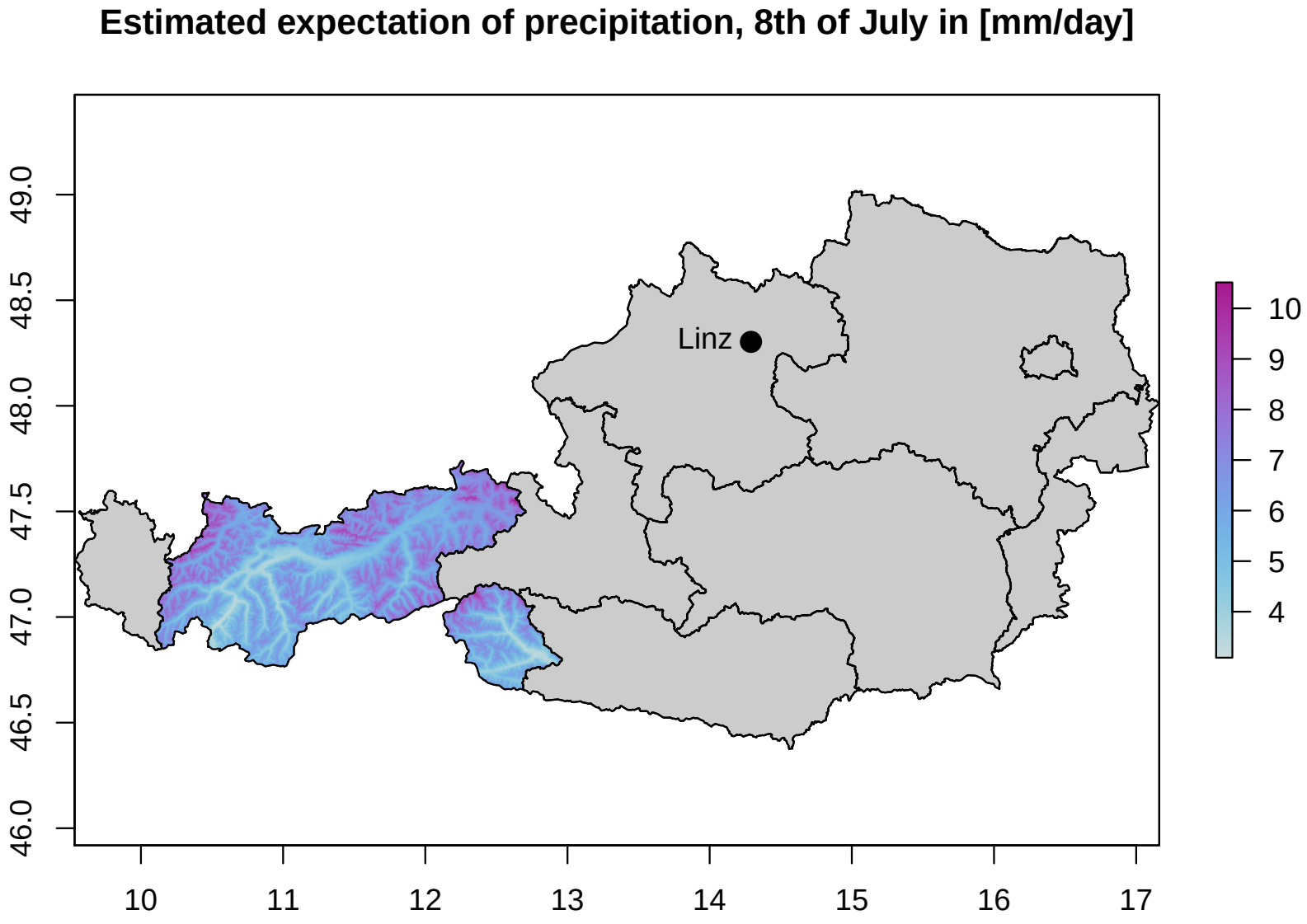
Spatio-Temporal Censored Model of Precipitation Climatology

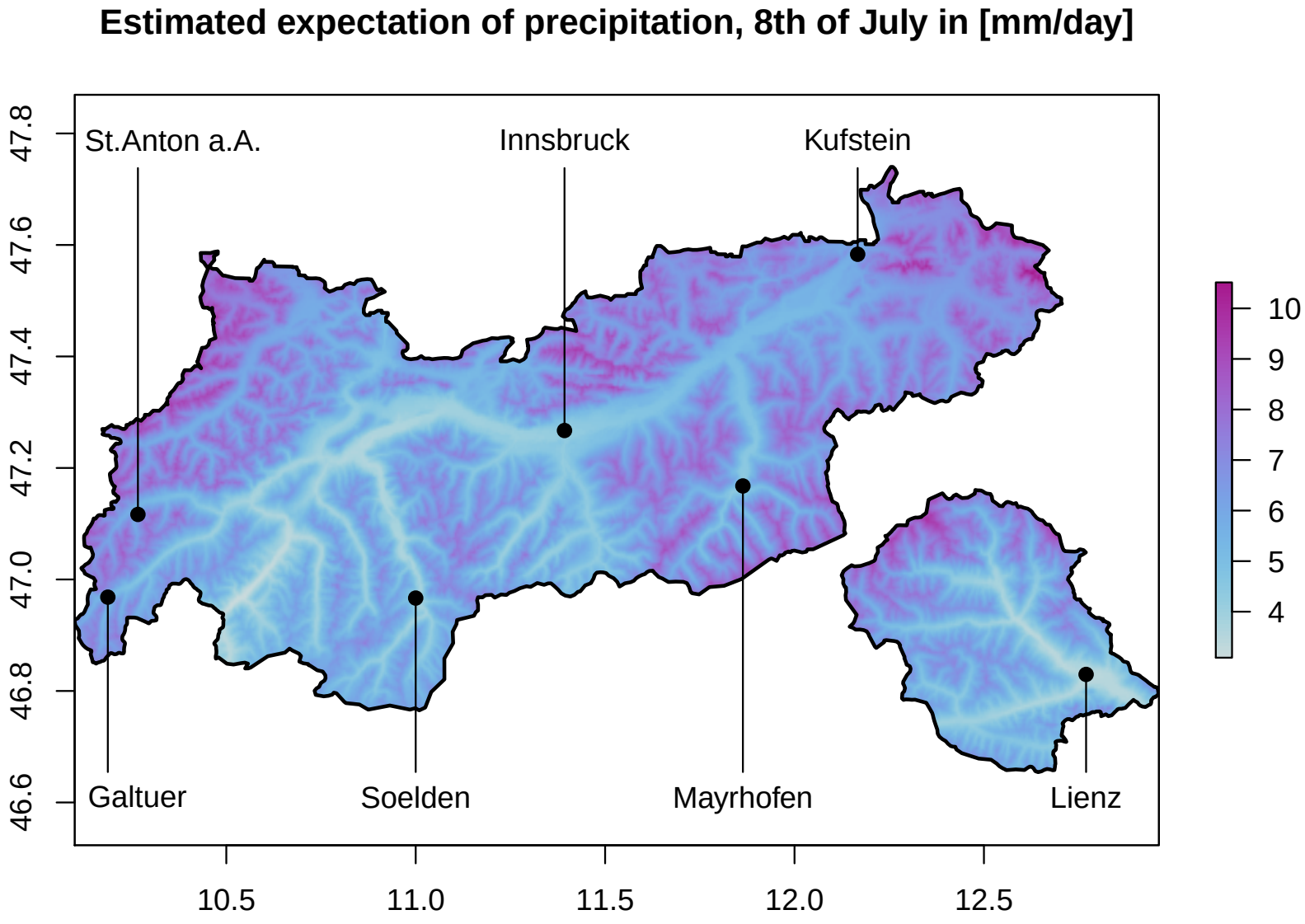
Reto Stauffer, Georg J. Mayr, Jakob Messner, Nikolaus Umlauf, Achim Zeileis

