Bayesian spatio-temporal quantile regression models

Jorge Castillo-Mateo

jorgecm@unizar.es

1 December 2025

Sapienza Università di Roma

Departamento de Métodos Estadísticos, Universidad Zaragoza

Introduction

Spatio-temporal quantile regression model[1]

The Annals of Applied Statistics
2023, Vol. 17, No. 3, 2305–2325
https://doi.org/10.1214/22-AOAS1719

SPATIAL QUANTILE AUTOREGRESSION FOR SEASON WITHIN YEAR DAILY MAXIMUM TEMPERATURE DATA

BY JORGE CASTILLO-MATEO[1,a], JESÚS ASÍN[1,b], ANA C. CEBRIÁN[1,c], ALAN E. GELFAND[2,e] AND JESÚS ABAURREA[1,d]

[1]*Department of Statistical Methods, University of Zaragoza,* [a]*jorgecm@unizar.es,* [b]*jasin@unizar.es,* [c]*acebrian@unizar.es,* [d]*abaurrea@unizar.es*

[1]Castillo-Mateo et al. 2023.

Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**

Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**

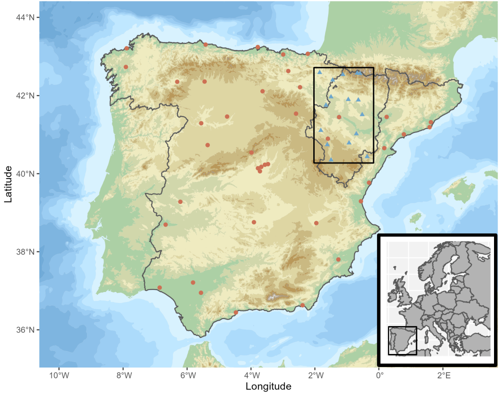
Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**
3. data sources – **AEMET – Spanish Meteorological Agency**

Data source

Daily maximum temperature over Aragon and Navarra, Spain



- $Y_{t\ell}(\mathbf{s})$: **daily max temp** for day ℓ within year t at location $\mathbf{s}$
- Sample size $\sim 55{,}200$
- Days: $\ell = 152, \ldots, 253$ (JJA)
- Years: $t = 1, \ldots, T$, $T = 40$ (1976–2015)
- Sites: $\mathbf{s}_i \in \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$, $n = 15$
- Interest in $\mathbf{s} \in D$, D Aragón and Navarra

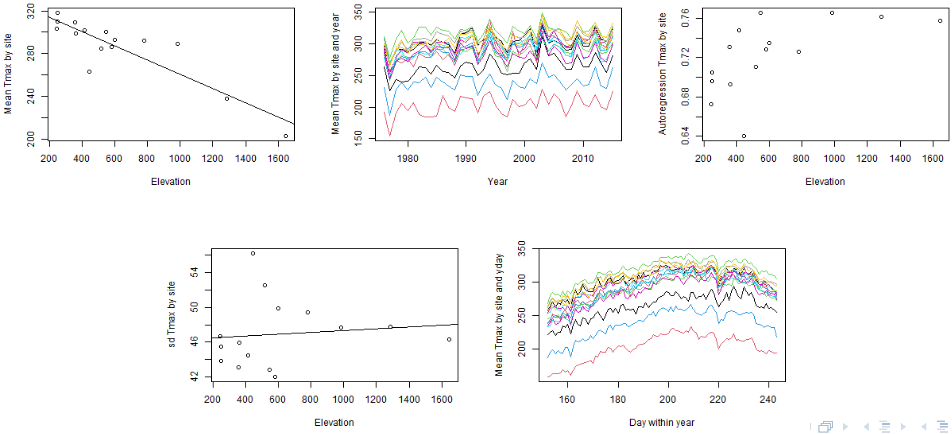
Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**
3. data sources – **AEMET – Spanish Meteorological Agency**
4. exploratory data analysis – **basic characteristics**

Exploratory data analysis

Characteristics of the mean



Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**
3. data sources – **AEMET – Spanish Meteorological Agency**
4. exploratory data analysis – **basic characteristics**
5. generative models – **mean models?**

Generative models

Linear regression models the relationship between response and covariates:

$$Y_i = \mathbf{x}_i \boldsymbol{\beta} + \epsilon_i, \qquad i = 1, \ldots, n,$$

- Y_i response variable,
- $\mathbf{x}_i$ row vector of p explanatory variables,
- $\boldsymbol{\beta}$ column vector of regression coefficients,
- ϵ_i pure error term.

