



Probabilistic Spatial Forecasting of Daily Precipitation Sums Over Complex Terrain



Supervised by: Georg J. Mayr & Achim Zeileis
Thesis Review: Daniel Wilks & Thomas Hamill

Ph.D. Defense Presentation 2016-11-14

The Project: SnowSafeFX

Project goals

- improve skill of **new snow amount** forecasts
- **spatial probabilistic** prediction, Tyrol

The Project: SnowSafeFX

Project goals

- improve skill of **new snow amount** forecasts
- **spatial probabilistic** prediction, Tyrol

Early remark: Our “snow” is still a bit liquid.

The Project: SnowSafeFX

Importance

- outdoor sportsmen:
210 **avalanche events**,
31 deaths (winter 14/15, A)¹



¹Winterbericht 2014/2015, Lawinenwarndienst Tirol.

²Statistisches Jahrbuch Bundesland Tirol, Amt der Tiroler Landesregierung.

The Project: SnowSafeFX

Importance

- outdoor sportsmen:
210 **avalanche events**,
31 deaths (winter 14/15, A)¹
- tourism:
5.3 million **visitors** (winter 13/14)²

¹Winterbericht 2014/2015, Lawinenwarndienst Tirol.

²Statistisches Jahrbuch Bundesland Tirol, Amt der Tiroler Landesregierung.



The Project: SnowSafeFX

Importance

- outdoor sportsmen:
210 **avalanche events**,
31 deaths (winter 14/15, A)¹
- tourism:
5.3 million **visitors** (winter 13/14)²
- public:
safety of infrastructure and people,
transport, ...

¹Winterbericht 2014/2015, Lawinenwarndienst Tirol.

²Statistisches Jahrbuch Bundesland Tirol, Amt der Tiroler Landesregierung.



The Project: SnowSafeFX

Difficulties

- ensemble prediction system (EPS) **too coarse**
- **small-scale** features **not** well represented
- EPS often underdispersive¹

¹Mullen and Buizza (2001), Hagedorn et al. (2012).

The Project: SnowSafeFX

Difficulties

- ensemble prediction system (EPS) **too coarse**
- **small-scale** features **not** well represented
- EPS often underdispersive¹

Approach to solution

- **down-scaling** (in general)
- statistical **post-processing** (MOS)

¹Mullen and Buizza (2001), Hagedorn et al. (2012).

The Project: SnowSafeFX

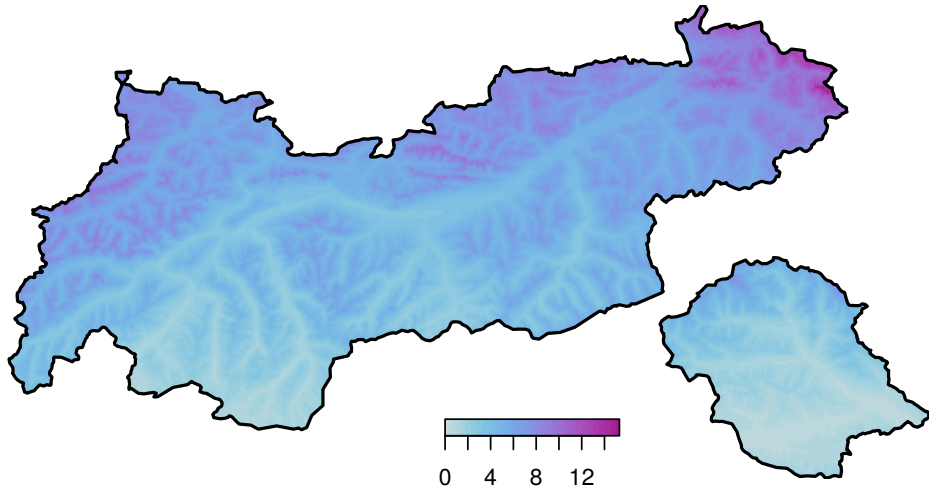


Figure 1: Precipitation forecast for North Tyrol [mm day^{-1}].

[illegible]

Reto Stauffe

**Spatio-temporal precipitation climatology
in terrain using a censored additive regression**

Cheng-I. Mo^a, Jacob W. Messner^{a,b}, Nikolai Unluar^a and
^aDepartment of Statistics, Faculty of Environment and Society, University of Waterloo, Waterloo, Ontario, Canada
^bDepartment of Atmospheric and Civil Engineering Sciences, University of Waterloo, Waterloo, Ontario, Canada

[illegible]

Received 4 April 2016; Revised 6 July 2016; Accepted 12 September 2016

1. Introduction

[illegible]

Publication I

Stauffer, R., G. J. Mayr, J. W. Messner, N. Umlauf, and A. Zeileis (2016): Spatio-Temporal Precipitation Climatology over Complex Terrain Using a Censored Additive Regression Model. *International Journal of Climatology*, doi: 10.1002/joc.4913, in print.