reto.stauffer@uibk.ac.at

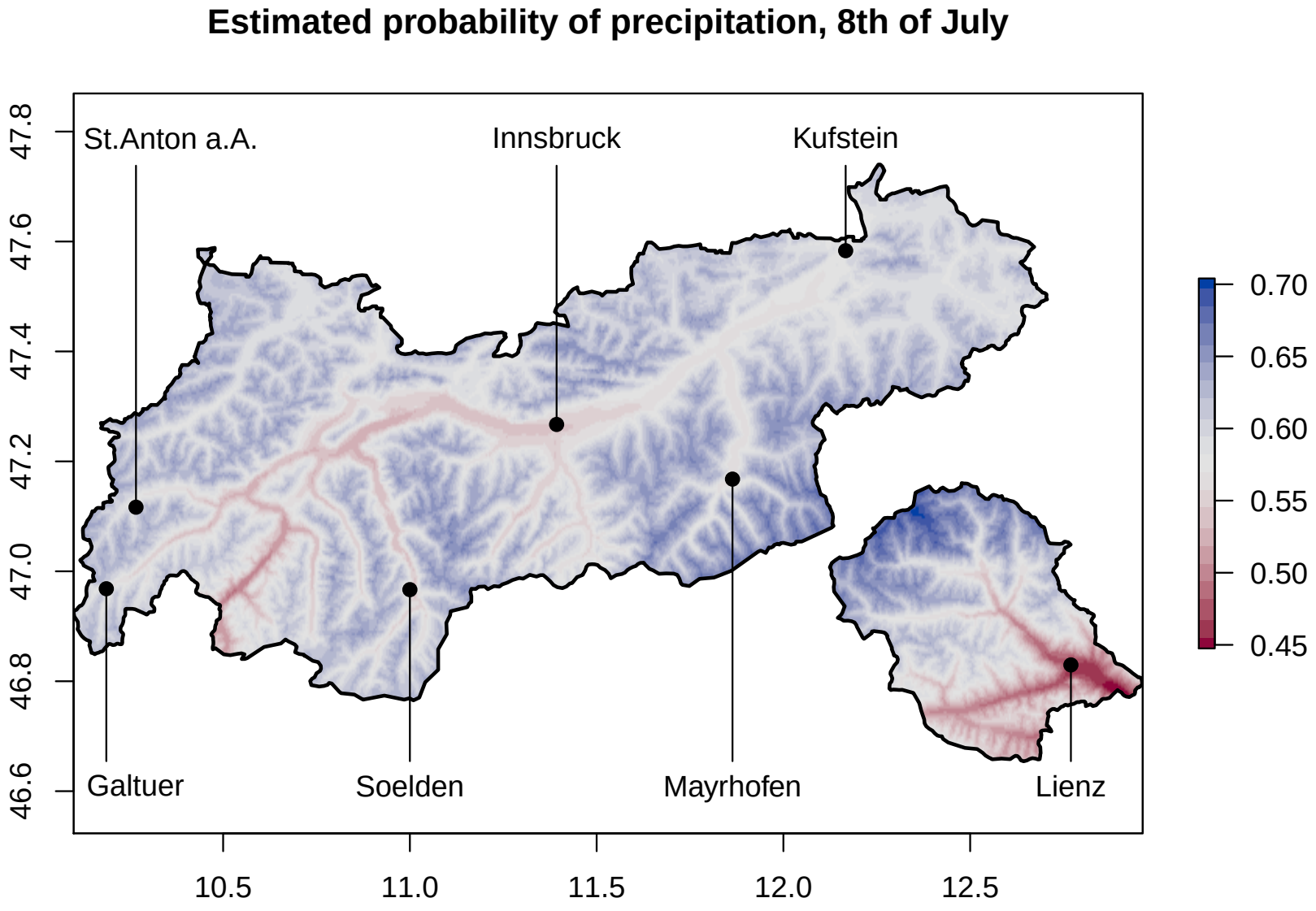
Introduction



Introduction

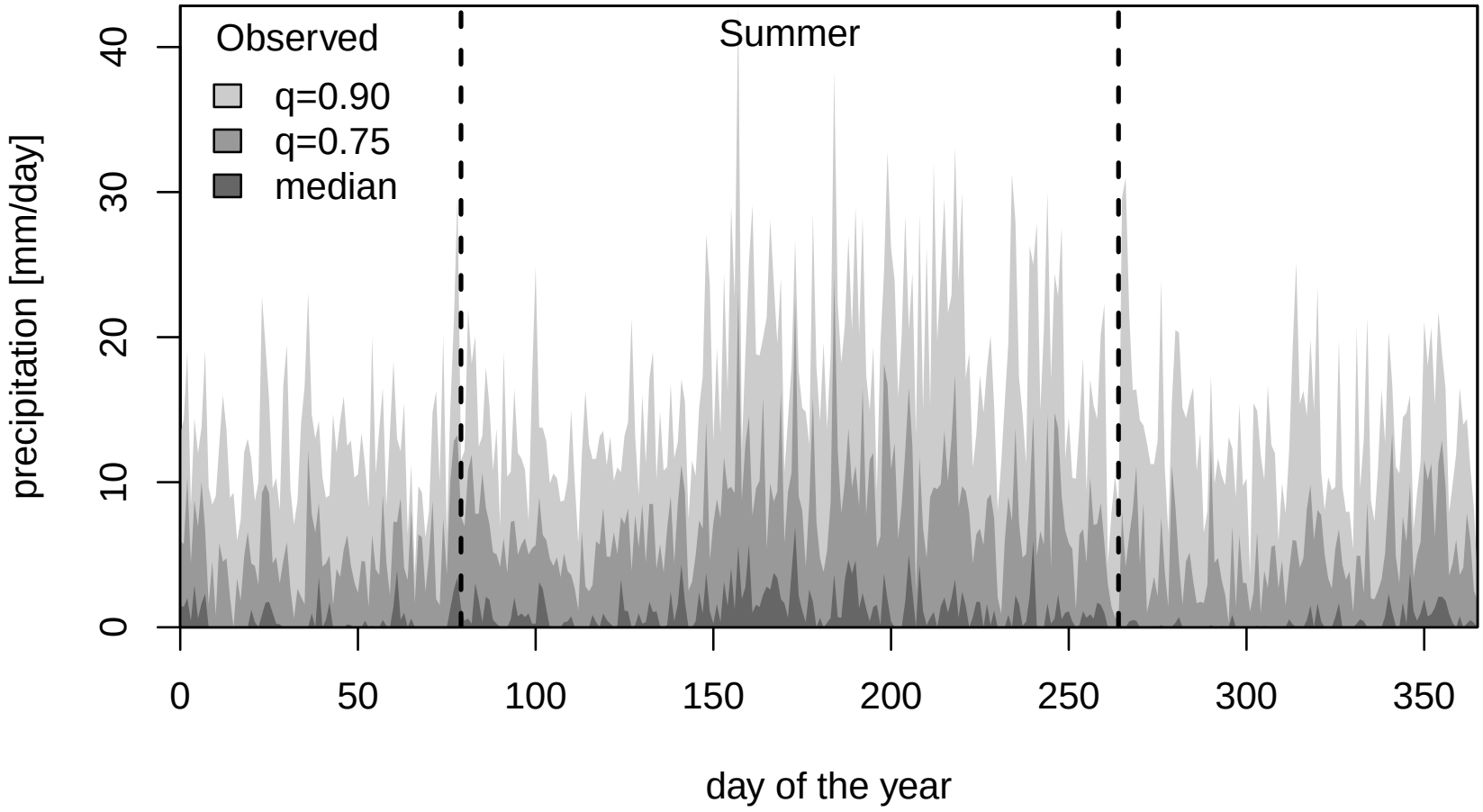


Introduction



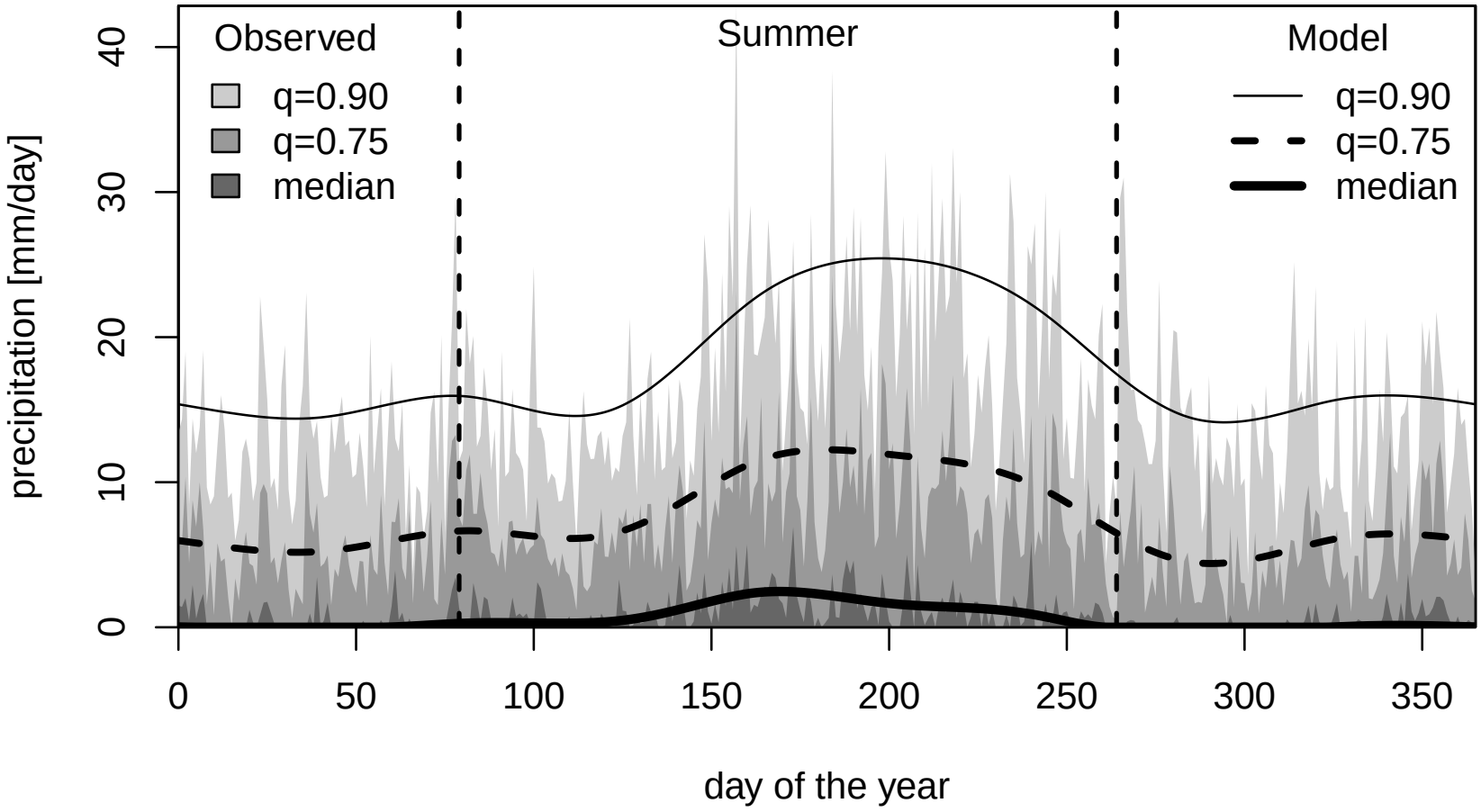
Introduction

Jungholz



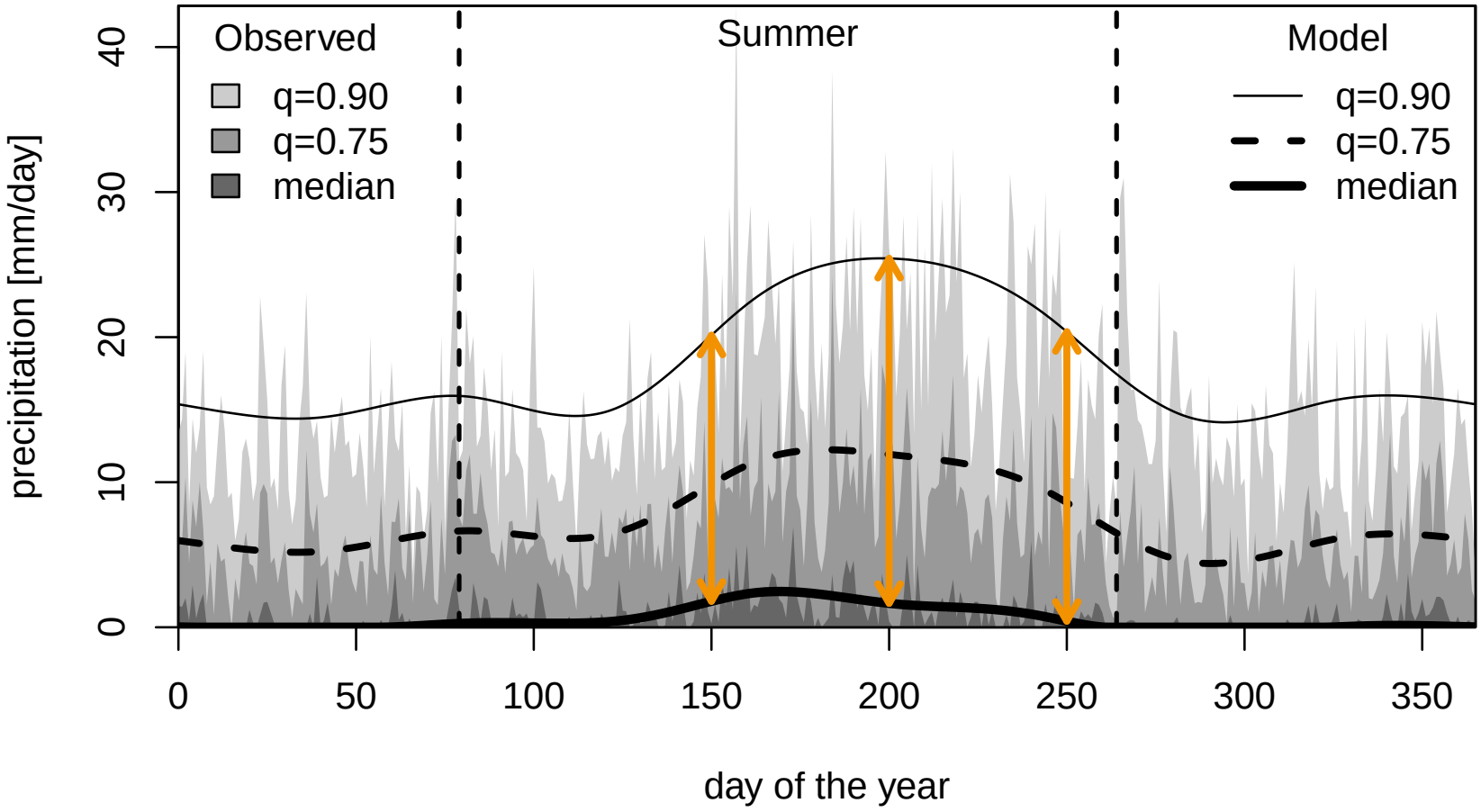
Introduction

Jungholz



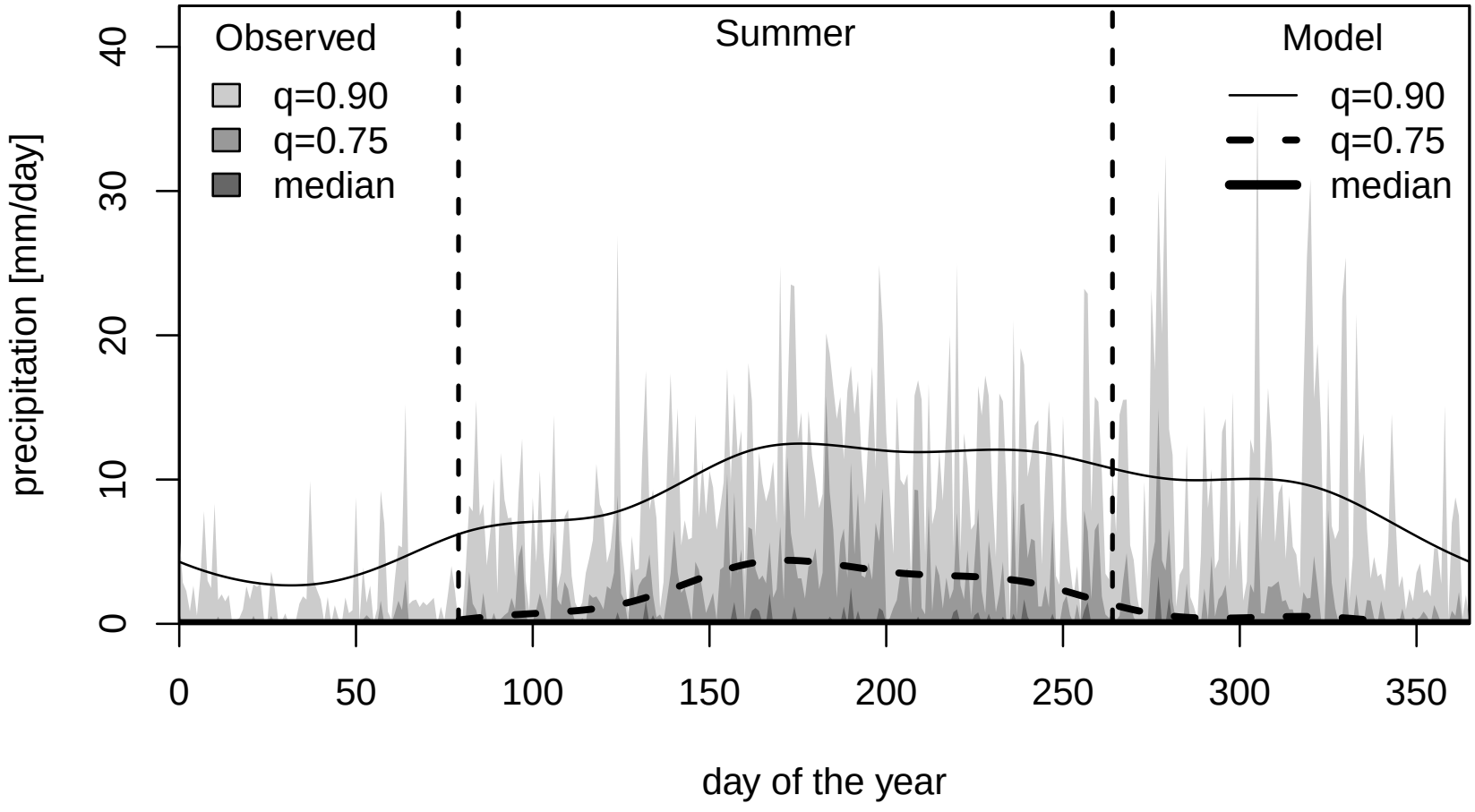
Introduction

Jungholz



Introduction

Nikolsdorf



Introduction

Our goal: precipitation climatology with . . .