Model fitting usually requires estimating the regression coefficients $\boldsymbol{\beta}$ such that the error terms ϵ_i's are minimized in some sense.

Generative models

Mean regression: Ordinary least squares

For mean regression:

$$Y_i = \mathbf{x}_i \boldsymbol{\beta} + \epsilon_i, \qquad i = 1, \ldots, n,$$

- interest falls on $E(Y \mid \mathbf{x}) = \mathbf{x}\boldsymbol{\beta}$; it is assumed that $E(\epsilon \mid \mathbf{x}) = 0$.
- The classical (mean) model estimates $\boldsymbol{\beta}$ by minimizing the **squared error loss**:

$$L_2(\boldsymbol{\beta}) = \sum_{i=1}^{n} (Y_i - \mathbf{x}_i \boldsymbol{\beta})^2.$$

- The minimizer is the **ordinary least squares (OLS)** estimator:

$$\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y},$$

where $\mathbf{X}$ is the $n \times p$ design matrix and $\mathbf{Y}$ is the vector of responses.

Generative models

Mean regression: Classical or frequentist framework

Under the assumption of iid normal errors,

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon, \qquad \epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n),$$

the OLS estimator has the following properties:

- It is **unbiased** and normally distributed:

$$\widehat{\beta}_{\text{OLS}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}, \qquad \widehat{\beta}_{\text{OLS}} \sim N\left(\beta, \sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1}\right).$$

- The residual variance estimator is

$$\widehat{\sigma}^2 = \frac{1}{n-p} (\mathbf{Y} - \mathbf{X}\widehat{\beta})^\top (\mathbf{Y} - \mathbf{X}\widehat{\beta}), \qquad \frac{(n-p)\widehat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p}.$$

Generative model

Mean regression: Maximum likelihood

- The OLS estimator coincides with the maximum likelihood estimator if the errors ϵ_i are normally distributed:

$$N(\epsilon \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{\epsilon - \mu}{\sigma}\right)^2\right\},$$

with zero mean $\mu = 0$ and variance σ^2; equivalently, $\epsilon_i \sim$ iid $N(0, \sigma^2)$.

- In multivariate notation for the response,

$$N_n(\mathbf{Y} \mid \mathbf{X}\beta, \sigma^2\mathbf{I}_n) = \frac{1}{\sqrt{(2\pi\sigma^2)^n}} \exp\left\{-\frac{1}{2}(\mathbf{Y} - \mathbf{X}\beta)^\top(\mathbf{Y} - \mathbf{X}\beta)\right\}.$$

Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**
3. data sources – **AEMET – Spanish Meteorological Agency**
4. exploratory data analysis – **basic characteristics**
5. generative models – **mean models?**
6. Bayesian framework for quantifying uncertainty

Bayesian framework

- Let **Y** denote **data** generated from the **likelihood** $[\mathbf{Y} \mid \theta]$.
- Let $[\theta]$ be a **prior** distribution for the parameters.
- The central goal of Bayesian statistics is to approximate the **joint posterior distribution**:

$$[\theta \mid \mathbf{Y}] = \frac{[\mathbf{Y} \mid \theta][\theta]}{[\mathbf{Y}]} \propto [\mathbf{Y} \mid \theta][\theta].$$

- The **marginal likelihood** $[\mathbf{Y}]$ is given by the integral:

$$[\mathbf{Y}] = \int [\mathbf{Y} \mid \theta][\theta] \, d\theta.$$

  which is typically high-dimensional and not analytically tractable.
- Direct computation of the posterior is usually infeasible.

Bayesian framework

The Gibbs sampler: Sampling from the posterior using MCMC

Initialization: Choose an arbitrary starting value: $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_i^{(0)}, \theta_{i+1}^{(0)}, \ldots, \theta_K^{(0)})$
Iterate a cycle:

