Bayesian Group Fused Priors

F. Mortier

M. Denis, J. Gibaud, B. Heuclin, S. Tisné & C. Trottier

Rencontres au sommet,
March 26, 2024

AGRICULTURAL RESEARCH FOR DEVELOPMENT

GEORGETOWN UNIVERSITY

This project has received funding from The European Unions's Horizon 2020 Research and innovation programme Under grant agreement No 840383.

Abscission process

Abscission: the process shedding of various parts of an organism, such as a plant dropping a leaf, fruit, flower, or seed.

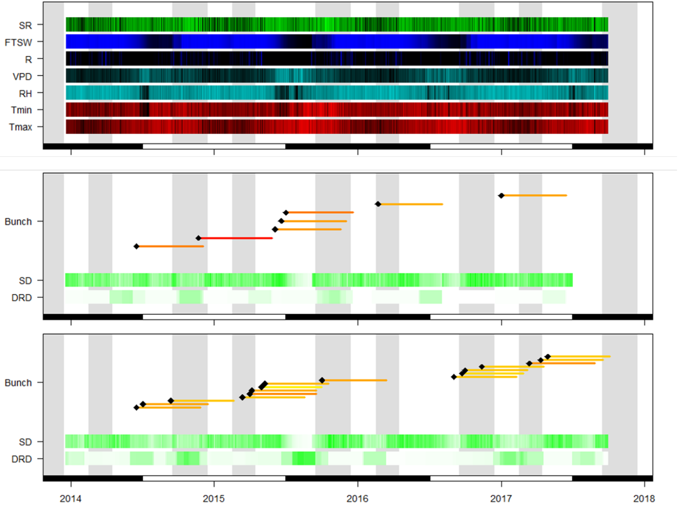
An optimal execution of the abscission process is of major importance for species survival

Environmental variations impact abscission process with varying effects over developmental stages

Question:
Identify the most relevant set of environmental factors and the time periods at which they modulate the abscission process

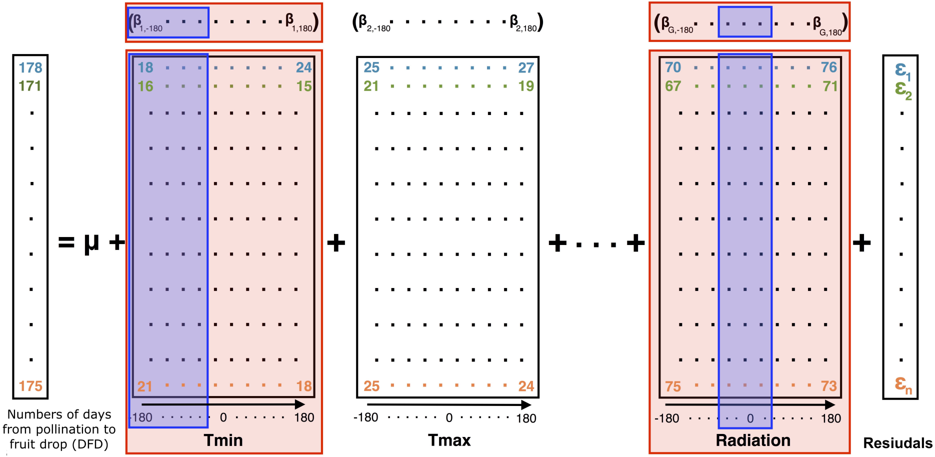
Oil palm: fruit abscission process

The fruit abscission of oil palm trees (Dataset provided by ”le Centre de Recherches Agricoles-Plantes Pérennes (CRA-PP)” (Tisné et al., 2020))



Oil palm: fruit abscission process

To identify the environmental variables and the time periods affecting the oil palm fruit abscission process



Statistical model

$$y = \mu + \sum_{g=1}^{G} X_g \beta_g + \varepsilon, \ \varepsilon \sim \mathcal{N}_n(0, \sigma^2 I_n)$$

with

- $\mathbf{y} = (y_1, \ldots, y_n)'$ the n-vector of outcomes,
- $X_g = \left[\mathbf{x}'_{g1}, \ldots, \mathbf{x}'_{gT}\right]$ a $n \times T$ matrix of covariates measured at T regular spaced times for $g = 1, \ldots, G$,
- $\boldsymbol{\beta}_g = (\beta_{g1}, \ldots, \beta_{gT})'$ the T-vector of coefficients associated to group g,
- $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)'$ the n-vector of residuals,
- σ^2 the residual variance.

