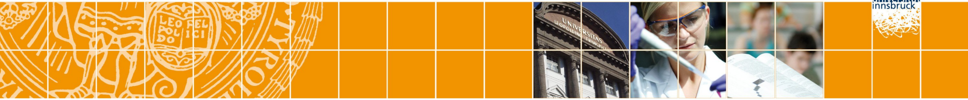
Universität Innsbruck



Heteroscedastic Censored Regression for Weather Forecasts

Jakob W. Messner, Georg J. Mayr, Achim Zeileis

Weather forecasts

Numerical Weather Prediction (NWP)

- Observations → estimate current atmospheric state.
- Simulate atmospheric processes with numerical models.

⇒ Compute future weather

Weather forecasts

Numerical Weather Prediction (NWP)

- Observations → estimate current atmospheric state.
- Simulate atmospheric processes with numerical models.

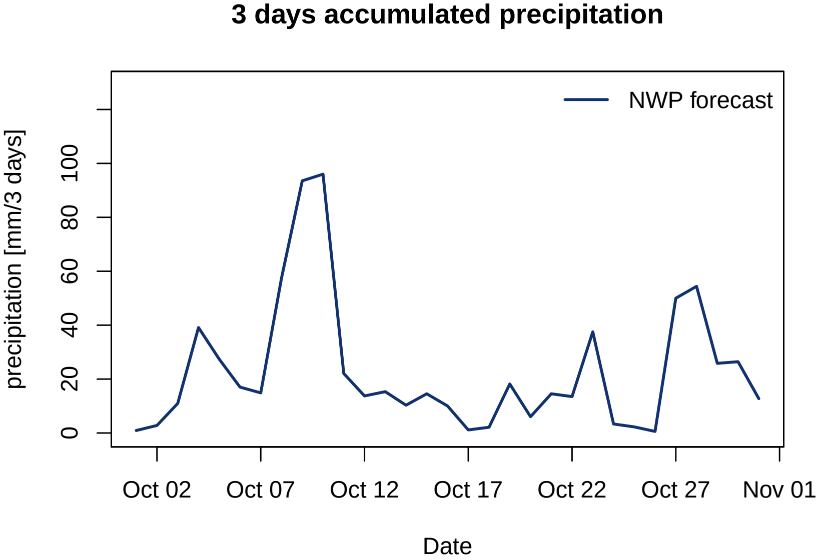
⇒ Compute future weather

Problems:

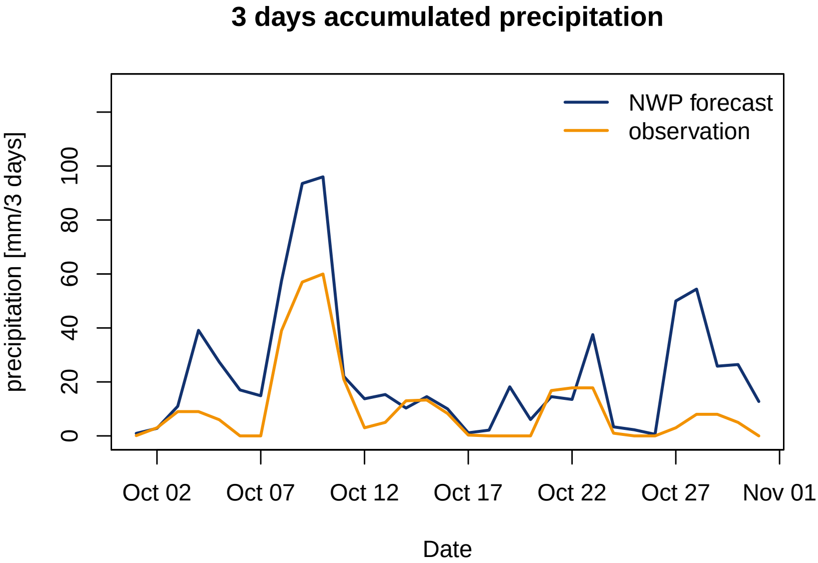
- Few observations
- Observation errors
- Not perfectly known atmospheric processes
- Unresolved processes