Introduction

Our goal: precipitation climatology with . . .

- scalable spatial resolution (station $\Leftrightarrow$ grid)

Introduction

Our goal: precipitation climatology with . . .

- scalable spatial resolution (station $\Leftrightarrow$ grid)
- **daily** resolution

Introduction

Our goal: precipitation climatology with . . .

- scalable spatial resolution (station $\Leftrightarrow$ grid)
- **daily** resolution
- **full** climatological **distribution**

Introduction

Our goal: precipitation climatology with . . .

- scalable spatial resolution (station $\Leftrightarrow$ grid)
- **daily** resolution
- **full** climatological **distribution**
- **no extensive** or **manual tuning** for new areas

Introduction

Our goal: precipitation climatology with . . .

- scalable spatial resolution (station $\Leftrightarrow$ grid)
- **daily** resolution
- **full** climatological **distribution**
- **no extensive** or **manual tuning** for new areas

Applications

- alpine risk assessment
- tourism

Introduction

Our goal: precipitation climatology with . . .

- scalable spatial resolution (station $\Leftrightarrow$ grid)
- **daily** resolution
- **full** climatological **distribution**
- **no extensive** or **manual tuning** for new areas

Applications

- alpine risk assessment
- tourism
- **improving** snow **forecasts**

Strategy

Strategy

Model setup

- **generalized additive model** for μ and σ^2 (GAMLSS)

Strategy

Model setup

- **generalized additive model** for μ and σ^2 (GAMLSS)
- **altitude**: linear effect

Strategy

Model setup

- **generalized additive model** for μ and σ^2 (GAMLSS)
- **altitude**: linear effect
- **season**: cyclic cubic spline

Strategy

Model setup

- **generalized additive model** for μ and σ^2 (GAMLSS)
- **altitude**: linear effect
- **season**: cyclic cubic spline
- **spatial**: two dimensional thin-plate splines

Strategy

Model setup

- **generalized additive model** for μ and σ^2 (GAMLSS)
- **altitude**: linear effect
- **season**: cyclic cubic spline
- **spatial**: two dimensional thin-plate splines
- **spatially variable season**: spatial variability on seasonal pattern

Strategy

Model setup

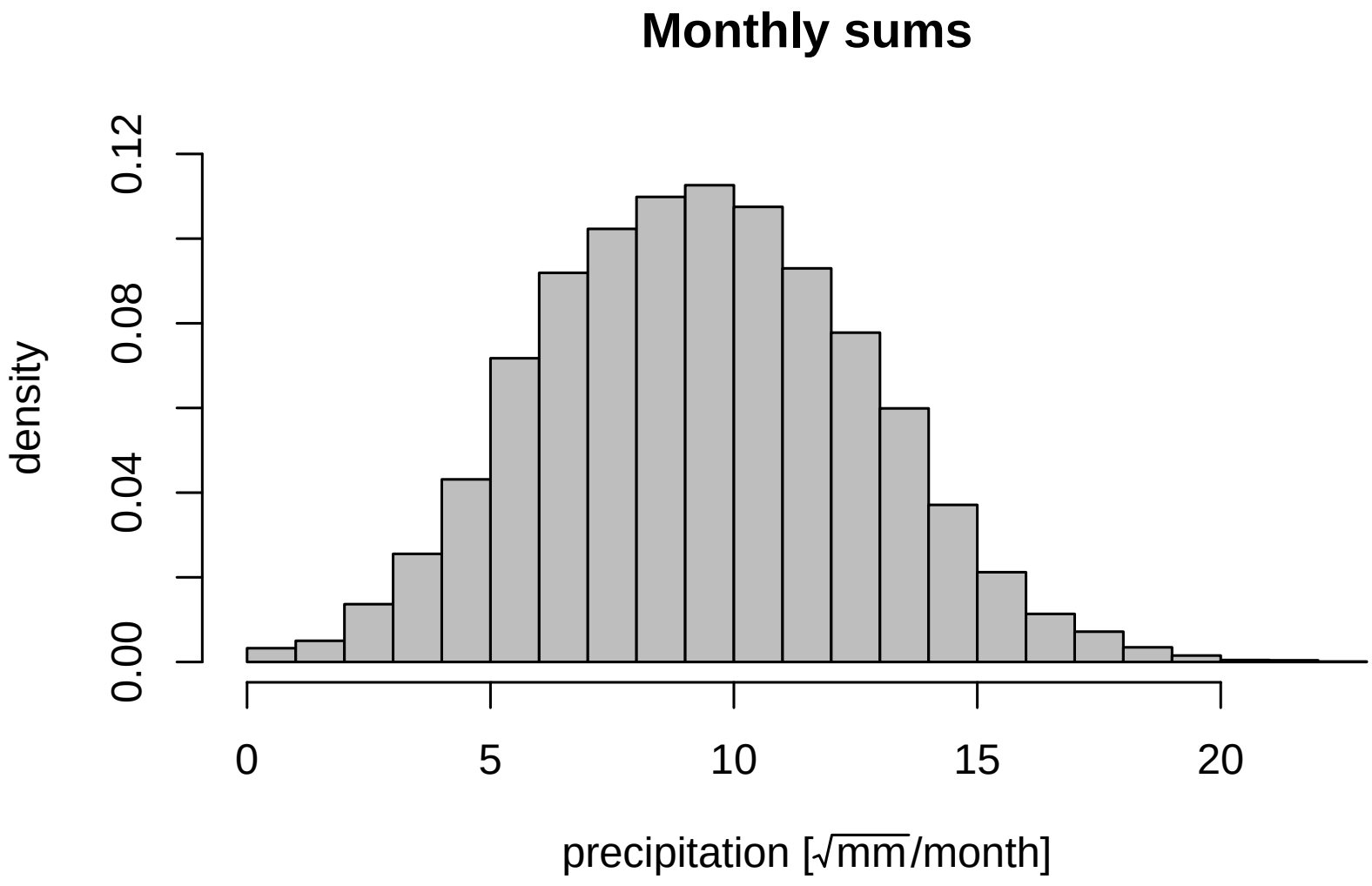
- **generalized additive model** for μ and σ^2 (GAMLSS)
- **altitude**: linear effect
- **season**: cyclic cubic spline
- **spatial**: two dimensional thin-plate splines
- **spatially variable season**: spatial variability on seasonal pattern

Open question

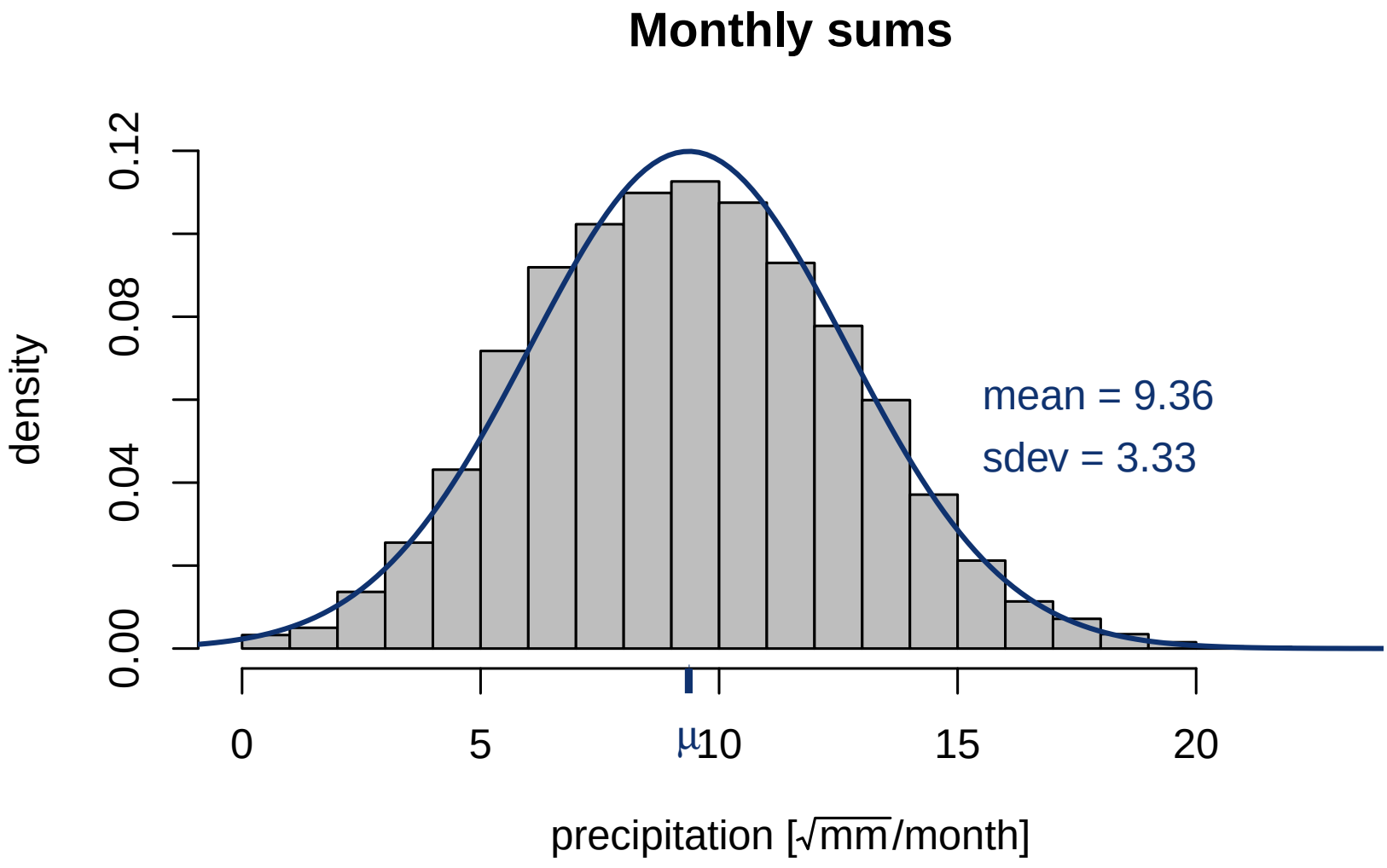
- which distribution?