$$
\begin{aligned}
\text{Step 1:} \quad & \theta_1^{(b)} \sim [\theta_1 \mid \theta_2^{(b-1)}, \theta_3^{(b-1)}, \ldots, \theta_K^{(b-1)}, \mathbf{Y}] \\
\text{Step 2:} \quad & \theta_2^{(b)} \sim [\theta_2 \mid \theta_1^{(b)}, \theta_3^{(b-1)}, \ldots, \theta_K^{(b-1)}, \mathbf{Y}] \\
& \vdots \\
\text{Step } i\text{:} \quad & \theta_i^{(b)} \sim [\theta_i \mid \theta_1^{(b)}, \ldots, \theta_{i-1}^{(b)}, \theta_{i+1}^{(b-1)}, \ldots, \theta_K^{(b-1)}, \mathbf{Y}] \\
\text{Step } i+1\text{:} \quad & \theta_{i+1}^{(b)} \sim [\theta_{i+1} \mid \theta_1^{(b)}, \ldots, \theta_i^{(b)}, \theta_{i+2}^{(b-1)}, \ldots, \theta_K^{(b-1)}, \mathbf{Y}] \\
& \vdots \\
\text{Step } K\text{:} \quad & \theta_K^{(b)} \sim [\theta_K \mid \theta_1^{(b)}, \theta_2^{(b)}, \ldots, \theta_{K-1}^{(b)}, \mathbf{Y}]
\end{aligned}
$$

end Iterate

Bayesian framework

Gibbs sampler for mean regression

- The full posterior distribution is calculated proportional to the likelihood and priors:

$$[\theta_i \mid \boldsymbol{\theta}_{-i}, \mathbf{Y}] \propto [\boldsymbol{\theta} \mid \mathbf{Y}] \propto [\mathbf{Y} \mid \boldsymbol{\theta}][\boldsymbol{\theta}].$$

- In our mean regression model, the likelihood and priors combine as:

$$[\mathbf{Y} \mid \boldsymbol{\theta}][\boldsymbol{\theta}] = N_n(\mathbf{Y} \mid \mathbf{X}\beta, \sigma^2 \mathbf{I}_n) \times N_p(\beta \mid \mathbf{a}, \mathbf{B})\, G(1/\sigma^2 \mid a, b).$$

Bayesian framework

Gibbs sampler for mean regression

- Prior $\beta \sim N(\mathbf{a}, \mathbf{B})$, full conditional

$$\beta \mid \mathbf{Y}, \sigma^2 \sim N(\hat{\beta}, \mathbf{\Omega}_\beta);$$
$$\hat{\beta} = \mathbf{\Omega}_\beta \left(\frac{1}{\sigma^2} \mathbf{X}^\top \mathbf{Y} + \mathbf{B}^{-1} \mathbf{a} \right),$$
$$\mathbf{\Omega}_\beta^{-1} = \frac{1}{\sigma^2} \mathbf{X}^\top \mathbf{X} + \mathbf{B}^{-1}.$$

- Prior $1/\sigma^2 \sim G(a, b)$, full conditional

$$\frac{1}{\sigma^2} \mid \mathbf{Y}, \beta \sim G\left(\frac{n}{2} + a, \frac{1}{2} (\mathbf{Y} - \mathbf{X}\beta)^\top (\mathbf{Y} - \mathbf{X}\beta) + b \right).$$

The R package spTReg

Local models

Time to practice...

- Load the R package `spTReg`.
- Use the `iidm` function to fit a mean model for Zaragoza station.
- The model should include an intercept, harmonic terms, and a linear trend in years.
- Obtain a summary of the coefficients using the `confint` function.
- Use the `predict` function and `hist` or `qqnorm` to check normality in the posterior mean of the residuals.

Generative models

Mean vs quantile regression

The starting point for many regression methods is the mean model:

$$Y_i = \mathbf{x}_i \beta + \epsilon_i, \qquad i = 1, \ldots, n.$$

Assumes a mean relationship and homoscedastic, symmetric errors.

Limitations:

- Sensitive to outliers, heavy–tailed, or asymmetric errors
- Describes only the average behavior of Y
- Cannot capture heterogeneous effects across the distribution

This motivates:

- Quantile regression for modeling all quantile levels of the distribution.
- Bayesian spatial models to handle spatial dependence and uncertainty.