⇒ NWP errors

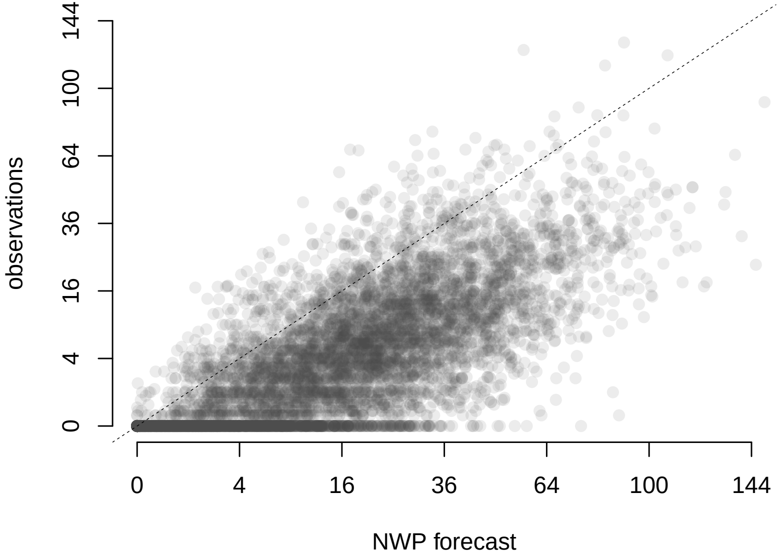
NWP errors


3 days accumulated precipitation
NWP forecast
precipitation [mm/3 days]
0
20
40
60
80
100
Oct 02
Oct 07
Oct 12
Oct 17
Oct 22
Oct 27
Nov 01
Date

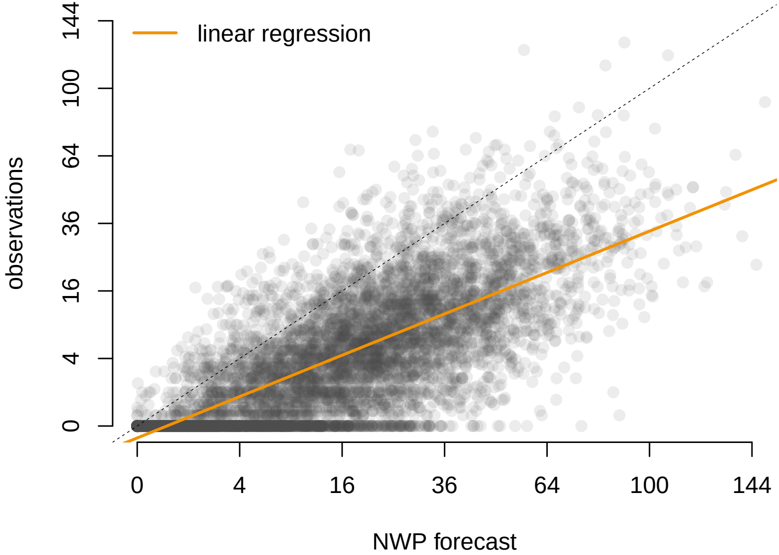
NWP errors


3 days accumulated precipitation
NWP forecast
observation
precipitation [mm/3 days]
0
20
40
60
80
100
Oct 02
Oct 07
Oct 12
Oct 17
Oct 22
Oct 27
Nov 01
Date

NWP errors


observations
0
4
16
36
64
100
144
NWP forecast
0
4
16
36
64
100
144

NWP errors


linear regression
observations
144
100
64
36
16
4
0
NWP forecast
0
4
16
36
64
100
144

Ensemble prediction

NWP error sources:

- Initial conditions
- Model formulations

Ensemble prediction

NWP error sources:

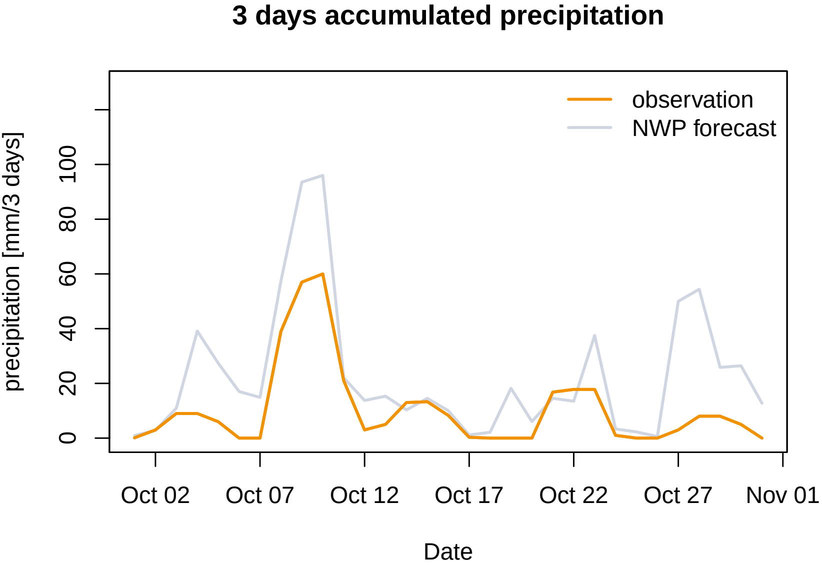
- Initial conditions
- Model formulations

Idea:

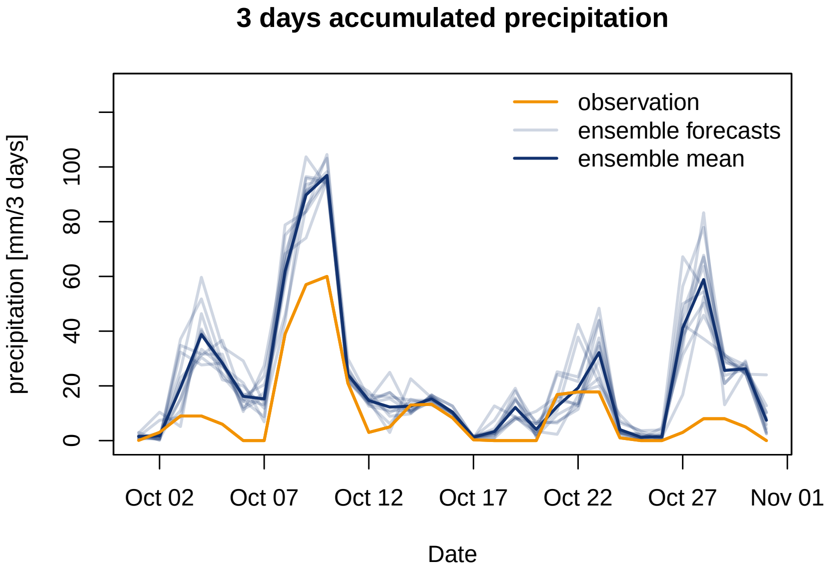
- Perturbed initial conditions
- Different model formulations

⇒ Compute different weather scenarios

NWP errors


3 days accumulated precipitation
observation
NWP forecast
precipitation [mm/3 days]
0
20
40
60
80
100
Oct 02
Oct 07
Oct 12
Oct 17
Oct 22
Oct 27
Nov 01
Date

NWP errors


3 days accumulated precipitation
observation
ensemble forecasts
ensemble mean
precipitation [mm/3 days]
0
20
40
60
80
100
Oct 02
Oct 07
Oct 12
Oct 17
Oct 22
Oct 27
Nov 01
Date

Statistical models

Challenges:

- utilize uncertainty information from ensemble forecasts
- limited (non-negative) response

Heteroscedastic censored regression

$$\begin{aligned} \textit{rain}^* &\sim \mathcal{N}(\mu, \sigma^2) \\ \mu &= \beta_0 + \beta_1 * \textit{ensmean} \\ \log(\sigma) &= \gamma_0 + \gamma_1 * \textit{enssd} \end{aligned}$$

- $\textit{rain}^*$: (latent) precipitation
- *ensmean*: ensemble mean forecast
- *enssd*: ensemble standard deviation
- $\beta_0, \beta_1, \gamma_0, \gamma_1$: regression coefficients