Statistical model

$$y = \mu + \sum_{g=1}^{G} X_g \beta_g + \varepsilon,\ \varepsilon \sim \mathcal{N}_n(0, \sigma^2 I_n)$$

Objectives

- To identify environmental variables $\rightarrow$ Selection of groups of temporally correlated variables (or time series)
- To identify time periods $\rightarrow$ Selection of correlated variables within groups

$\rightarrow$ Need to develop statistical approaches **selecting** groups of correlated variables and variables within groups while considering the group **structure** and the natural order of variables within groups

Statistical challenges

- High correlation between consecutive variables $\rightarrow$ ill-conditioned and over-fitting problems
- Double selection

$\rightarrow$ Bayesian framework: selection and integration of dependence structure between variables taken into account by specifying specific priors

Selection

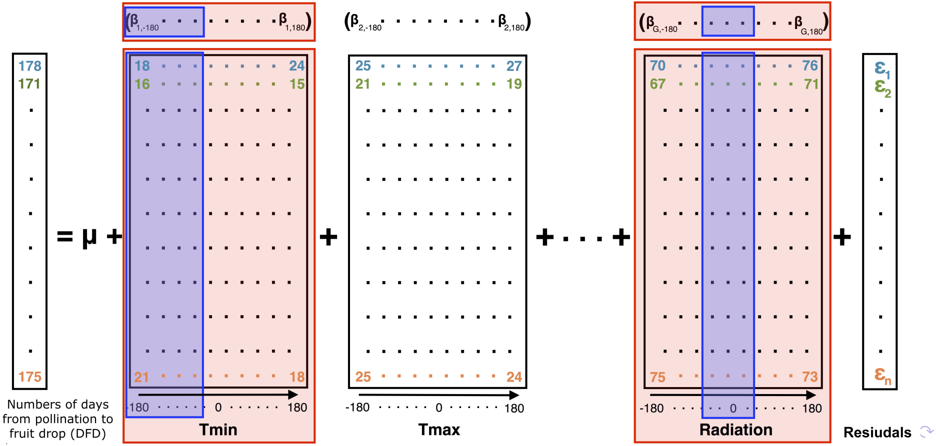
To shrink towards zero small coefficients while leaving large signals large: **Shrinkage** priors

Structure

Priors with a **variance-covariance matrix** related to structure information between variables

In fruit abscission process context

Continuous shrinkage priors (Park and Casella, 2008; Polson and Scott, 2011) with a specific covariance structure to identify the environmental variables and the time periods affecting the oil palm fruit abscission process



Bayesian fused priors

Usual approach to take into account time structure within covariate matrix while imposing sparsity on coefficients

Global-local parametrization

$$\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\upsilon^2\gamma_t^2\sigma^2}} \exp\left(-\frac{\beta_t^2}{2\upsilon^2\gamma_t^2\sigma^2}\right) \quad \text{and} \quad \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\lambda^2\omega_t^2\sigma^2}} \exp\left(-\frac{(\beta_t - \beta_{t-1})^2}{2\lambda^2\omega_t^2\sigma^2}\right)$$

- Global shrinkage parameters ν and λ: perform shrinkage on all coefficients and their differences
- Local shrinkage parameters γ_t and ω_t: allow true large signals to escape from the overall shrinkage

Bayesian fused priors

To investigate the trade-off between strong shrinkage prior on coefficients and their differences