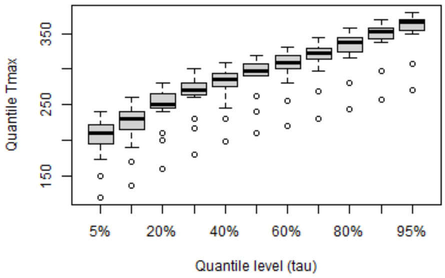
Introduction

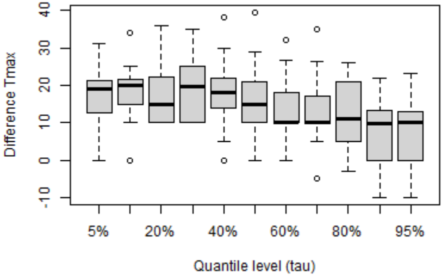
Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**
3. data sources – **AEMET – Spanish Meteorological Agency**
4. exploratory data analysis – **specific characteristics**
5. generative models

Exploratory data analysis

Characteristics of the quantiles





Generative models

Quantile regression: Asymmetrically weighted error loss

For quantile regression:

- interest falls on $Q_\tau(Y \mid \mathbf{x}) = \mathbf{x}\boldsymbol{\beta}$ for a desired quantile order $\tau \in (0, 1)$; it is assumed that $Q_\tau(\epsilon \mid \mathbf{x}) = 0$.
- The quantile model estimates $\boldsymbol{\beta}$ by minimizing an asymmetrically weighted error loss,[2]

$$L_{1,\tau}(\boldsymbol{\beta}) = \left[\sum_{i \in \{i : Y_i \geq \mathbf{x}_i \boldsymbol{\beta}\}} \tau \, |Y_i - \mathbf{x}_i^\top \boldsymbol{\beta}| + \sum_{i \in \{i : Y_i < \mathbf{x}_i \boldsymbol{\beta}\}} (1 - \tau) \, |Y_i - \mathbf{x}_i \boldsymbol{\beta}| \right].$$

[2] Koenker and Bassett 1978.

Generative models

Quantile regression: Asymmetrically weighted error loss

For quantile regression:

- The quantile regression estimator is defined as the minimizer of the sum of asymmetrically weighted absolute residuals; it does not have a closed-form solution, but can be efficiently computed using linear programming methods:

$$\widehat{\beta}_\tau = \arg \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau \left(Y_i - \mathbf{x}_i \beta \right).$$

where the **check loss function**:

$$\rho_\tau(u) = \begin{cases} \tau\, u, & u \geq 0, \\ (\tau - 1)\, u, & u < 0. \end{cases}$$

Generative models

Quantile regression: Median minimizes sum of absolute deviations

Claim: The median of a set of numbers $\{x_1, \ldots, x_n\}$ minimizes the sum of absolute deviations:

$$f(c) = \sum_{i=1}^{n} |x_i - c|.$$

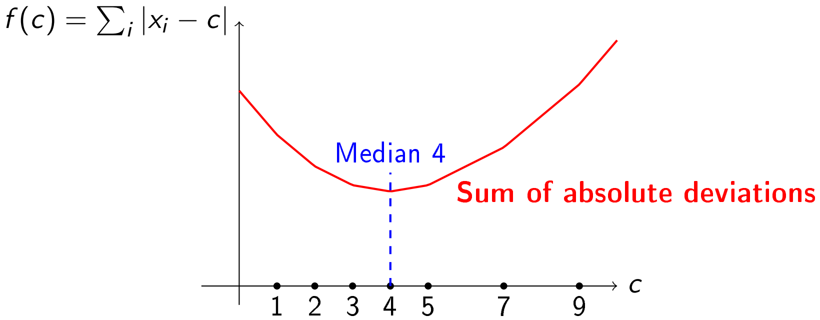
Proof Sketch:

- Assume $x_1 \leq x_2 \leq \cdots \leq x_n$ (sorted order).
- Consider $f(c) = \sum_{i=1}^{n} |x_i - c|$.
- For $c < x_1$, increasing c decreases $|x_i - c|$ for all i, so $f(c)$ decreases.
- For $c > x_n$, decreasing c decreases $|x_i - c|$ for all i, so $f(c)$ decreases.
- Between x_1 and x_n, each time c crosses a data point, the slope of $f(c)$ changes by $+1$ or -1.
- The function $f(c)$ is piecewise linear and convex, with minimum at the median.

Conclusion: The median is the value that minimizes $\sum_i |x_i - c|$.

Generative models

Quantile regression: Median minimizes sum of absolute deviations (convex function)



- $f(c)$ is **convex** and piecewise linear.
- Its minimum occurs at the **median** of the data ($c = 4$).
- Moving c left or right from the median increases $f(c)$.