Introduction

Research goals

- develop a **flexible** statistical model
- **spatio-temporal** climatology
- **daily** resolution
- **full** climatological **distribution**

Introduction

Research goals

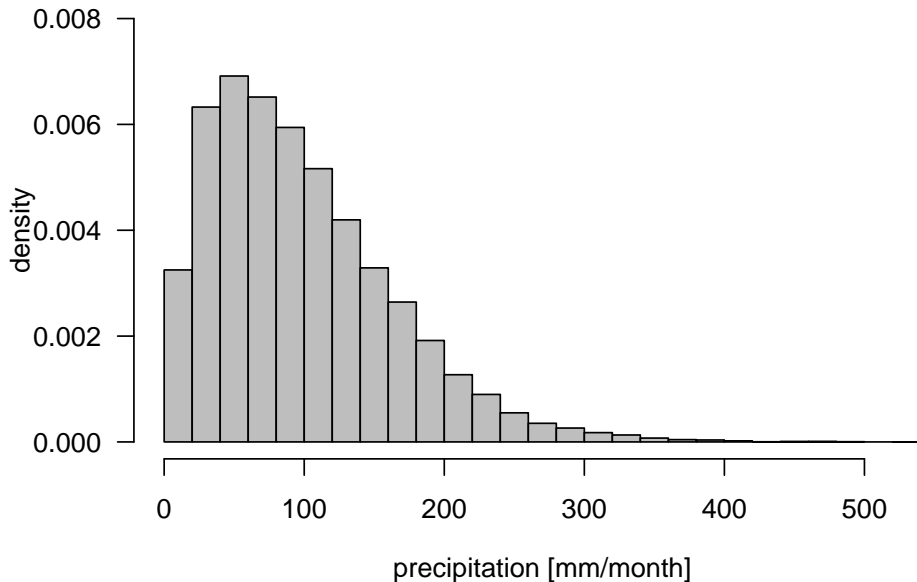
- develop a **flexible** statistical model
- **spatio-temporal** climatology
- **daily** resolution
- **full** climatological **distribution**