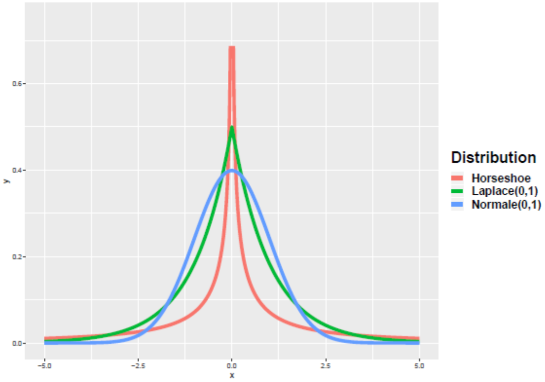


Figure: Continuous shrinkage prior distributions

Prior names	Difference prior	Coefficient prior	Reference
Fused NE-NE Bayesian fused Lasso	$\lambda^2 \sim \mathcal{IG}(a, b)$ $\omega_t^2 \sim \mathcal{E}xp(1/2)$	$\upsilon^2 \sim \mathcal{IG}(s, r)$ $\gamma_t^2 \sim \mathcal{E}xp(1/2)$	(Kyung et al., 2010)
Fused NEG-NE	$\lambda^2 = 1$ $\omega_t^2 \vert \psi_t \sim \mathcal{E}xp(\psi_t)$ $\psi_t \sim \mathcal{G}(a, b)$	$\upsilon^2 \sim \mathcal{IG}(s, r)$ $\gamma_t^2 \sim \mathcal{E}xp(1/2)$	(Shimamura et al., 2019)
Fused HS-NE	$\lambda \sim \mathcal{C}^+(0, 1)$ $\omega_t \sim \mathcal{C}^+(0, 1)$	$\upsilon^2 \sim \mathcal{IG}(s, r)$ $\gamma_t^2 \sim \mathcal{E}xp(1/2)$	(Kakikawa et al., 2023)
Fused HS-HS	$\lambda \sim \mathcal{C}^+(0, 1)$ $\omega_t \sim \mathcal{C}^+(0, 1)$	$\upsilon \sim \mathcal{C}^+(0, 1)$ $\gamma_j \sim \mathcal{C}^+(0, 1)$	
Fused HS-NhC	$\lambda \sim \mathcal{C}^+(0, 1)$ $\omega_t \sim \mathcal{C}^+(0, 1)$	$\upsilon = 1$ $\gamma_t \sim \mathcal{C}^+(0, 1)$	

Table: Fused priors in the one group context ($G = 1$).

Bayesian multi-group fused priors

$$\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi \upsilon_g^2 \gamma_{gt}^2 \sigma^2}} \exp\left(-\frac{\beta_{gt}^2}{2\upsilon_g^2\gamma_{gt}^2\sigma^2}\right) \quad \text{and} \quad \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi \lambda_g^2 \omega_{gt}^2 \sigma^2}} \exp\left(-\frac{(\beta_{gt}-\beta_{g,t-1})^2}{2\lambda_g^2\omega_{gt}^2\sigma^2}\right).$$

Prior names	**Difference prior**	**Coefficient prior**
Specific HS-NE	$\lambda_g \sim \mathcal{C}^+(0,1)$ $\omega_{gt} \sim \mathcal{C}^+(0,1)$	$\upsilon_g^2 \sim \mathcal{IG}(s,r)$ $\gamma_{gt}^2 \sim \mathcal{E}xp(1/2)$
Global HS-NE	$\lambda_g = \lambda$ $\lambda \sim \mathcal{C}^+(0,1)$ $\omega_{gt} \sim \mathcal{C}^+(0,1)$	$\upsilon_g^2 \sim \mathcal{IG}(s,r)$ $\gamma_{gt}^2 \sim \mathcal{E}xp(1/2)$
Specific HS-NhC	$\lambda_g \sim \mathcal{C}^+(0,1)$ $\omega_{gt} \sim \mathcal{C}^+(0,1)$	$\gamma_{gt} \sim \mathcal{C}^+(0,1)$
Global HS-NhC	$\lambda_g = \lambda$ $\lambda \sim \mathcal{C}^+(0,1)$ $\omega_{gt} \sim \mathcal{C}^+(0,1)$	$\gamma_{gt} \sim \mathcal{C}^+(0,1)$

Table: Fused priors in the multi-group context.