Generative models

Quantile regression: Maximum likelihood

- The quantile regression estimator coincides with the maximum likelihood estimator if the errors ϵ_i follow an asymmetric Laplace distribution:[3]

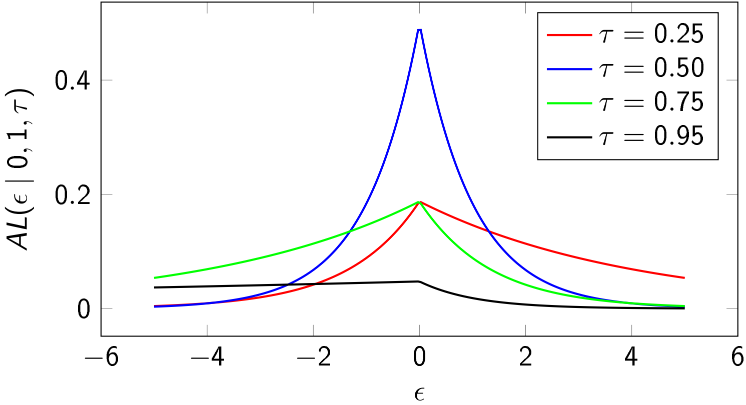
$$AL(\epsilon \mid \mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\rho_\tau\left(\frac{\epsilon - \mu}{\sigma}\right)\right\},$$

where $\rho_\tau(u) = u(\tau - \mathbf{1}(u < 0))$, and with zero location $\mu = 0$, scale σ, for the quantile order τ; or equivalently, $\epsilon_i \sim$ iid $AL(0, \sigma, \tau)$.

[3] Yu and Moyeed 2001.

Generative models

Quantile regression: Asymmetric Laplace density



Bayesian framework

Gibbs sampler for quantile regression

- The full posterior distribution is calculated proportional to the likelihood and priors:

$$[\theta_i \mid \boldsymbol{\theta}_{-i}, \mathbf{Y}] \propto [\boldsymbol{\theta} \mid \mathbf{Y}] \propto [\mathbf{Y} \mid \boldsymbol{\theta}][\boldsymbol{\theta}].$$

- In our quantile regression model, the likelihood and priors combine as:

$$[\mathbf{Y} \mid \boldsymbol{\theta}][\boldsymbol{\theta}] = AL_n(\mathbf{Y} \mid \mathbf{X}\beta, \sigma \mathbf{I}_n, \tau) \times N_p(\beta \mid \mathbf{a}, \mathbf{B})\, G(1/\sigma \mid a, b).$$

- $[\beta \mid \cdots]$ and $[1/\sigma \mid \cdots]$ are no longer known in closed form (Metropolis-Hastings algorithm?).

Bayesian framework

Gibbs sampler for quantile regression

Convenient **data augmentation** approach:[4]

- ϵ_i follows a location-scale mixture of normal form with a marginal AL distribution;
- in particular,

$$\epsilon_i = \theta\xi_i + \omega\sqrt{\xi_i\sigma}Z_i$$

  where

$$\xi_i = U_i\sigma, \quad \theta = \frac{1-2\tau}{\tau(1-\tau)}, \quad \omega = \sqrt{\frac{2}{\tau(1-\tau)}},$$

  and $U_i \sim$ iid $Exp(1)$, $Z_i \sim$ iid $N(0,1)$;
- equivalently, $\epsilon_i \mid \xi_i \sim$ ind. $N(\theta\xi_i, \omega^2\xi_i\sigma)$ is normally distributed making both models, for the mean and for the quantiles, relatively equivalent to fit given the latent exponential variables.

[4] Kozumi and Kobayashi 2011.

Bayesian framework

Gibbs sampler for quantile regression

- In the augmented parameter space, the augmented likelihood and priors combine as:

$$[\mathbf{Y} \mid \boldsymbol{\theta}][\boldsymbol{\theta}] = N_n(\mathbf{Y} \mid \mathbf{X}\beta + \theta\boldsymbol{\xi}, \omega^2\sigma\boldsymbol{\Xi}) \times \prod_i Exp(\xi_i \mid 1/\sigma) \times N_p(\beta \mid \mathbf{a}, \mathbf{B})\, G(1/\sigma \mid a, b),$$

where $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)^\top$ and $\boldsymbol{\Xi} = \mathrm{diag}(\boldsymbol{\xi})$.