  $y \sim \mathcal{N}(\mu, \sigma^2)$ not suitable for all aggregation levels

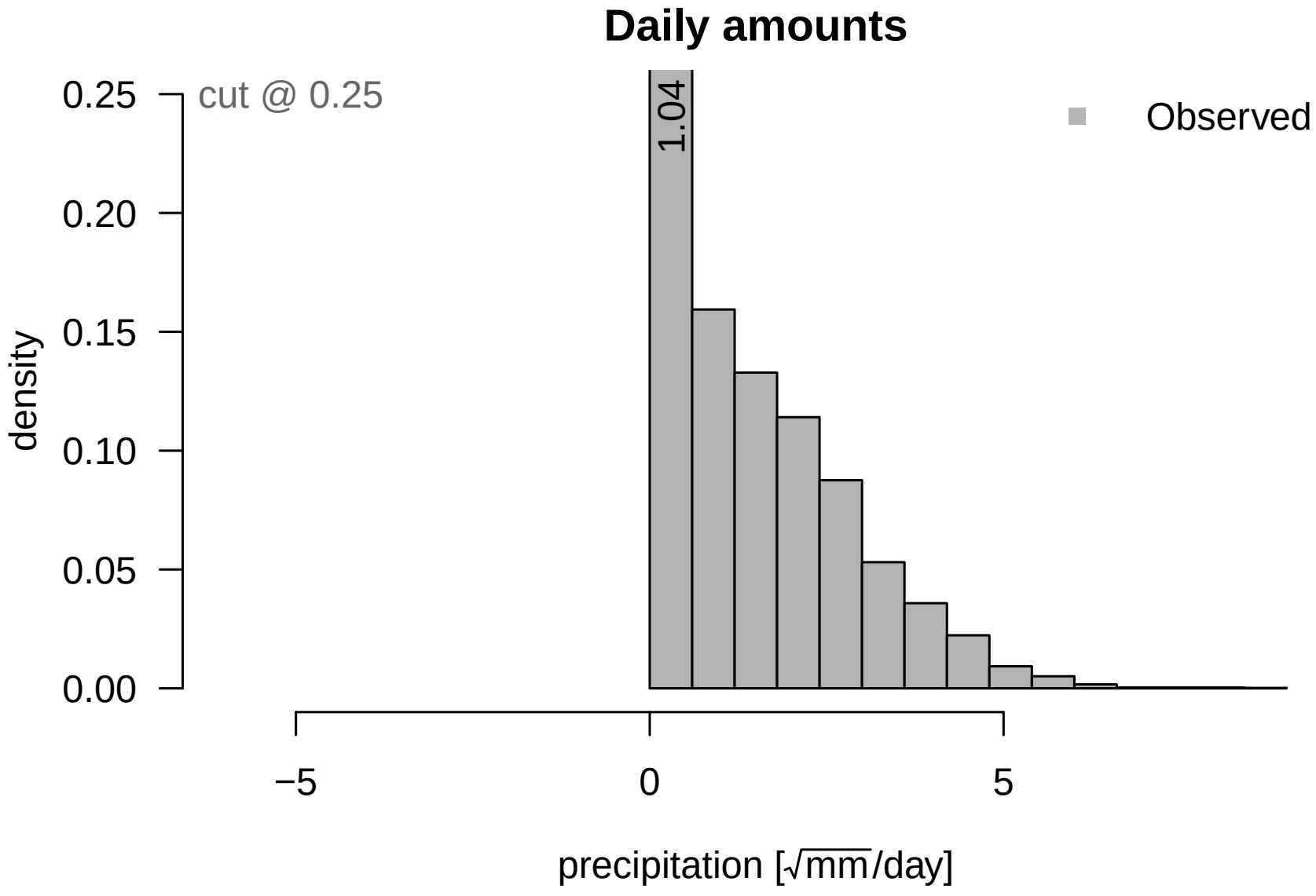
The data: monthly sums



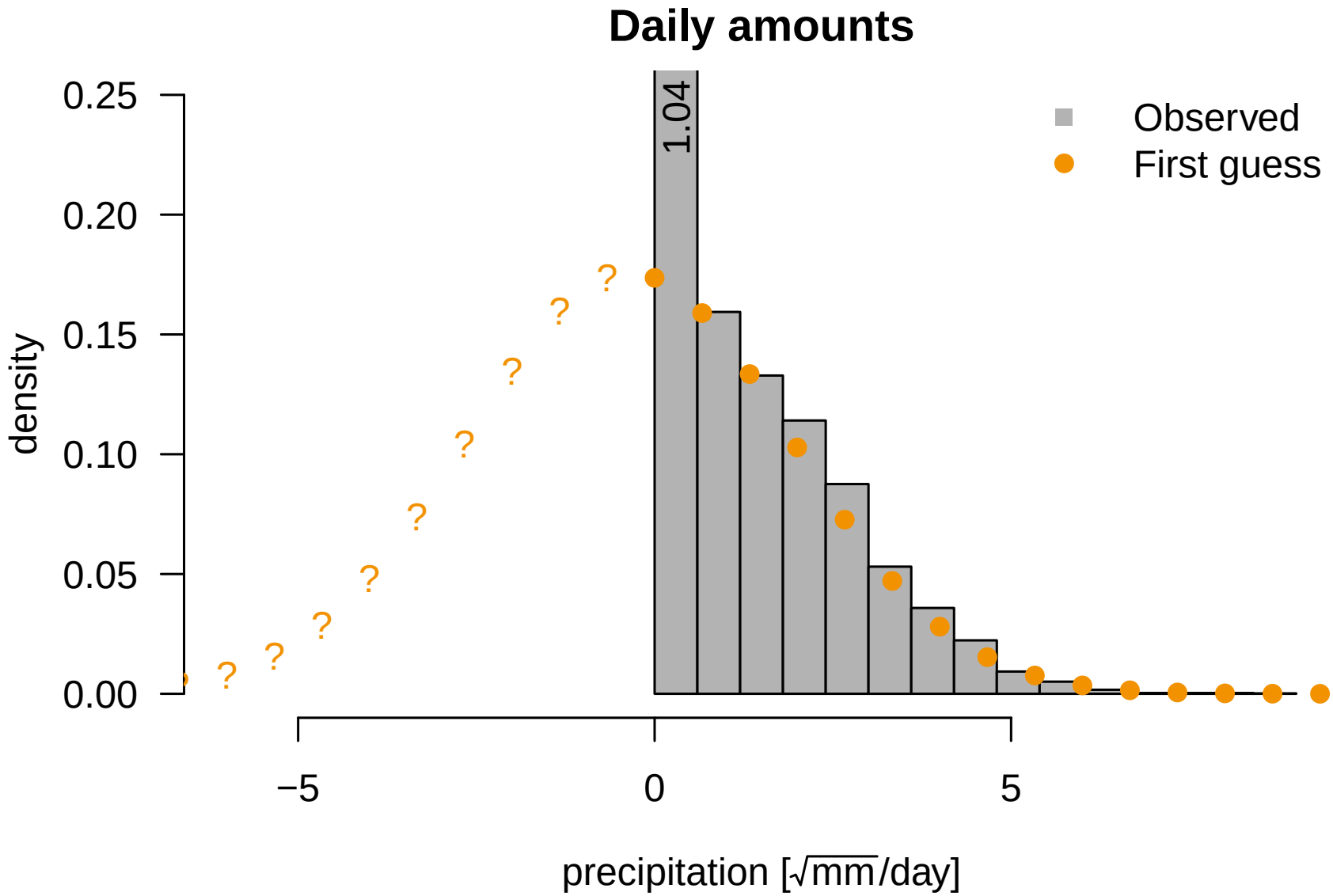
The data: monthly sums



The data: daily amounts



The data: daily amounts



Censoring

Can be seen as censored if . . .

- **limited** to a threshold
- values exceeding threshold **cannot occur** or **not observable**

Censoring

Can be seen as censored if . . .

- **limited** to a threshold
- values exceeding threshold **cannot occur** or **not observable**

Examples

- survival in 5 *yr* clinical study: **right** $y_i \leqslant 5$
- hours worked this week: **two sided** $0 \leqslant y_i \leqslant 168$

Censoring

Can be seen as censored if . . .

- **limited** to a threshold
- values exceeding threshold **cannot occur** or **not observable**

Examples

- survival in 5 *yr* clinical study: **right** $y_i \leqslant 5$
- hours worked this week: **two sided** $0 \leqslant y_i \leqslant 168$
- precipitation: **left** $0 \leqslant y_i$

Censoring

Can be seen as censored if . . .

- **limited** to a threshold
- values exceeding threshold **cannot occur** or **not observable**

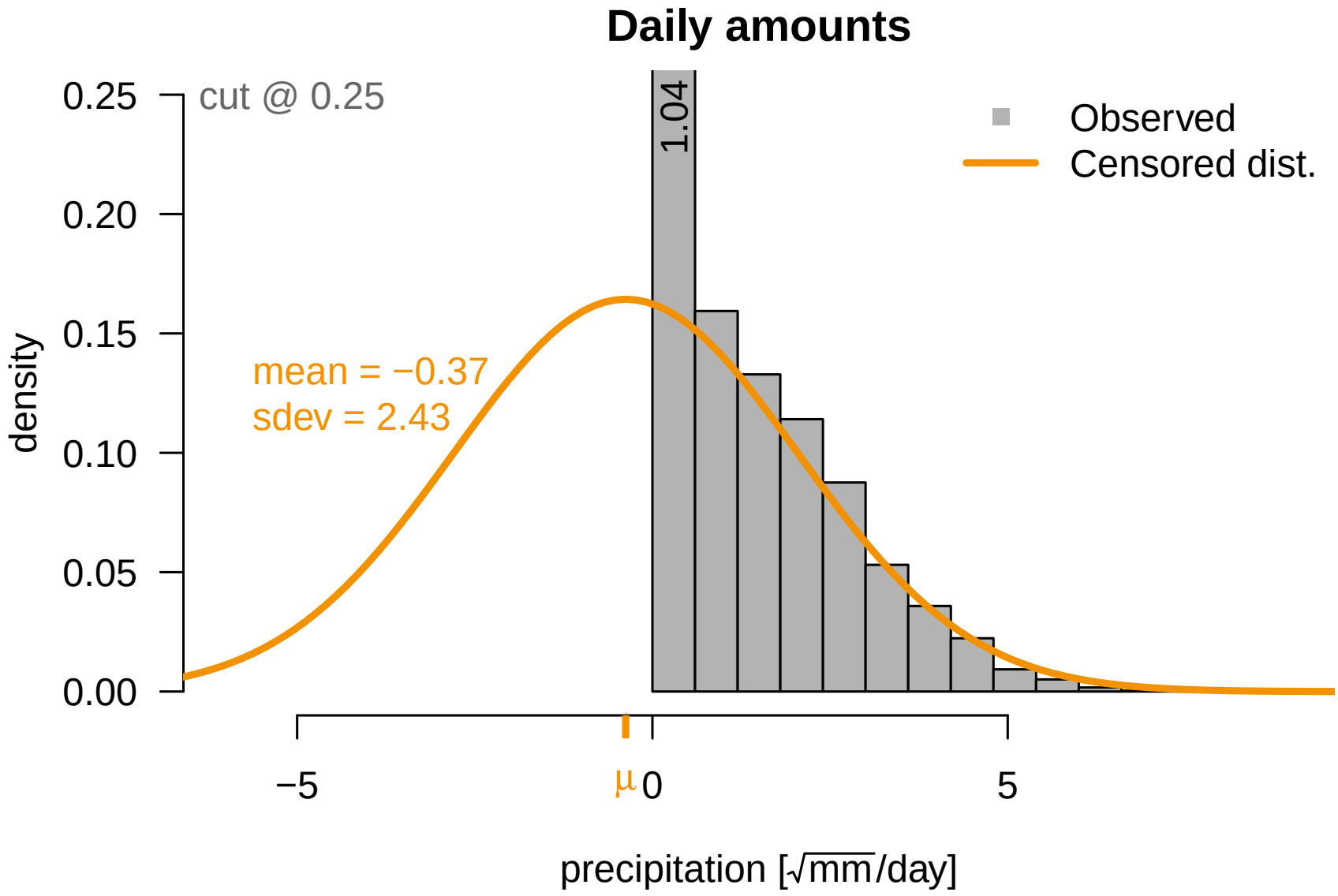
Examples

- survival in 5 *yr* clinical study: **right** $y_i \leqslant 5$
- hours worked this week: **two sided** $0 \leqslant y_i \leqslant 168$
- precipitation: **left** $0 \leqslant y_i$