Inference: technical aspects

Conjugate prior distributions:

$$\beta_g|\gamma_g, \lambda^2, \omega_g, \sigma^2 \sim \mathcal{N}_T(0, \sigma^2 \boldsymbol{Q}_g^{-1}), \quad g = 1, \ldots, G$$

where Q_g is equal to

$$\boldsymbol{Q}_g = \left(\boldsymbol{\Upsilon}_g^{-1} + \boldsymbol{D}_g^\top \boldsymbol{\Omega}_g^{-1} \boldsymbol{D}_g / \lambda^2 \right).$$

For all shrinkage parameters (Makalic and Schmidt, 2015):

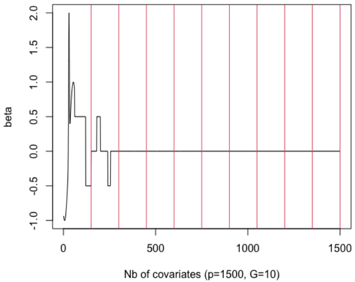
$$x \sim \mathcal{C}^+(0,1) \quad \Leftrightarrow \quad x^2|\xi \sim IG(1/2, 1/\xi), \quad \xi \sim IG(1/2, 1),$$

All full conditional distributions have closed forms → Gibbs sampler (GitHub: `https://github.com/Heuclin/GroupFusedHorseshoe`)

Simulation study

Simulation design:

- $p = 1500$ covariates divided into $G = 1, 10, 30, 100$ groups
- $T = \min\left(\frac{p}{\max(10, G)}, 60\right)$
- $n = 150$ and $\sigma^2 = 1, 4$



$$\beta_t = \begin{cases} \sin(4t/T - 2) + 2e^{-30(4t/T-2)^2} & t < T \\ 0.5 & t \in [T+1, 2T] \\ -0.5 & t \in [2T+1, (2+1/2)T] \\ 0.5 & t \in [3T+1, (3+1/3)T] \\ -0.5 & t \in [4T+1, (4+1/4)T] \\ 0 & \text{otherwise} \end{cases}$$

$X_g \sim \mathcal{N}(0, AR_{0.95})$.

Simulation results

Priors / Diff_Coeff	MCC	MSE_z	MSE_{nz}
HS_NhC	**0.90688**	**0.00003**	**0.01485**
HS_NE	0.90018	0.00067	0.05000
HS_HS	0.06456	0.00000	2.33525
NE_NE	0.21203	0.00273	0.04602

Table: Matthews Correlation Coefficient (MCC), mean squared errors of true zeroes (MSE_z), and mean squared errors of non-zero coefficients MSE_{nz} using the different priors with residual variance σ^2 equal to 1 and $G = 1$.

Simulation results

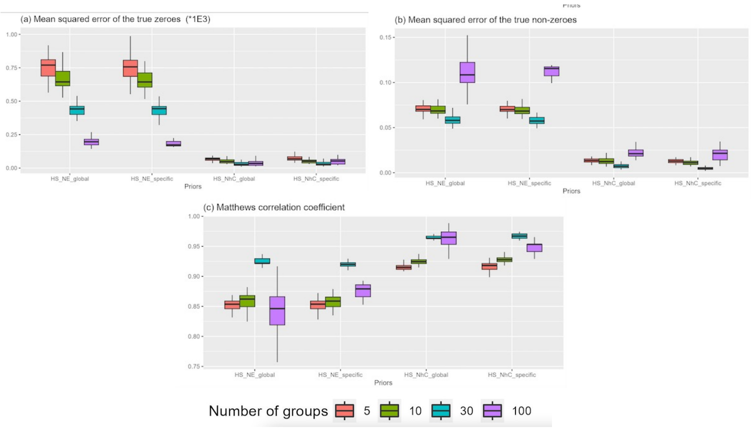


Figure: (a) Mean squared errors of true zero coefficients MSE_z, (b) Mean squared errors of non-zeroes (MSE_{nz}), and (c) Matthews Correlation Coefficient (MCC) using the different priors with residual variance σ^2 equal to 1, and $G = 5, 10, 30, 100$.