Bayesian framework

Gibbs sampler for quantile regression

- Prior $\beta \sim N(\mathbf{a}, \mathbf{B})$, full conditional

$$\beta \mid \mathbf{Y}, \sigma, \xi \sim N(\hat{\beta}, \mathbf{\Omega}_\beta);$$

$$\hat{\beta} = \mathbf{\Omega}_\beta \left(\frac{1}{\omega^2 \sigma} \sum_{i=1}^{n} \frac{\mathbf{x}_i^\top (Y_i - \theta \xi_i)}{\xi_i} + \mathbf{B}^{-1}\mathbf{a} \right),$$

$$\mathbf{\Omega}_\beta^{-1} = \frac{1}{\omega^2 \sigma} \sum_{i=1}^{n} \frac{\mathbf{x}_i^\top \mathbf{x}_i}{\xi_i} + \mathbf{B}^{-1}.$$

- Prior $1/\sigma \sim G(a, b)$, full conditional

$$\frac{1}{\sigma} \mid \mathbf{Y}, \beta, \xi \sim G\left(\frac{3}{2} n + a, \sum_{i=1}^{n} \left[\frac{(Y_i - \mathbf{x}_i \beta - \theta \xi_i)^2}{2 \xi_i \omega^2} + \xi_i \right] + b \right).$$

Bayesian framework

Gibbs sampler for quantile regression

- For $i = 1, \ldots, n$, prior $\xi_i \sim$ iid $Exp(1/\sigma)$, full conditional (generalized inverse Gaussian)

$$\xi_i \mid Y_i, \boldsymbol{\beta}, \sigma \sim GIG\left(\frac{1}{\sigma}\left(2 + \frac{\theta^2}{\omega^2}\right), \frac{1}{\sigma}\frac{(Y_i - \mathbf{x}_i\boldsymbol{\beta})^2}{\omega^2}, \frac{1}{2}\right).$$

Note. $GIG(a, b, p) = 1/GIG(b, a, -p)$ and $GIG(b, a, -1/2) = IG(\sqrt{a/b}, a)$ (inverse Gaussian).

Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**
3. data sources – **AEMET – Spanish Meteorological Agency**
4. exploratory data analysis – **basic characteristics**
5. generative models – **quantile models**
6. Bayesian framework for quantifying uncertainty
7. model checking and validation – **quantile metrics**

Model checking and validation

Quantile regression

1. **Pseudo-R^2 ($R^1(\tau)$)**

$$R^1(\tau) = 1 - \frac{\sum_{i=1}^{n} \rho_\tau(y_i - \hat{y}_i)}{\sum_{i=1}^{n} \rho_\tau(y_i - \tilde{y})},$$

- $\rho_\tau(u) = u(\tau - \mathbf{1}_{\{u<0\}})$ is the check (pinball) loss.
- $\hat{y}_i$ are predicted quantiles; $\tilde{y}$ is a reference (e.g., unconditional τ-quantile).
- Measures the improvement of the model over a naive baseline.

2. **Quantile Calibration: Probability of Exceedance**

$$\hat{p}_\tau = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\{y_i < \hat{y}_i\} \approx \tau$$

- Checks whether the estimated τ-quantile is correctly calibrated.

Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**
3. data sources – **AEMET – Spanish Meteorological Agency**
4. exploratory data analysis – **basic characteristics**
5. generative models – **quantile models**
6. Bayesian framework for quantifying uncertainty
7. model checking and validation – **quantile metrics**

9. make software free and easily available – `spTReg`

Software

The R package spTReg

The aim is to offer an R package, **spTReg** (available on GitHub, coming soon to CRAN), that fits, spatially predicts, and temporarily forecasts space-time dependent data with space-time varying slope coefficients, autoregressive coefficients and scale parameters including:

- Mean regression using Gaussian processes and Gaussian errors,
- Quantile regression using Gaussian processes and asymmetric Laplace errors,
- Probit and logit binary regression using Gaussian processes (soon).

Technical specifications:

- Bayesian framework
- Model fitting by MCMC $\rightarrow$ Gibbs sampling $\rightarrow$ Metropolis-within-Gibbs sampling
- Data aumentation
- C++ $\rightarrow$ **RcppArmadillo** $\rightarrow$ **Rcpp** $\rightarrow$ R

Generative models

Spatio-temporal models: General model specification

The package aims to fit the following general model (and particularizations),

$$
\begin{aligned}
Y_{t\ell}(\mathbf{s}) &= \mu_{t\ell}(\mathbf{s}) + \epsilon_{t\ell}(\mathbf{s}) \\
&= O_{t\ell}(\mathbf{s}) + \rho(\mathbf{s})(Y_{t,\ell-1}(\mathbf{s}) - O_{t,\ell-1}(\mathbf{s})) + w_{t\ell}(\mathbf{s}) + \epsilon_{t\ell}(\mathbf{s}),
\end{aligned}
$$