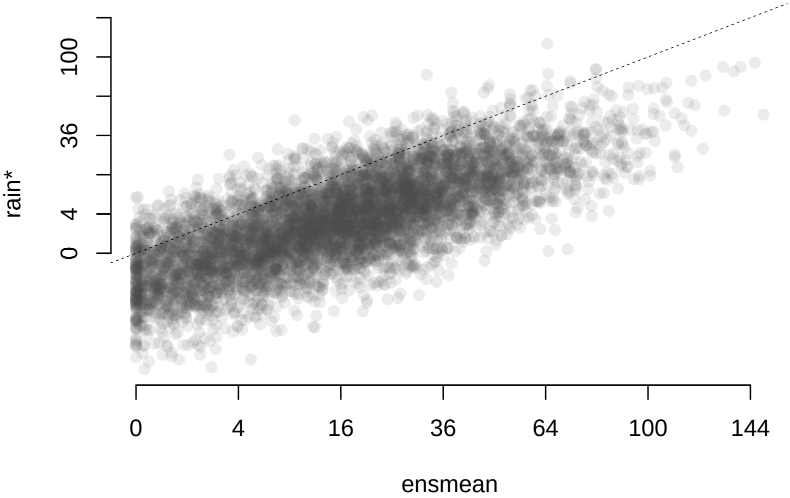
Heteroscedastic censored regression

$$\begin{aligned} \textit{rain}^* &\sim \mathcal{N}(\mu, \sigma^2) \\ \mu &= \beta_0 + \beta_1 * \textit{ensmean} \\ \log(\sigma) &= \gamma_0 + \gamma_1 * \textit{enssd} \end{aligned}$$

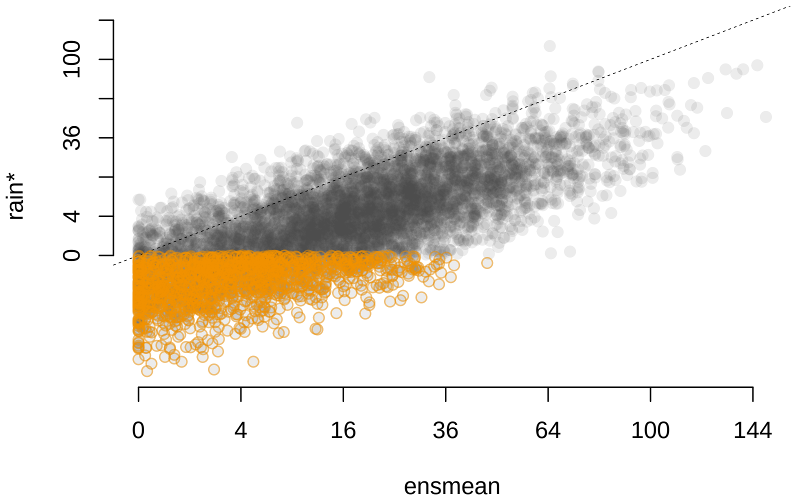
Consider non-negativity:

$$\textit{rain} = \begin{cases} 0 & \textit{rain}^* \leq 0 \\ \textit{rain}^* & \textit{rain}^* > 0 \end{cases}$$