Application on oil palm

Data

- 1,173 bunches (statistical unit)
- Outcome (y): number of days from pollination to fruit drop
- 5 climatic variables: Tmax, Tmin, Relative air humidity (RH), Rainfall (R), Solar radiation (SR)
- 5 ecophysiological variables: Maximum daily vapor pressure deficit (VPD), Fraction of transpirable soil water (FTSW), Supply-demand ratio (SD), Daily reproductive demand (DRD)
- 121 time points for each environmental variables: $p = 1,210$ predictors greater than $n = 1,173$ observations

Application on oil palm

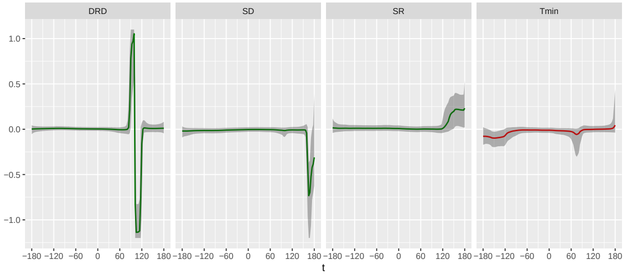


Figure: Estimated coefficient profiles for DRD, SD, SR, Tmin. Gray shadows represent the 95% credible interval.

- Identification of 4 environmental variables: DRD, SD, SR, Tmin
- Identification of relevant time periods
  - Tmin: smooth effect during the inflorescence development
  - DRD and SD: punctual effects at the end of the fruit bunch development
  - SR: smooth effect at the end of the fruit bunch development

Conclusion/Perspectives

Conclusion

Four Bayesian priors proposed:

- Able to handle different structures and to select relevant variables and/or groups of variables,
- Easily adaptable to a broad type of dependence structures : applicable in various applications (varying coefficient models, near infrared spectroscopy context (NIRS), QTL mapping, ...)

Perspectives

- To integrate prior knowledge on strengths of connections,
- To integrate dependence structures between observations,
- To extend to multivariate case (Y multivariate),
- To consider a multi-dimensional indexation.

Bibliography

Kakikawa, Y., Shimamura, K., and Kawano, S. (2023). Bayesian fused lasso modeling via horseshoe prior. *Japanese Journal of Statistics and Data Science*, 6(2):705–727.

Kyung, M., Gill, J., Ghosh, M., Casella, G., et al. (2010). Penalized regression, standard errors, and bayesian lassos. *Bayesian Analysis*, 5(2):369–411.

Makalic, E. and Schmidt, D. F. (2015). A simple sampler for the horseshoe estimator. *IEEE Signal Processing Letters*, 23(1):179–182.

Park, T. and Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686.

Polson, N. G. and Scott, J. G. (2011). Shrink Globally, Act Locally: Sparse Bayesian Regularization and Prediction. In *Bayesian Statistics 9*. Oxford University Press.

Shimamura, K., Ueki, M., Kawano, S., and Konishi, S. (2019). Bayesian generalized fused lasso modeling via neg distribution. *Communications in Statistics-Theory and Methods*, 48(16):4132–4153.

Tisné, S., Denis, M., Domonhédo, H., Pallas, B., Cazemajor, M., Tranbarger, T. J., and Morcillo, F. (2020). Environmental and trophic determinism of fruit abscission and outlook with climate change in tropical regions. *Plant-Environment Interactions*, 1(1):17–28.

Kakikawa, Y., Shimamura, K., and Kawano, S. (2023). Bayesian fused lasso modeling via horseshoe prior. *Japanese Journal of Statistics and Data Science*, 6(2):705–727.

Kyung, M., Gill, J., Ghosh, M., Casella, G., et al. (2010). Penalized regression, standard errors, and bayesian lassos. *Bayesian Analysis*, 5(2):369–411.

Makalic, E. and Schmidt, D. F. (2015). A simple sampler for the horseshoe estimator. *IEEE Signal Processing Letters*, 23(1):179–182.

Park, T. and Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686.

Polson, N. G. and Scott, J. G. (2011). Shrink Globally, Act Locally: Sparse Bayesian Regularization and Prediction. In *Bayesian Statistics 9*. Oxford University Press.

Shimamura, K., Ueki, M., Kawano, S., and Konishi, S. (2019). Bayesian generalized fused lasso modeling via neg distribution. *Communications in Statistics-Theory and Methods*, 48(16):4132–4153.

Tisné, S., Denis, M., Domonhédo, H., Pallas, B., Cazemajor, M., Tranbarger, T. J., and Morcillo, F. (2020). Environmental and trophic determinism of fruit abscission and outlook with climate change in tropical regions. *Plant-Environment Interactions*, 1(1):17–28.