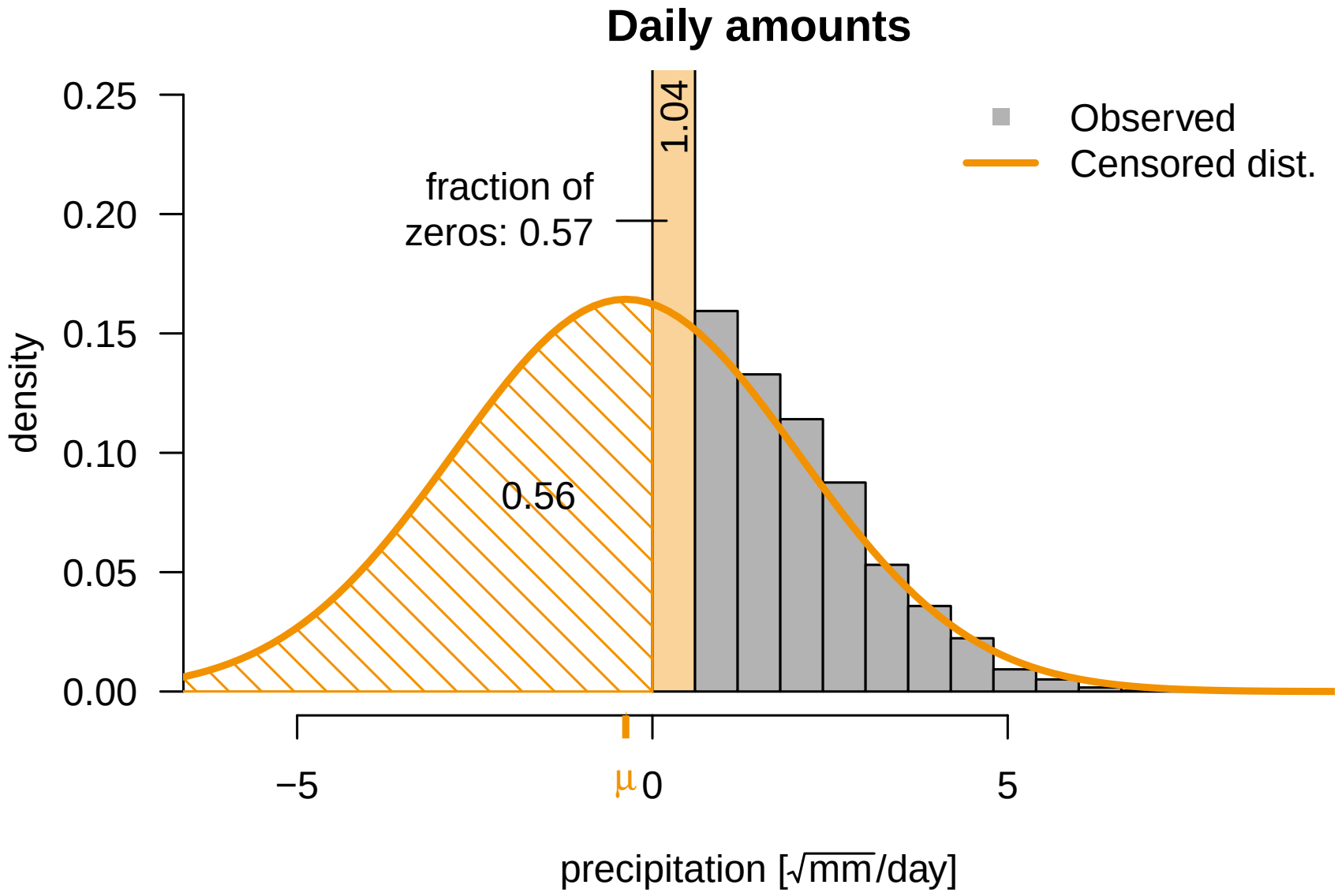
Left censored Gaussian distribution

$$y^* \sim \mathcal{N}(\mu, \sigma^2), \quad y = \max(\tau = 0, y^*)$$

The data: daily amounts



The data: daily amounts



Model setup

Assumptions

$$y \sim \mathcal{D}(\mu, \sigma^2), \quad \mu = \eta_\mu, \quad \log(\sigma) = \eta_\sigma$$

y	response; daily precipitation (transformed)
$\mathcal{D}(\dots)$	left censored Gaussian distribution on 0
μ/σ	location/scale

Model setup

Assumptions

$$y \sim \mathcal{D}(\mu, \sigma^2), \quad \mu = \eta_\mu, \quad \log(\sigma) = \eta_\sigma$$

Linear predictor

$$\eta = \beta_0 + \underbrace{\beta_1 \text{ alt}}_{\text{altitude}} + \underbrace{f_1(\text{yday})}_{\text{seasonal}} + \underbrace{f_2(\text{long}, \text{lat})}_{\text{spatial}} + \underbrace{f_3(\underline{\text{yday}}, \text{long}, \text{lat})}_{\text{spatial season}}$$

y	response; daily precipitation (transformed)
$\mathcal{D}(\dots)$	left censored Gaussian distribution on 0
μ/σ	location/scale
η	linear predictor; identical for μ and $\log(\sigma)$

Model and Results

Model and Results

Model estimation

- **R** package **bamlss** (Umlauf et al. (2015); Poster 72)
- censored Gaussian distribution

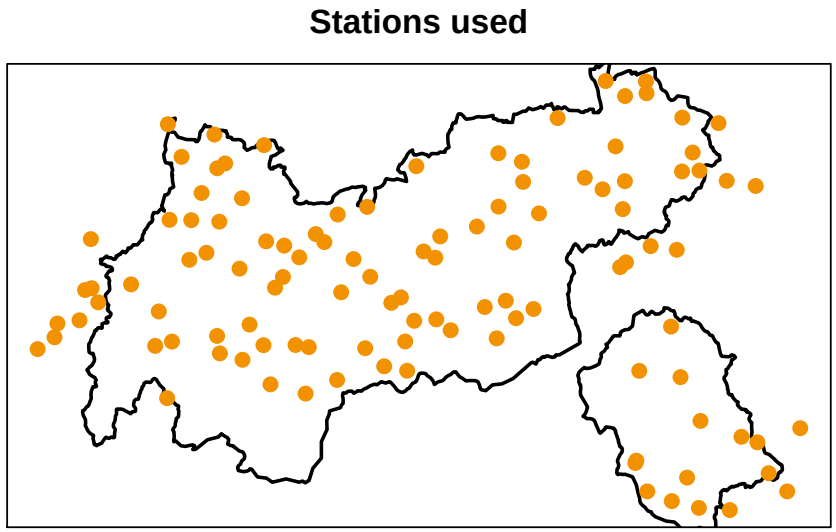
Model and Results

Model estimation

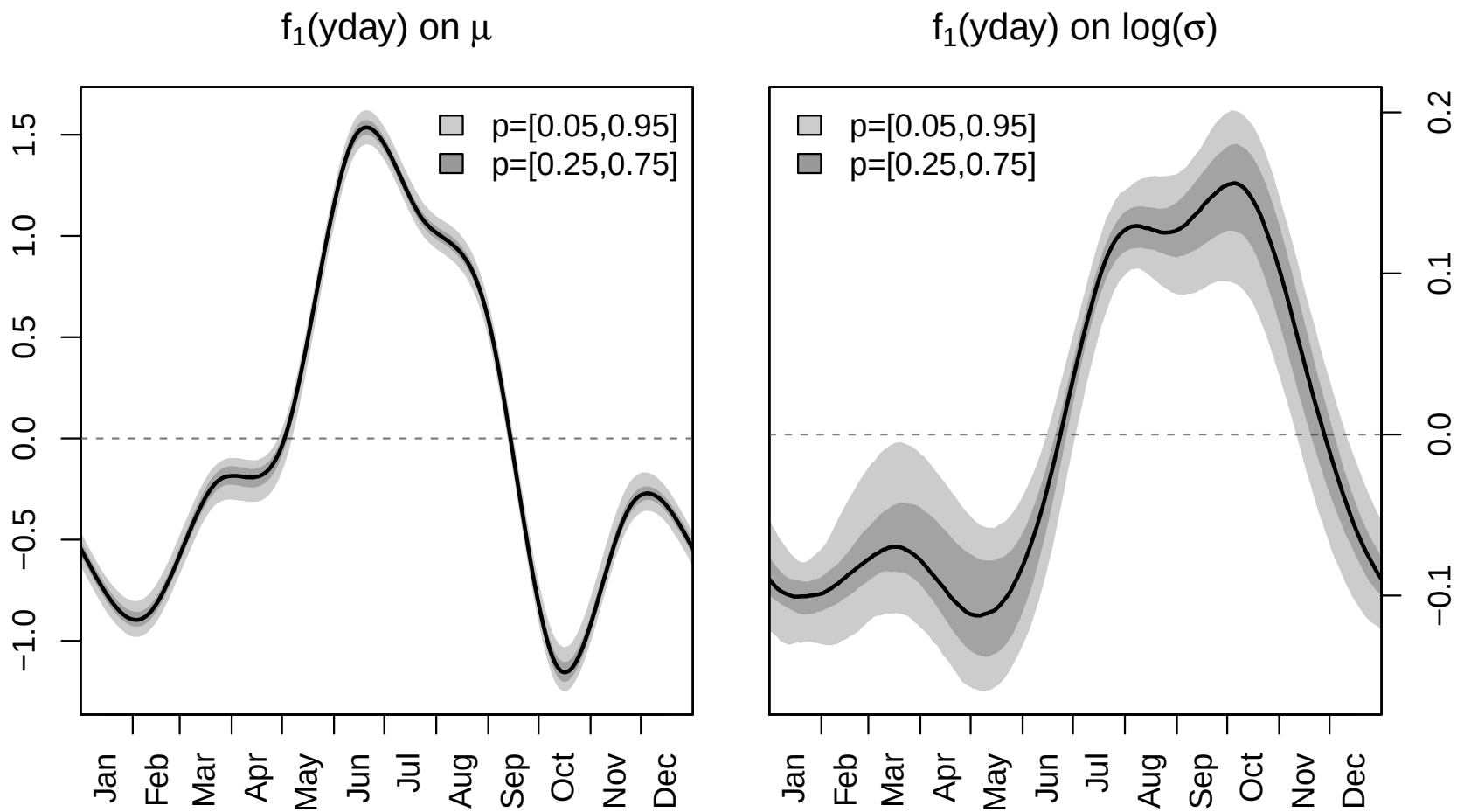
- **R** package **bamlss** (Umlauf et al. (2015); Poster 72)
- censored Gaussian distribution

Data set

- 110 stations
- 41 *years*
- daily rain/snow amounts
- $\sim$ 1.5 million observations