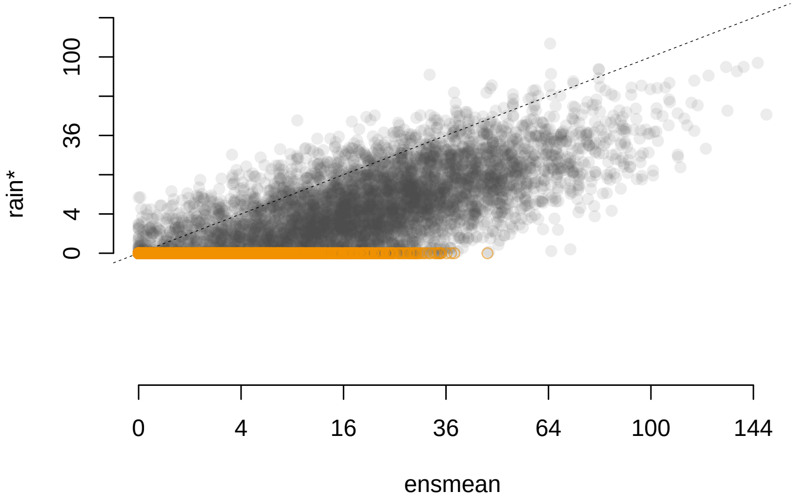
Heteroscedastic censored regression



Heteroscedastic censored regression



Heteroscedastic censored regression



Heteroscedastic censored regression

$$\begin{aligned} \mathit{rain}^* &\sim \mathcal{N}(\mu, \sigma^2) \\ \mu &= \beta_0 + \beta_1 * \mathit{ensmean} \\ \log(\sigma) &= \gamma_0 + \gamma_1 * \mathit{enssd} \end{aligned}$$

Consider non-negativity:

$$\mathit{rain} = \begin{cases} 0 & \mathit{rain}^* \leq 0 \\ \mathit{rain}^* & \mathit{rain}^* > 0 \end{cases}$$

Implementation in R

Model fitting

- `crch()` from package **crch**
- `glm()`-like interface

Implementation in R

Model fitting

- `crch()` from package **crch**
- `glm()`-like interface
- model specification via `formula` plus `data`
- two part formula, e.g., `y ~ x1 + x2 + x3 | z1 + z2`

Implementation in R

Model fitting

- `crch()` from package **crch**
- `glm()`-like interface
- model specification via `formula` plus `data`
- two part formula, e.g., `y ~ x1 + x2 + x3 | z1 + z2`
- log-likelihood maximized numerically via `optim()`
- distributions: Gaussian, logistic, student-t
- censored and truncated

Implementation in R

Model fitting

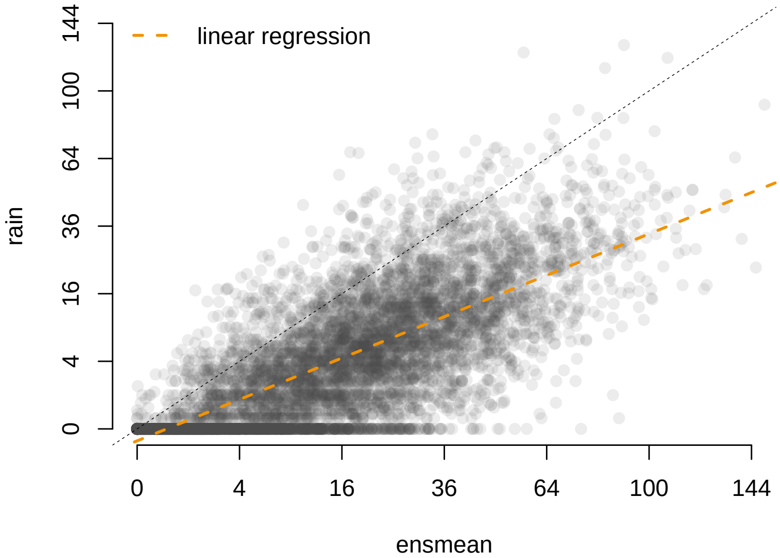
- `crch()` from package **crch**
- `glm()`-like interface
- model specification via `formula` plus `data`
- two part formula, e.g., `y ~ x1 + x2 + x3 | z1 + z2`
- log-likelihood maximized numerically via `optim()`
- distributions: Gaussian, logistic, student-t
- censored and truncated
- methods: `summary()`, `coef()`, `residuals()`, `logLik()`, `predict()`, `...`

```
R> CRCH <- crch(rain ~ ensmean | enssd, data = Rain, left = 0)
R>
```