- $Y_{t\ell}(\mathbf{s})$ observed point-referenced response;
- location $\mathbf{s} \in D$, D study region; two time scales, t longer scale, e.g., year, $t = 1, \ldots, T$; ℓ shorter scale, e.g., day, $\ell = 1, \ldots, L_t$;
- $O_{t\ell}(\mathbf{s})$ fixed and random effects;
- $w_{t\ell}(\mathbf{s})$ spatially correlated error term;
- $\epsilon_{t\ell}(\mathbf{s})$ independent normal or AL error term with $\sigma(\mathbf{s})$ scale parameter;
- $\rho(\mathbf{s})$ autoregression parameter.

Generative models

Spatio-temporal models: Fixed and random effects

$O_{t\ell}(\mathbf{s})$ includes fixed effects, yearly-varying coefficients, and spatially-varying coefficients as

$$O_{t\ell}(\mathbf{s}) = \sum_{j=1}^{p} x_{t\ell,j}(\mathbf{s})\beta_j + \sum_{k=1}^{q} u_{t\ell,k}(\mathbf{s})\beta_{t,k} + \sum_{m=1}^{r} v_{t\ell,m}(\mathbf{s})\beta_m(\mathbf{s}),$$

- $x_{t\ell,j}(\mathbf{s})$ jth regressor with fixed effect coefficient β_j, $j = 1, \ldots, p$;
- $u_{t\ell,k}(\mathbf{s})$ kth regressor in the same or another set of regressors with yearly-varying coefficient $\beta_{t,k}$, $k = 1, \ldots, q$;
- $v_{t\ell,m}(\mathbf{s})$ mth regressor in the same or another set of regressors with spatially-varying coefficient $\beta_m(\mathbf{s})$, $m = 1, \ldots, r$.

Generative models

Spatio-temporal models: Matrix notation

Let $\mathcal{S} = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ be the set of spatial locations for which $Y_{t\ell}(\mathbf{s}_i)$ has been observed. The model in matrix notation is

$$\mathbf{Y}_{t\ell} = \mathbf{O}_{t\ell} + \mathbf{P}(\mathbf{Y}_{t,\ell-1} - \mathbf{O}_{t,\ell-1}) + \mathbf{w}_{t\ell} + \boldsymbol{\epsilon}_{t\ell},$$

$$\mathbf{O}_{t\ell} = \mathbf{X}_{t\ell}\boldsymbol{\beta} + \mathbf{U}_{t\ell}\boldsymbol{\beta}_t + \sum_{m=1}^{r} \mathbf{V}_{t\ell,m}\boldsymbol{\beta}_m(\mathbf{s}),$$

- $\mathbf{Y}_{t\ell} = (Y_{t\ell}(\mathbf{s}_1), \ldots, Y_{t\ell}(\mathbf{s}_n))^\top$,
- $\mathbf{O}_{t\ell} = (O_{t\ell}(\mathbf{s}_1), \ldots, O_{t\ell}(\mathbf{s}_n))^\top$,
- $\mathbf{P} = \mathrm{diag}(\rho(\mathbf{s}_1), \ldots, \rho(\mathbf{s}_n))$,
- $\boldsymbol{\epsilon}_{t\ell} = (\epsilon_{t\ell}(\mathbf{s}_1), \ldots, \epsilon_{t\ell}(\mathbf{s}_n))^\top$,
- $\mathbf{w}_{t\ell} = (w_{t\ell}(\mathbf{s}_1), \ldots, w_{t\ell}(\mathbf{s}_n))^\top$.
- $\mathbf{X}_{t\ell}$ $n \times p$ design matrix, $\boldsymbol{\beta}$ fixed effects coefficients.
- $\mathbf{U}_{t\ell}$ $n \times q$ design matrix accompanying the yearly-varying effects, $\boldsymbol{\beta}_t$ is the column vector of yearly-varying coefficients.
- $\mathbf{V}_{t\ell,m} = \mathrm{diag}(v_{t\ell,m}(\mathbf{s}_1), \ldots, v_{t\ell,m}(\mathbf{s}_n))$ $n \times n$ diagonal matrix accompanying the spatially-varying coefficients $\boldsymbol{\beta}_m(\mathbf{s}) = (\beta_m(\mathbf{s}_1), \ldots, \beta_m(\mathbf{s}_n))^\top$, $m = 1, \ldots, r$.