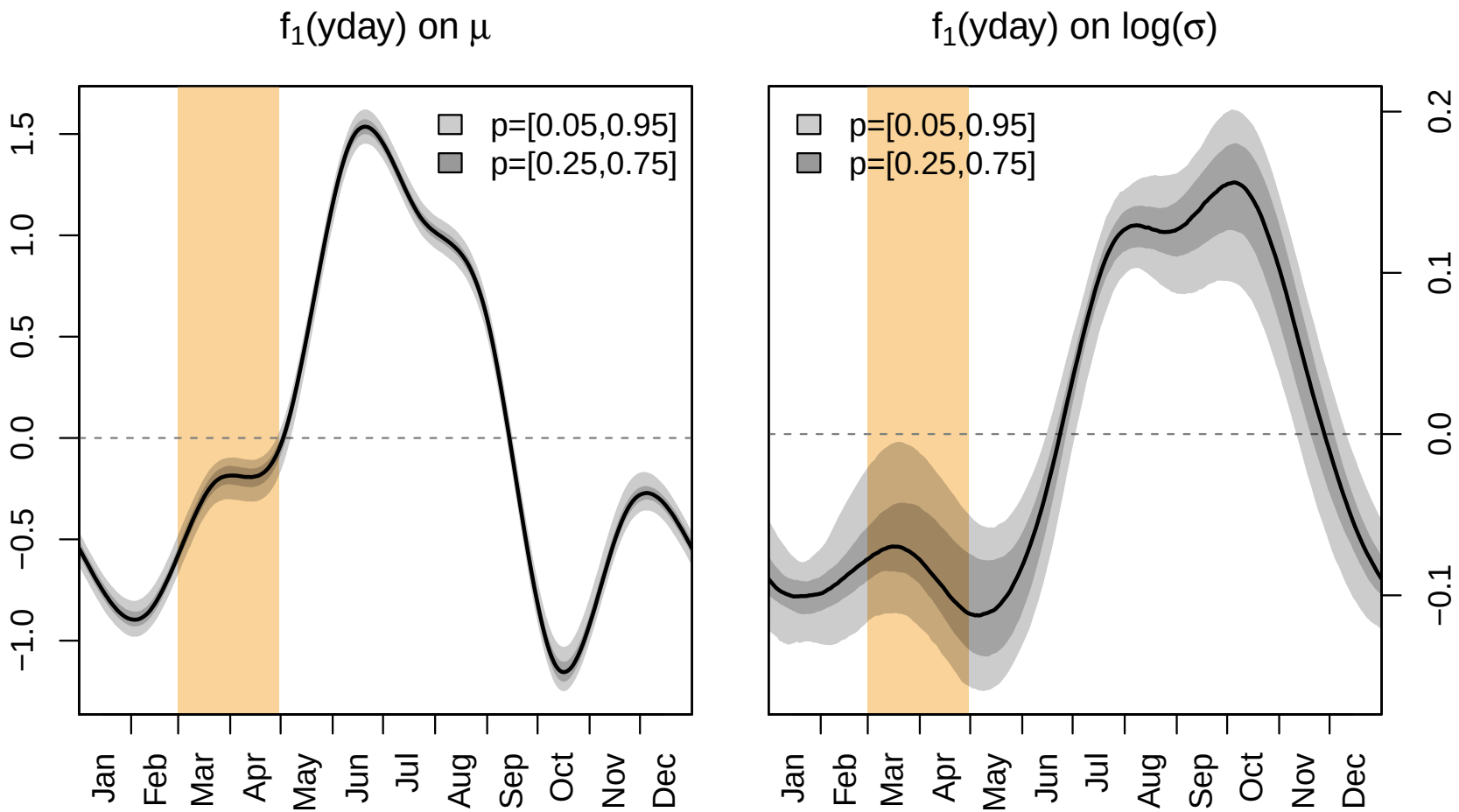


Model and Results



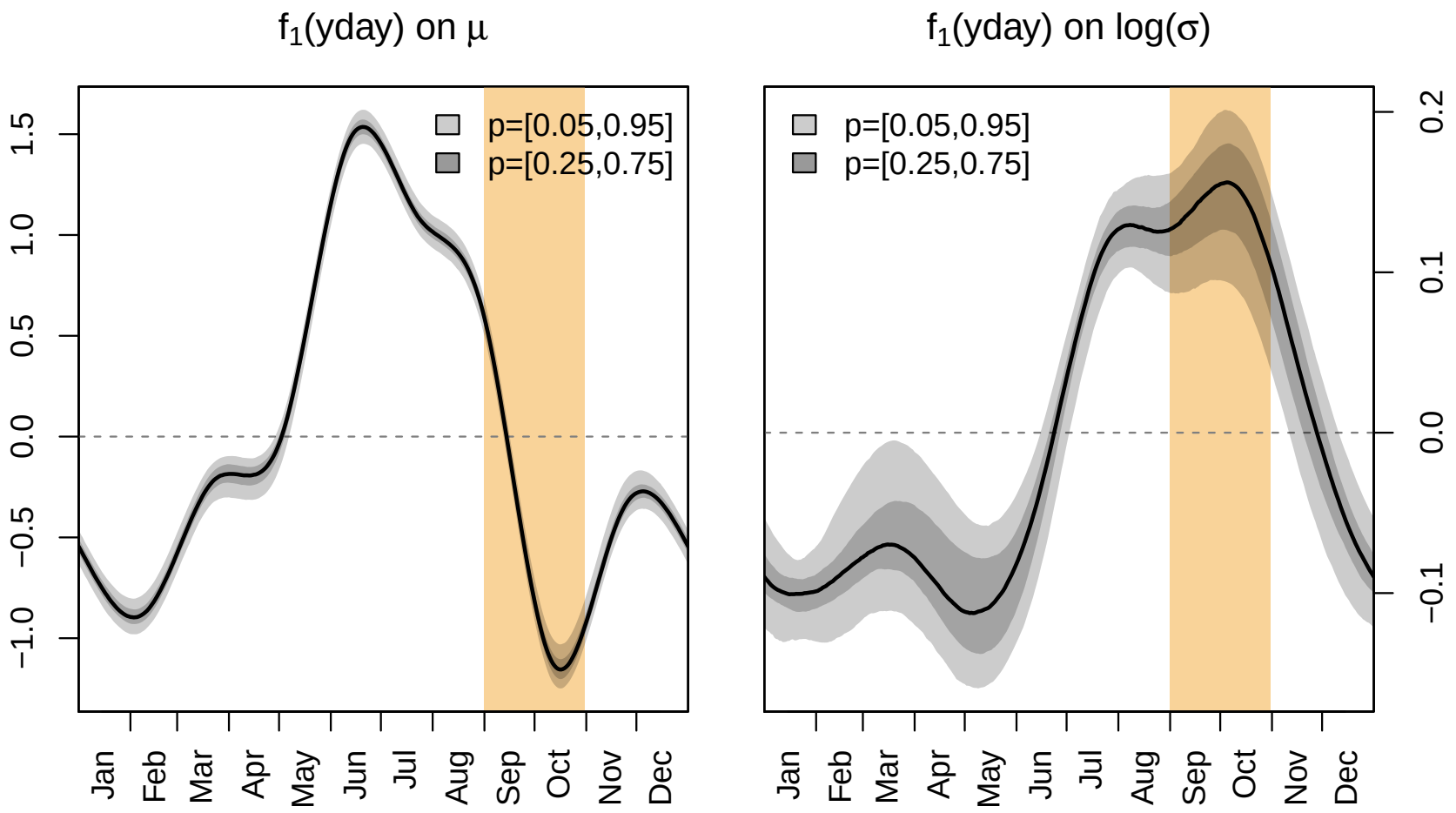
$$\eta = \beta_0 + \beta_1 \text{ alt} + \underbrace{f_1(\text{yday})}_{\textit{season}} + f_2(\text{long}, \text{lat}) + f_3(\text{yday}, \text{long}, \text{lat})$$

Model and Results



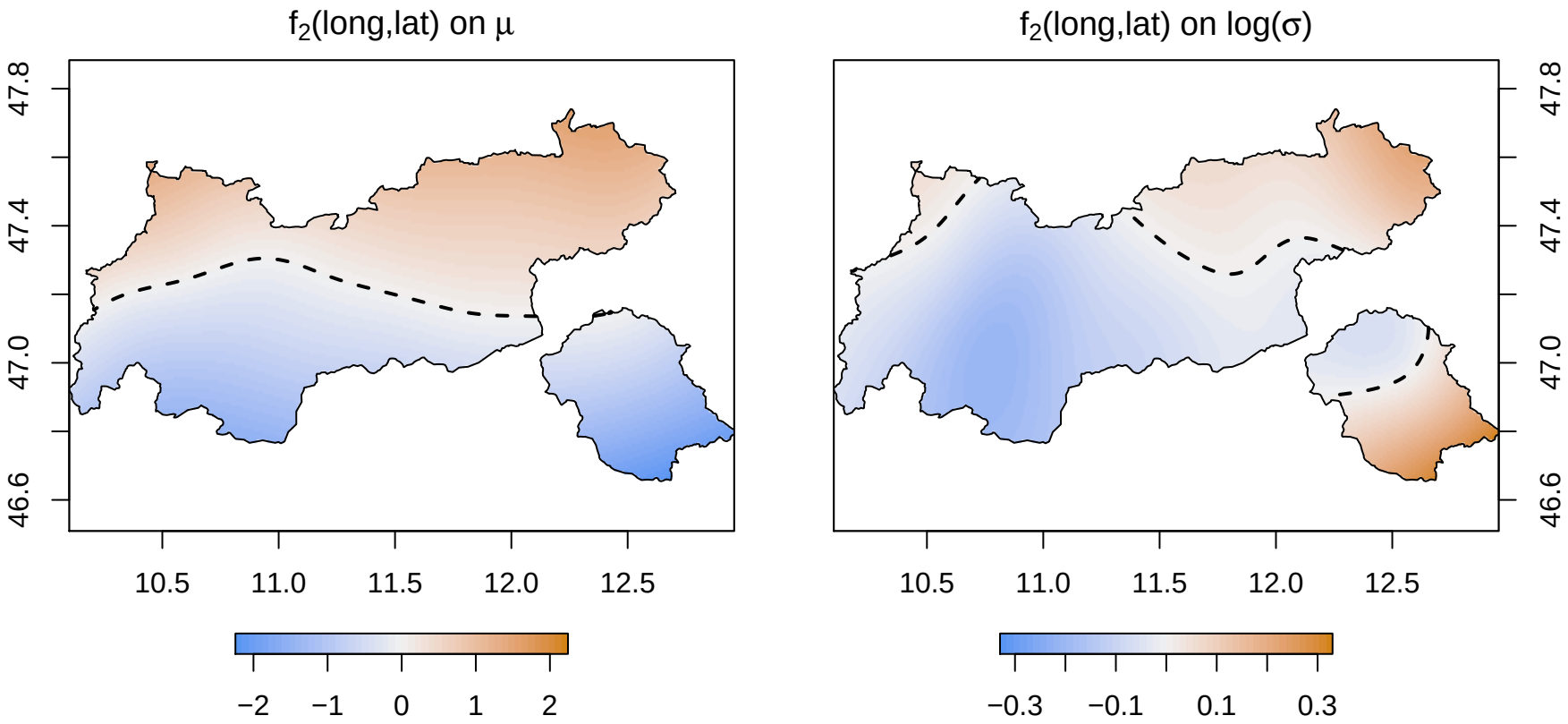
$$\eta = \beta_0 + \beta_1 \text{ alt} + \underbrace{f_1(\text{yday})}_{\textit{season}} + f_2(\text{long}, \text{lat}) + f_3(\text{yday}, \text{long}, \text{lat})$$

Model and Results



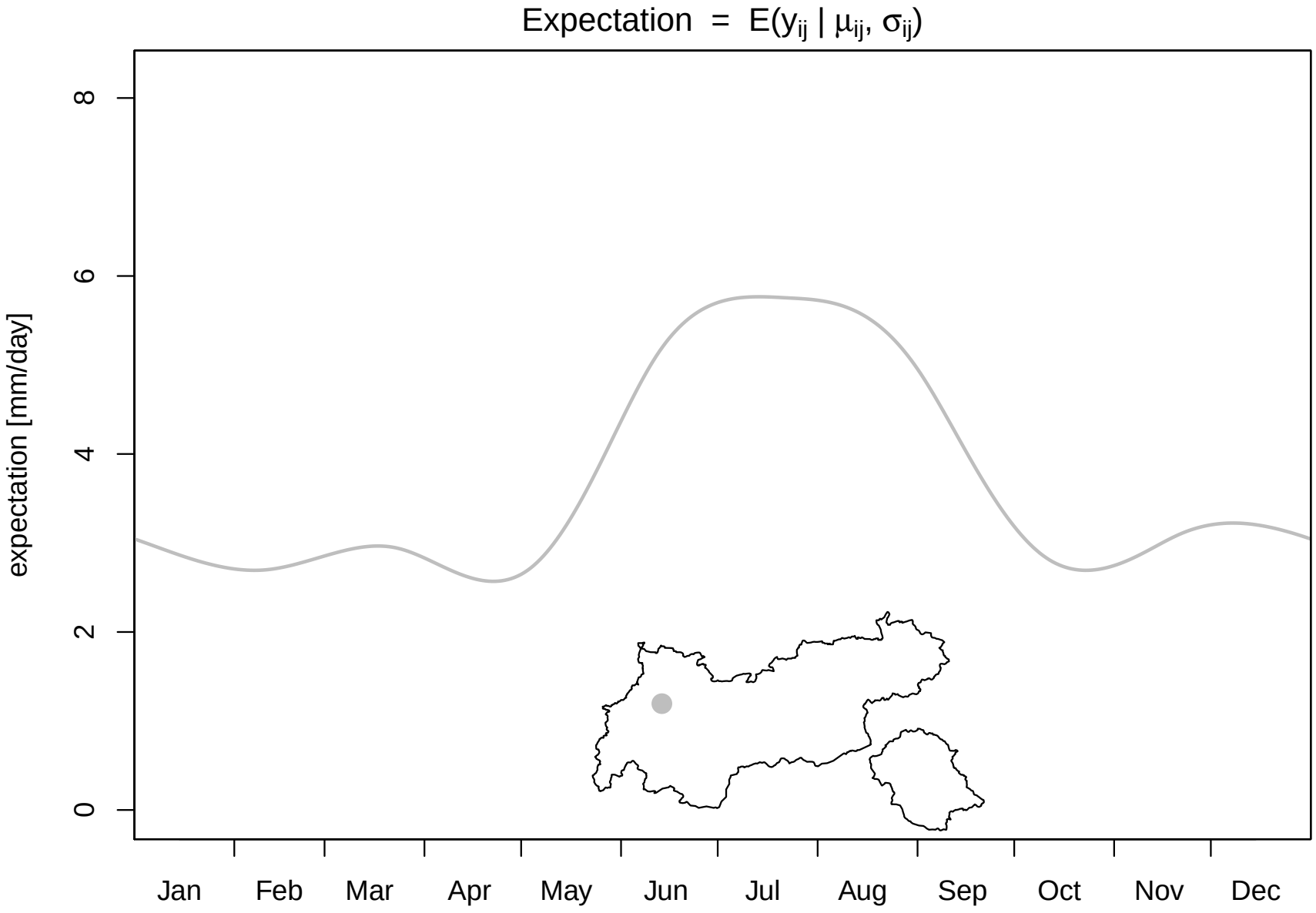
$$\eta = \beta_0 + \beta_1 \text{ alt} + \underbrace{f_1(\text{yday})}_{\textit{season}} + f_2(\text{long}, \text{lat}) + f_3(\text{yday}, \text{long}, \text{lat})$$