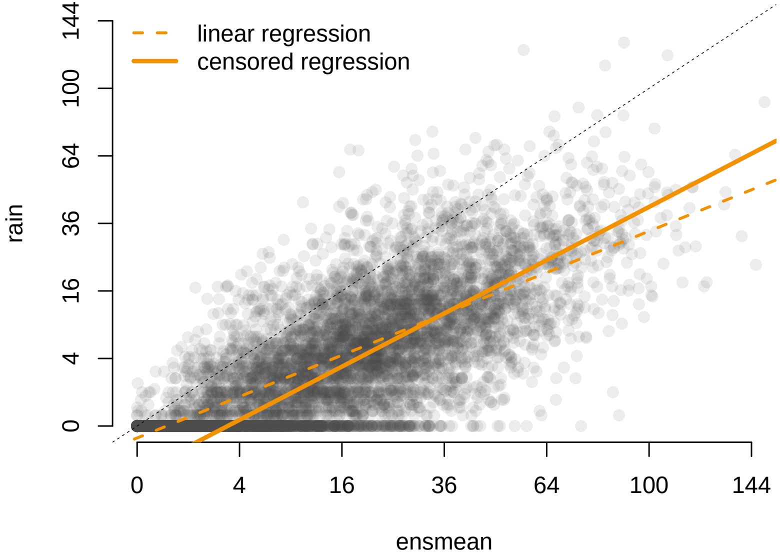
```
R> CRCH <- crch(rain ~ ensmean | enssd, data = Rain, left = 0)
R> summary(CRCH)
```

```
R> CRCH <- crch(rain ~ ensmean | enssd, data = Rain, left = 0)
R> summary(CRCH)

Call:
crch(formula = rain ~ ensmean | enssd, data = Rain, left = 0)


Standardized residuals:
    Min      1Q  Median      3Q     Max
-3.7622 -0.3298  0.2448  0.7536  3.8235

Coefficients (location model):
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.36061    0.04609  -29.52   <2e-16 ***
ensmean      0.78533    0.00962   81.63   <2e-16 ***

Coefficients (scale model with log link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.33189    0.02078  15.975  < 2e-16 ***
enssd        0.25445    0.03827   6.649 2.96e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Distribution: gaussian
Log-likelihood: -1.028e+04 on 4 Df
Number of iterations in BFGS optimization: 12
```