Generative models

Spatio-temporal models

Capturing spatial and temporal dependence:

- We treat the spatially correlated error process as a Gaussian process, $\mathbf{w}_{t\ell} \sim N(\rho_w \mathbf{w}_{t,\ell-1}, \boldsymbol{\Sigma}_w)$, where the covariance matrix $\boldsymbol{\Sigma}_w = \sigma_w^2 \mathbf{R}(\phi_w)$ is determined by the variance parameter σ_w^2 and the correlation matrix $R(\phi_w)$ defined from the exponential correlation function $R(\mathbf{s}, \mathbf{s}'; \phi) = \exp(-\phi||\mathbf{s} - \mathbf{s}'||)$, where $||\mathbf{s} - \mathbf{s}'||$ is the distance between sites $\mathbf{s}$ and $\mathbf{s}'$, and $\phi > 0$ is the *decay* parameter ($1/\phi$ is called the *range* parameter).
- We also assume the spatially-varying coefficient to follow a Gaussian process, $\boldsymbol{\beta}_m(\mathbf{s}) \sim N(\mathbf{z}_{\beta_m}(\mathbf{s})\gamma_{\beta_m}, \boldsymbol{\Sigma}_{\beta_m})$ hierarchically centered on the spatial covaraites for identifiability.
- $\beta_{t,k} \sim N(\rho_\beta \beta_{t-1,k}, \sigma_\beta^2)$.

Generative models

Prediction: Bayesian kriging

The details of the Bayesian kriging are as follows. In particular, we are interested in predicting the state of a GP, $W(\mathbf{s})$, at a new location $\mathbf{s}_0$. The joint distribution for $\mathbf{s} \in \{\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_n\}$ is a multivariate Gaussian distribution arising from the GP for $W(\mathbf{s})$, i.e.,

$$\begin{pmatrix} W(\mathbf{s}_0) \\ \{W(\mathbf{s}_i)\} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_0 \\ \boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} \Sigma_{00} & \{\Sigma_{i0}\}^\top \\ \{\Sigma_{i0}\} & \boldsymbol{\Sigma} \end{pmatrix} \right).$$

Therefore, the conditional distribution of the process at $\mathbf{s}_0$ is

$$\left[W(\mathbf{s}_0) \mid \{W(\mathbf{s}_i)\} = \mathbf{w}, \cdots \right] \sim N\left(\mu_0 + \{\Sigma_{i0}\}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{w} - \boldsymbol{\mu}), \Sigma_{00} - \{\Sigma_{i0}\}^\top \boldsymbol{\Sigma}^{-1}\{\Sigma_{i0}\}\right),$$

from which we would draw a sample.

Introduction

Good modeling practice

1. scientific questions – **climate change in daily temperatures?**
2. proper understanding of the process – **how do daily temperatures work?**
3. data sources – **AEMET – Spanish Meteorological Agency**
4. exploratory data analysis – **basic characteristics**
5. generative models – **quantile models**
6. Bayesian framework for quantifying uncertainty
7. model checking and validation – **quantile metrics**
8. answer the initial questions
9. make software free and easily available – `spTReg`

Answer initial questions

Fit your own models

We will fit a relatively simple spatial model:

- Use the function `spTReg` to fit a quantile model for $\tau = 0.05, 0.50, 0.95$ with a spatially varying intercept centered on its mean and altitude, and a spatially varying trend in years centered on its mean.
- Implement Bayesian kriging to predict in the Aragon and Navarra regions using the `krigeBayes` function.

References I

Castillo-Mateo, Jorge et al. (2023). "Spatial quantile autoregression for season within year daily maximum temperature data". In: *Annals of Applied Statistics* 17.3, pp. 2305–2325.

Koenker, Roger and Gilbert Bassett (1978). "Regression quantiles". In: *Econometrica* 46.1, pp. 33–50.

Kozumi, Hideo and Genya Kobayashi (2011). "Gibbs sampling methods for Bayesian quantile regression". In: *Journal of Statistical Computation and Simulation* 81.11, pp. 1565–1578.

Yu, Keming and Rana A. Moyeed (2001). "Bayesian quantile regression". In: *Statistics & Probability Letters* 54.4, pp. 437–447.