Model and Results

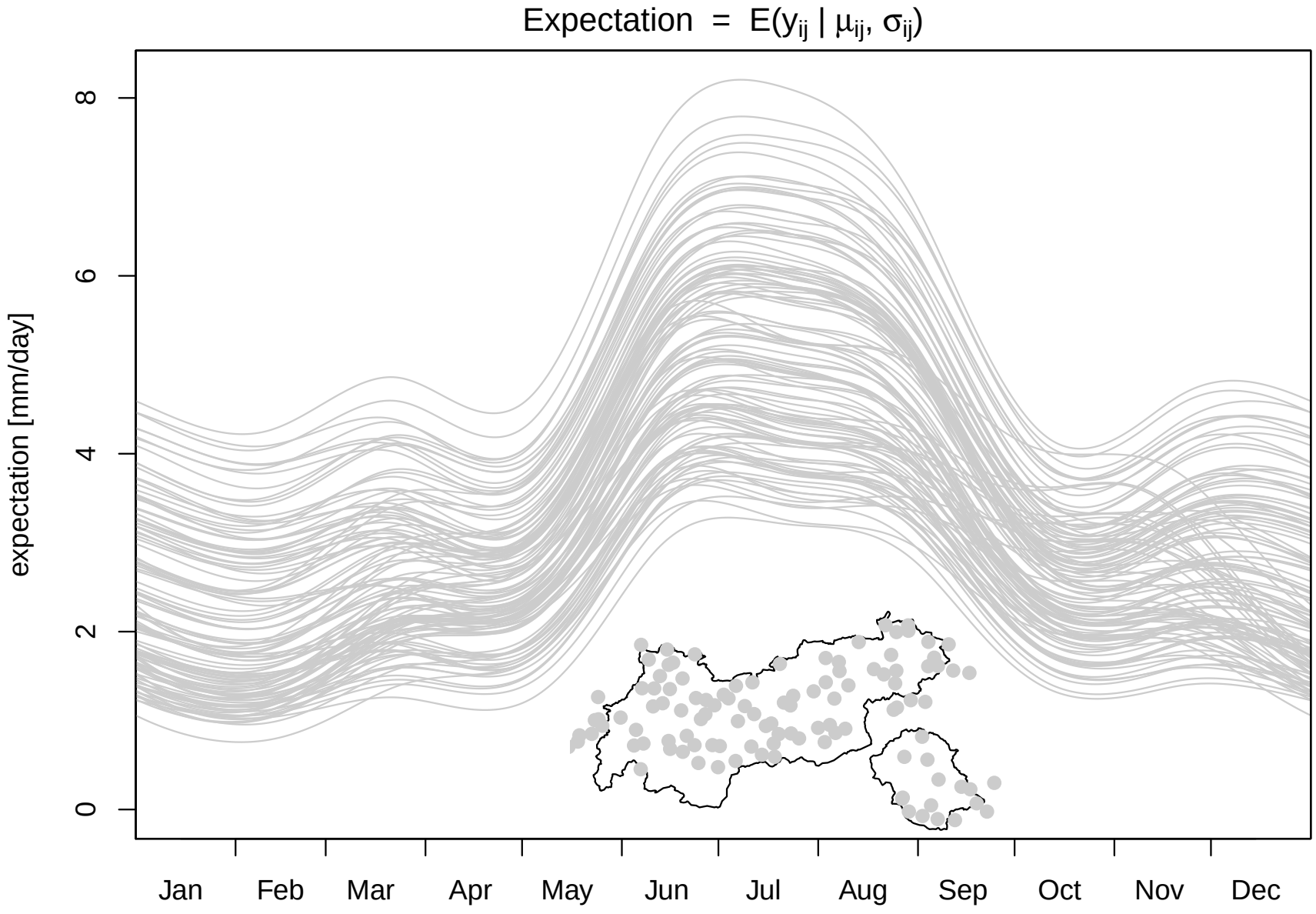


$$\eta = \beta_0 + \beta_1 \text{ alt} + f_1(\text{yday}) + \underbrace{f_2(\text{long}, \text{lat})}_{\textit{spatial}} + f_3(\text{yday}, \text{long}, \text{lat})$$