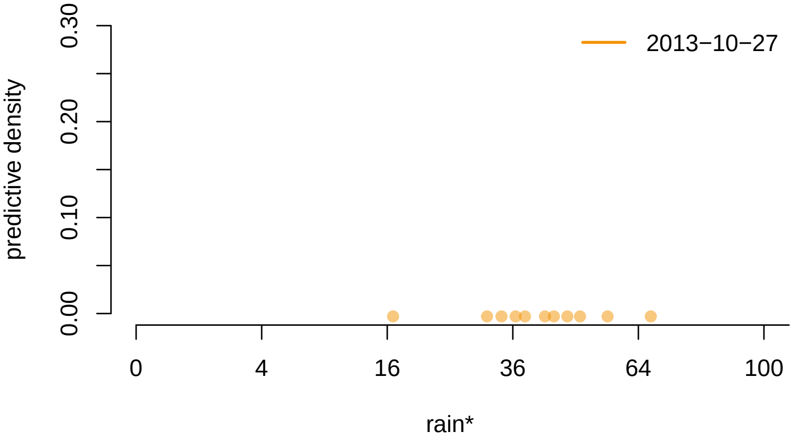
Censored regression



Censored regression

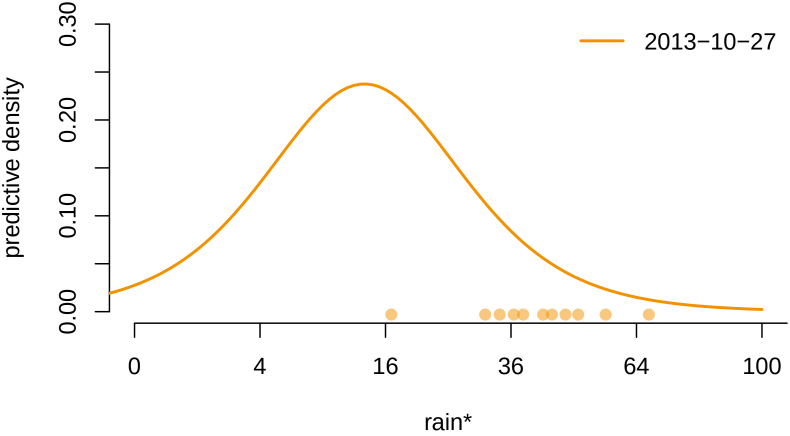


Predictions



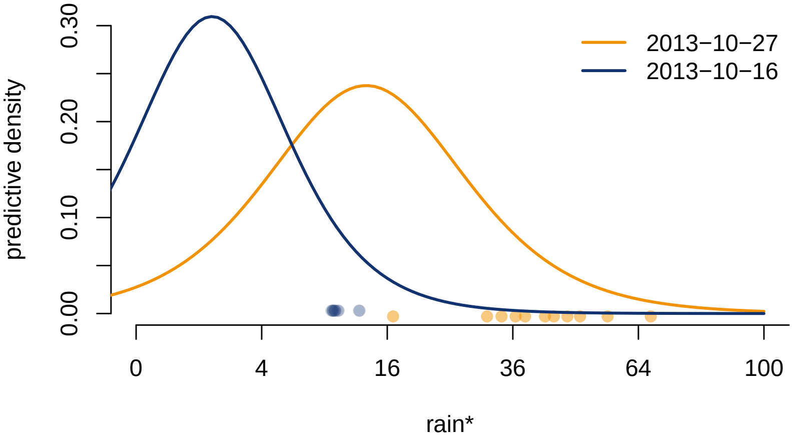
Predictions

```
R> location <- predict(CRCH, newdata = Rain[j,], type = "location")
R> scale    <- predict(CRCH, newdata = Rain[j,], type = "scale")
```



Predictions

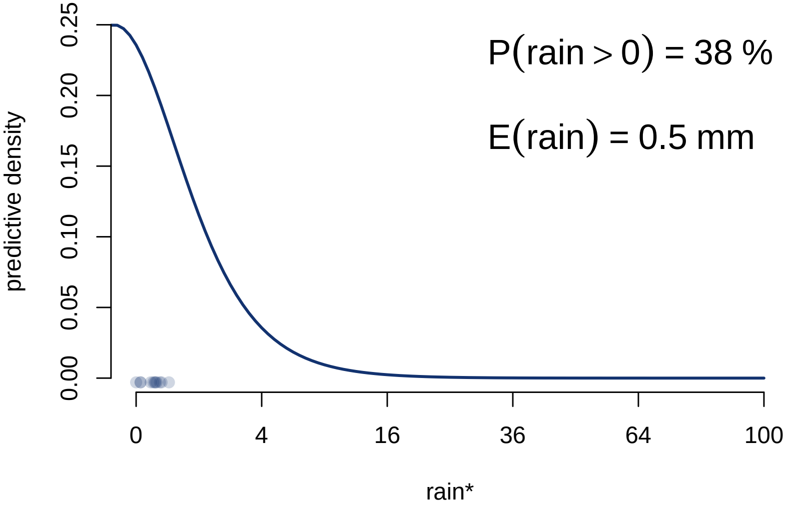
```
R> location <- predict(CRCH, newdata = Rain[j,], type = "location")
R> scale    <- predict(CRCH, newdata = Rain[j,], type = "scale")
```



UseR forecast

Will it rain in Aalborg the next 3 days?

UseR forecast



Summary

Censored regression with conditional heteroscedasticity:

- effective usage of ensemble information
- non-negativity of precipitation considered
- **crch** package:
  - model fitting with `crch()`
  - various methods

Summary

Censored regression with conditional heteroscedasticity:

- effective usage of ensemble information
- non-negativity of precipitation considered
- **crch** package:
  - model fitting with `crch()`
  - various methods

sunny weather for UseR!

Thank you!

Messner, J. W. and A. Zeileis, 2015: *crch: Censored Regression with Conditional Heteroscedasticity*. URL `http://CRAN.R-project.org/package=crch`, R package version 0.9-0.

jakob.messner@uibk.ac.at