understanding Predictive Information Criteria for Bayesian Models." *Statistics and Computing* 24 (6): 997–1016. https://doi.org/10.1007/S11222-013-9416-2.

Kabe, Dattatraya G. 1964. "A Note on the Bartlett Decomposition of a Wishart Matrix." *Journal of the Royal Statistical Society Series B: Statistical Methodology* 26 (2): 270–73. https://doi.org/https://doi.org/10.1111/j.2517-6161.1964.tb00558.x.

Leroux, Brian G., Xingye Lei, and Norman Breslow. 2000. "Estimation of Disease Rates in Small Areas: A New Mixed Model for Spatial Dependence." In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, edited by M. Elizabeth Halloran and Donald Berry, 179–91. New York, NY: Springer New York. https://doi.org/https://doi.org/10.1007/978-1-4612-1284-3_4.

Lindgren, Finn, Fabian Bachl, Janine Illian, Man Ho Suen, Håvard Rue, and Andrew E Seaton. 2024. "Inlabru: Software for Fitting Latent Gaussian Models with Non-Linear Predictors." *arXiv Preprint arXiv:2407.00791*. https://doi.org/https://doi.org/10.48550/arXiv.2407.00791.

Palmí-Perales, Francisco, Virgilio Gómez-Rubio, and Miguel A. Martinez-Beneito. 2021. "Bayesian Multivariate Spatial Models for Lattice Data with INLA." *Journal of Statistical Software* 98 (2): 1–29. https://doi.org/10.18637/jss.v098.i02.

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.

Stoner, Oliver, Theo Economou, and Gabriela Drummond Marques da Silva. 2019. "A Hierarchical Framework for Correcting Under-Reporting in Count Data." *Journal of the American Statistical Association*.

Urdangarin, Arantxa, Tomas Goicoa, T. Kneib, and M. D. Ugarte. 2024. "A Simplified Spatial+ Approach to Mitigate Spatial Confounding in Multivariate Spatial Areal Models." *Spatial Statistics* 59: 100804. https://doi.org/10.1016/j.spasta.2023.100804.

Van Niekerk, Janet, Elias Krainski, Denis Rustand, and Haavard Rue. 2023. "A New Avenue for Bayesian Inference with INLA." *Computational Statistics and Data Analysis* 181. https://doi.org/10.1016/j.csda.2023.107692.

Vicente, Gonzalo, Aritz Adin, Tomás Goicoa, and María Dolores Ugarte. 2023. "High-Dimensional Order-Free Multivariate Spatial Disease Mapping." *Statistics and Computing* 33 (5): 104.

Wøllo, Sara E. 2022. "Correcting for Under-Reporting of Violence Against Women in Italy Using INLA." NTNU Open. https://hdl.handle.net/11250/3026838.