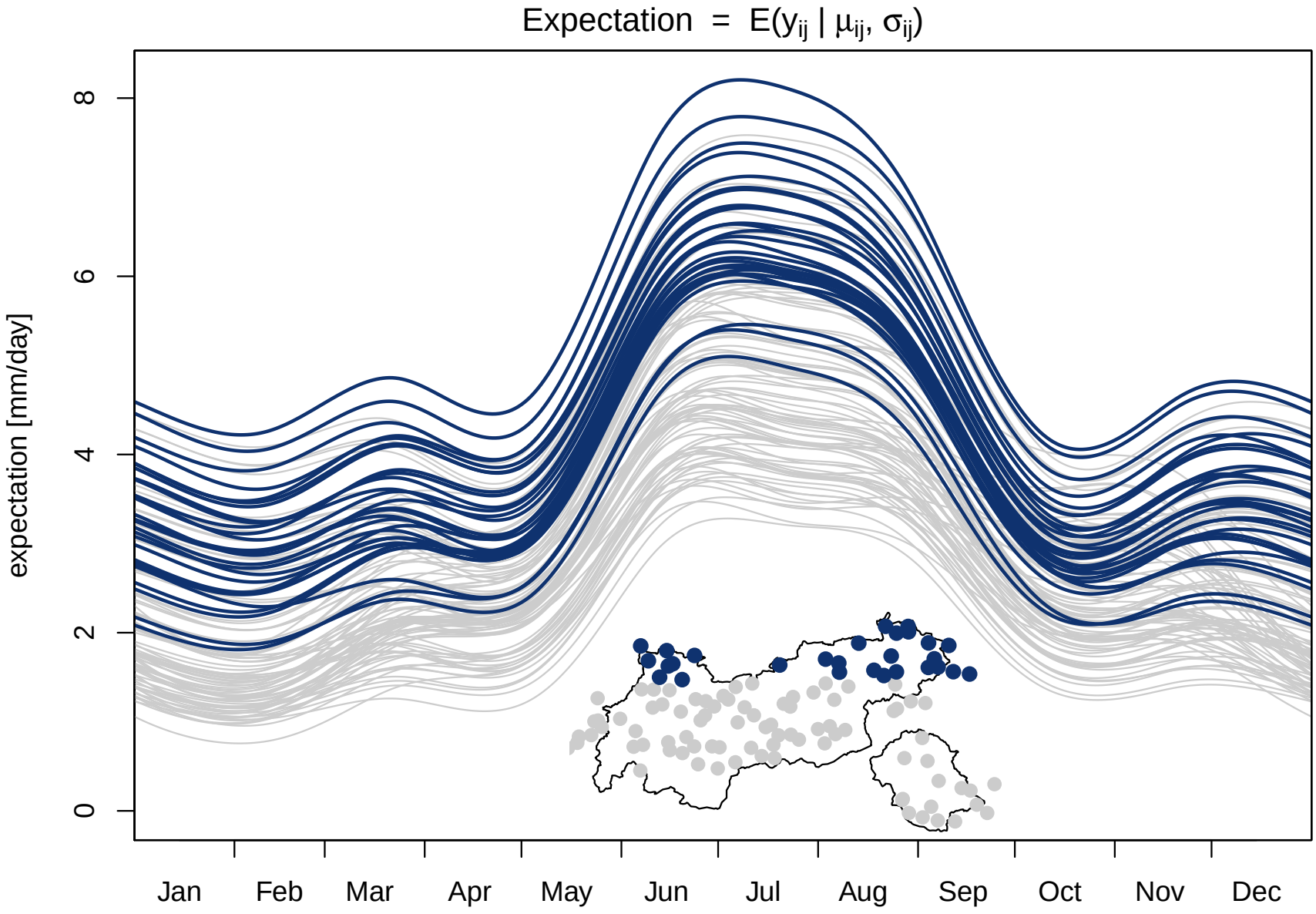
Model and Results



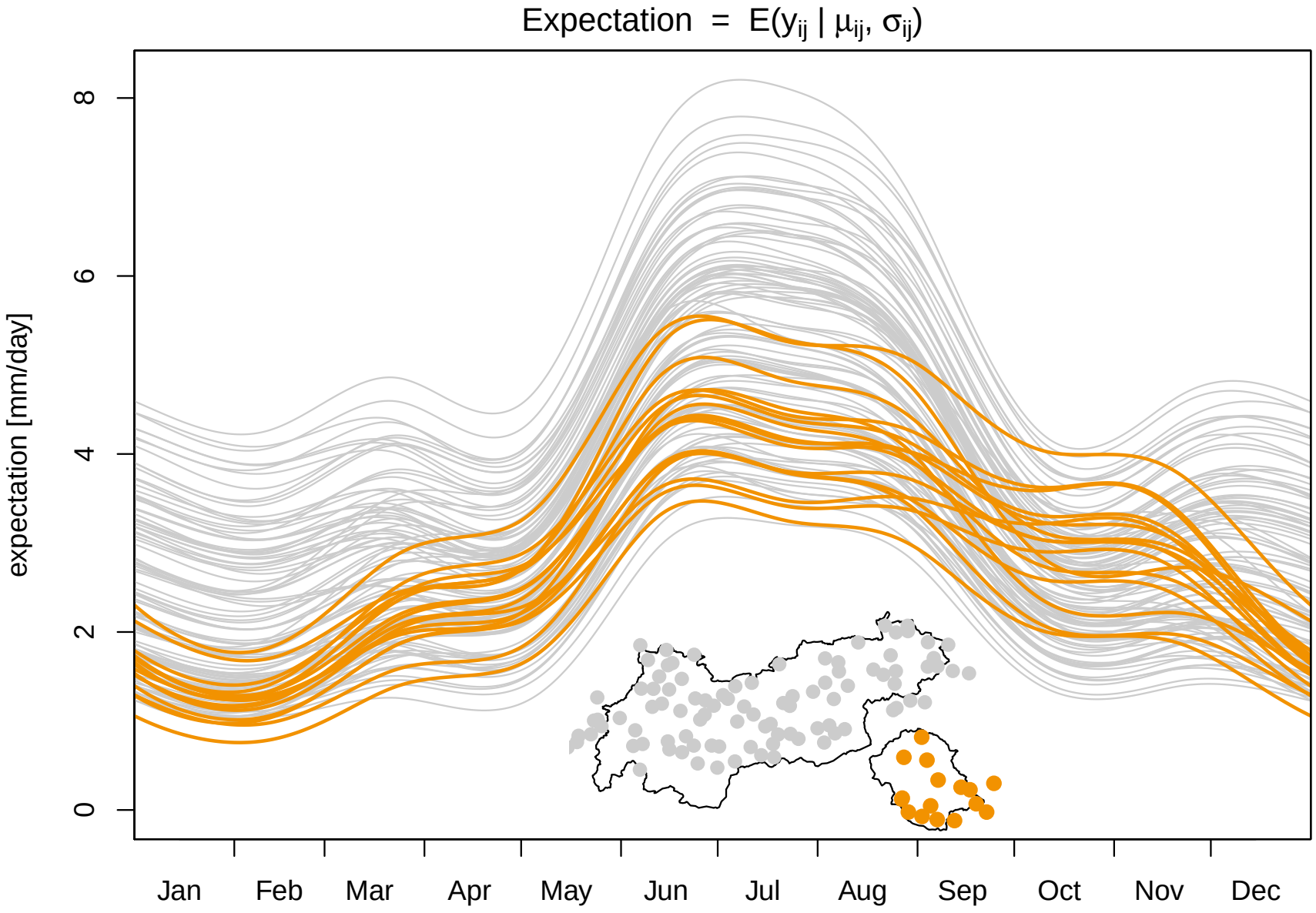
Model and Results



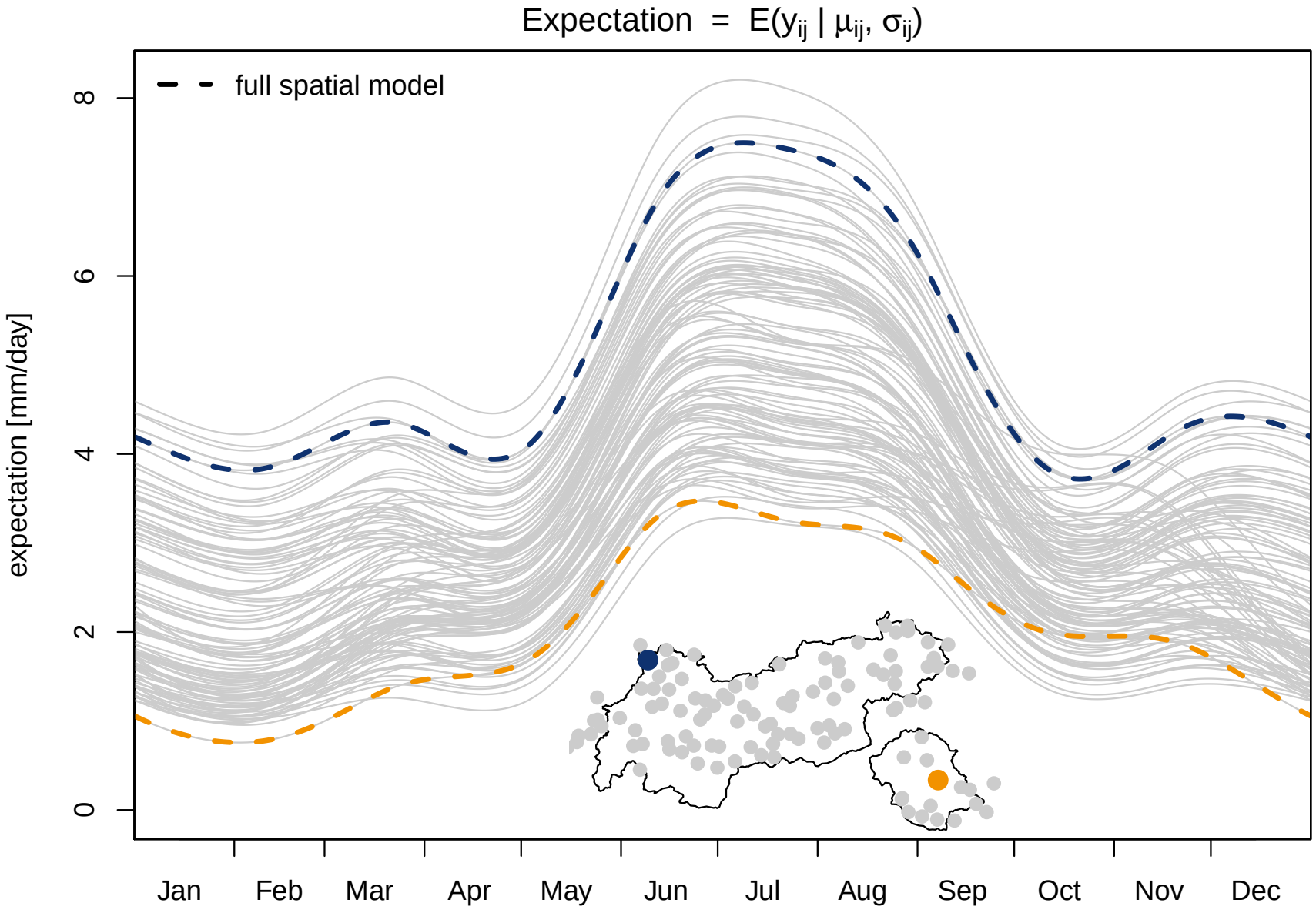
Model and Results



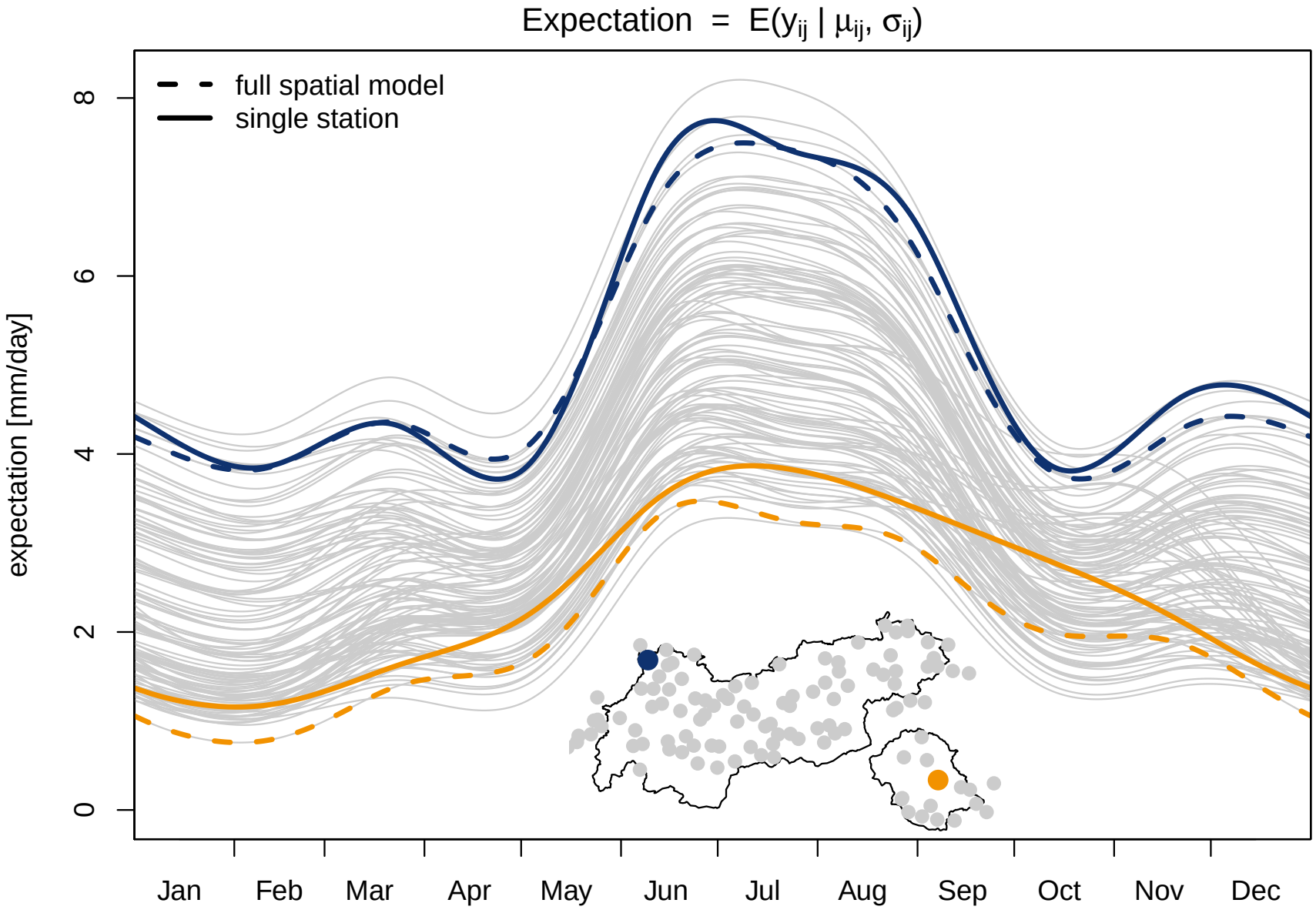
Model and Results



Model and Results

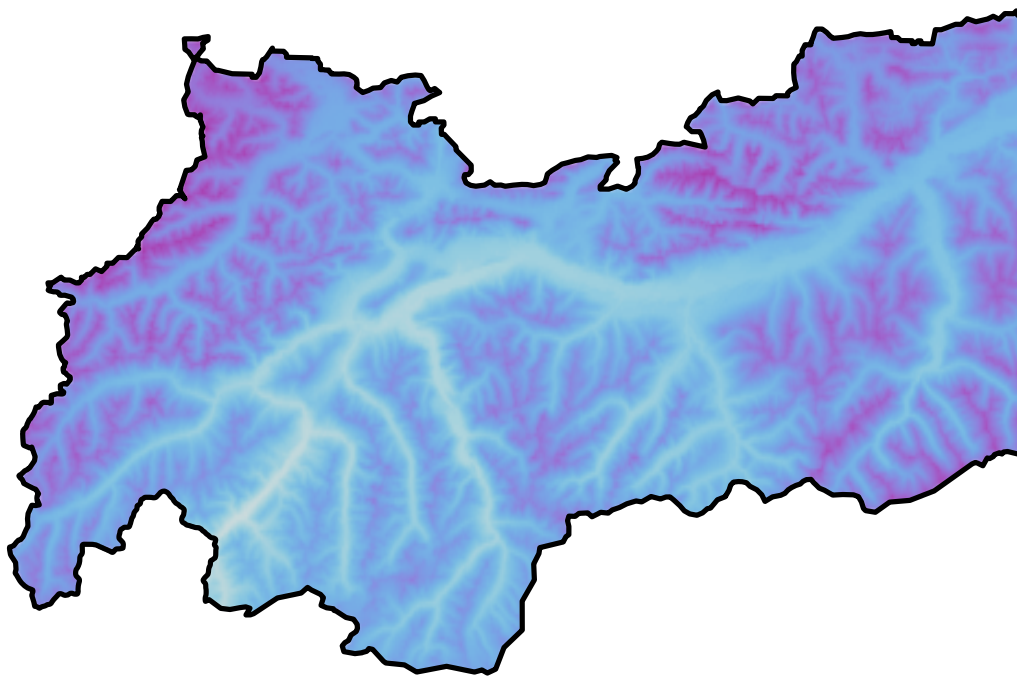


Model and Results



Summary

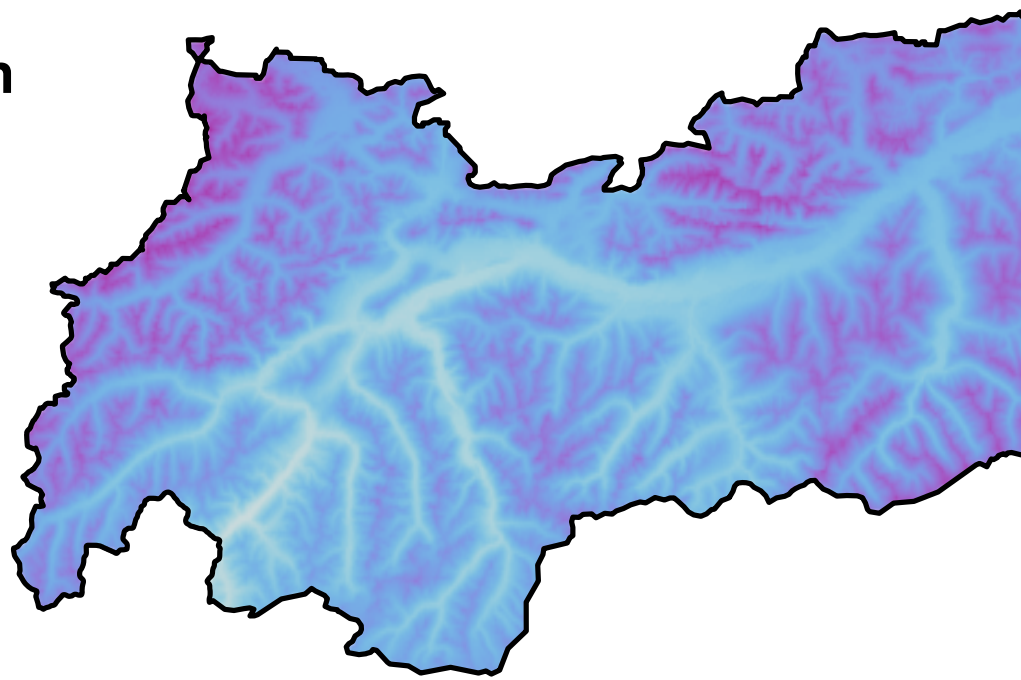
- “simple”, generalized setup
- full **spatio-temporal** model with **censored** response
- **spatial/temporal** resolution arbitrary scalable
- accurate estimate at station level



Summary

- “simple”, generalized setup
- full **spatio-temporal** model with **censored** response
- **spatial/temporal** resolution arbitrary scalable
- accurate estimate at station level

- handling zero-observations
- **full** climatological **distribution**
- quantiles
- probability of precipitation



Thank you for your attention!

Further details:

Scientific article in progress.

Contact:

Reto Stauffer, Department of Statistics, University of Innsbruck.
`reto.stauffer@uibk.ac.at`.

Acknowledgements:

Ongoing project funded by the Austrian Science Fund (FWF): TRP 290-N26.

The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC).

Data set provided by the “Ministerium für ein lebenswertes Österreich”, hydrographical service Tyrol (`ehyd.gv.at`).

References

Cohen, A. C. J., 1959: Simplified estimators for the normal distribution when samples are singly censored or truncated. *Technometrics*, **1 (3)**, 217–237.

Friederichs, P. and A. Hense, 2007: Statistical downscaling of extreme precipitation events using censored quantile regression. *Monthly Weather Review*, **135 (6)**, 2365–2378.

Messner, J. W., A. Zeileis, J. Broecker, and G. J. Mayr, 2014: Probabilistic wind power forecasts with an inverse power curve transformation and censored regression. *Wind Energy*, **17 (11)**, 1753–1766.

R Core Team, 2015: *R: A Language and Environment for Statistical Computing*. Vienna, Austria, R Foundation for Statistical Computing.

Umlauf, N., N. Klein, A. Zeileis, and S. Lang, 2015: *bamlss: Bayesian Additive Models for Location Scale and Shape (and Beyond)*. URL `https://r-forge.r-project.org/scm/?group_id=865`, r package version 0.1-1/r1358.

Model and Results

Model on monthly sums

- **R** package **mgcv**
- $y = (\text{monthly sums})^{\frac{1}{2}}$
- $y \sim \mathcal{N}(\mu, \sigma^2)$
- $\mu = \eta_\mu$, $\log(\sigma) = \eta_\sigma$

Model and Results

Model on monthly sums

- **R** package **mgcv**
- $y = (\text{monthly sums})^{\frac{1}{2}}$
- $y \sim \mathcal{N}(\mu, \sigma^2)$
- $\mu = \eta_\mu$, $\log(\sigma) = \eta_\sigma$

Model on daily amounts

- **R** package **bamlss**
- $y = (\text{daily amounts})^{\frac{1}{1.5}}$
- $y^* \sim \mathcal{N}(\mu, \sigma^2)$
- $\mu = \eta_\mu$, $\log(\sigma) = \eta_\sigma$

Error statistics ([*mm*/*month*])

	in-sample		out-of-sample	
	mgcv	bamlss	mgcv	bamlss
BIAS	**-0.13**	0.58	**-0.21**	0.77
MAE	37.62	**37.37**	38.41	**37.90**
RMSE	50.44	**50.17**	51.42	**50.84**