Challenges

- suitable **response distribution**
- **effects** to be considered

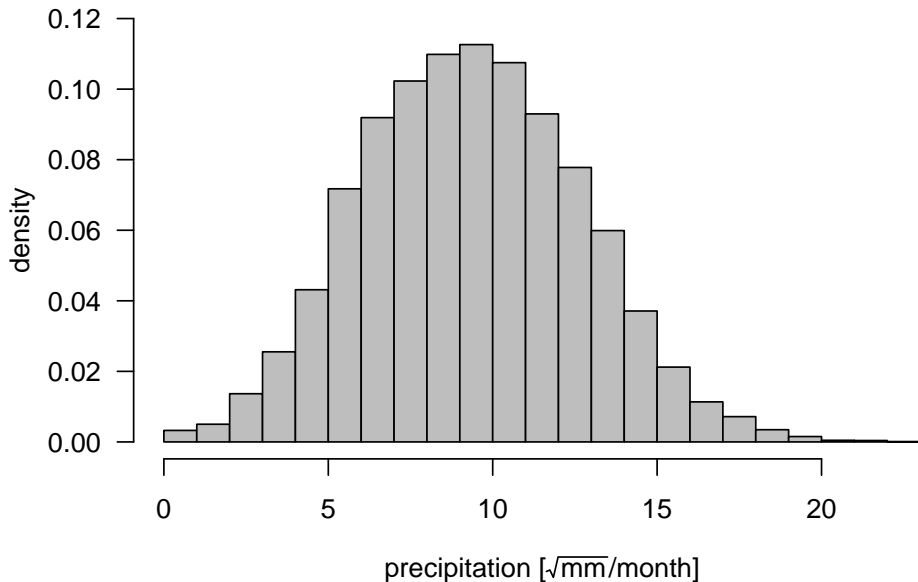
Response Distribution

Monthly precipitation sums



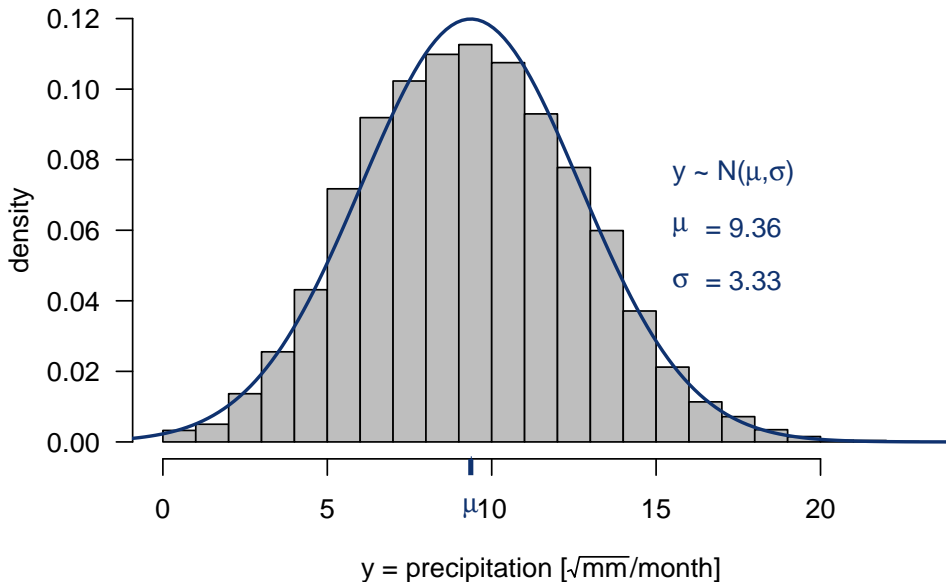
Response Distribution

Monthly precipitation sums



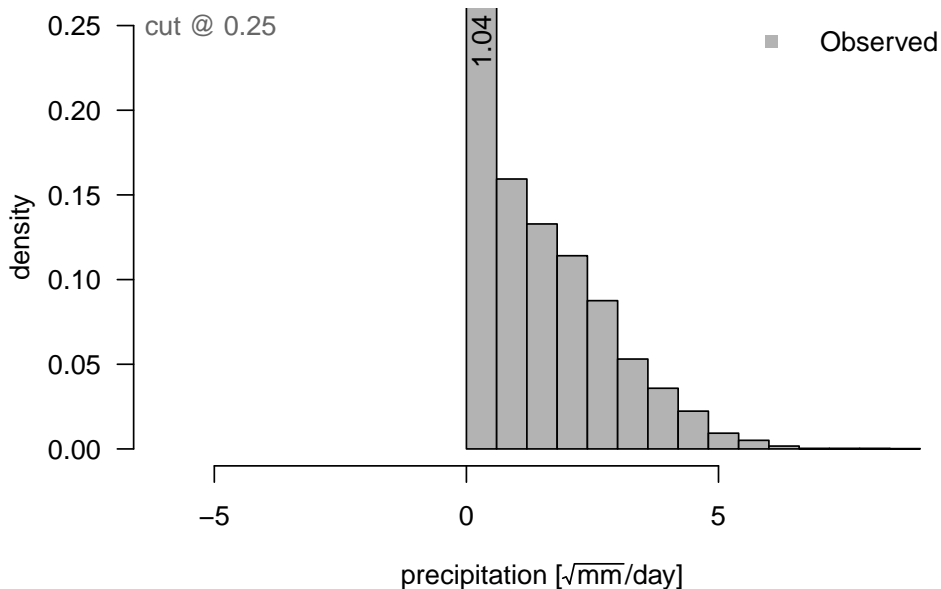
Response Distribution

Monthly precipitation sums

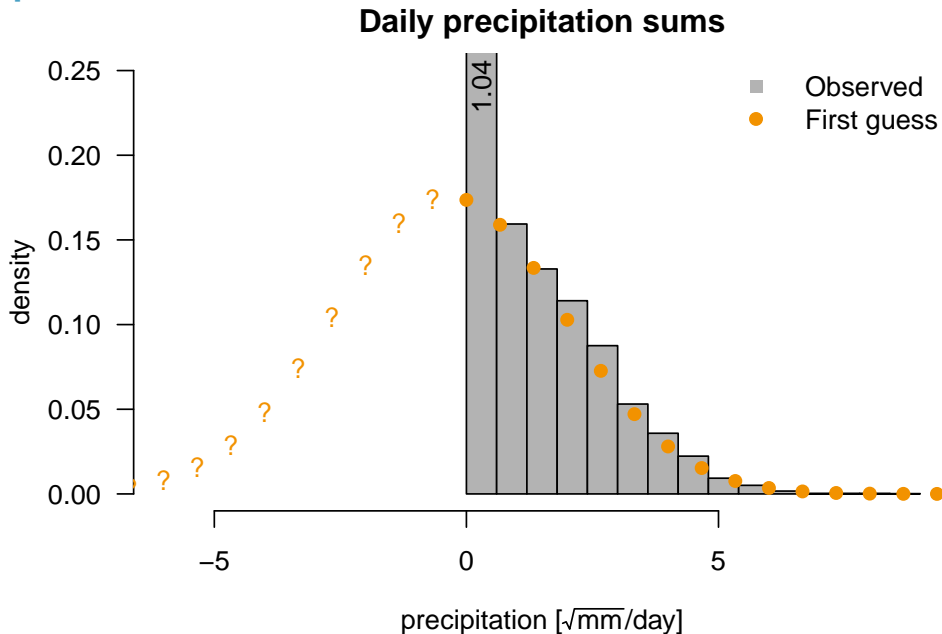


Response Distribution

Daily precipitation sums



Response Distribution



Censoring

Can be seen as censored if:

- **limited** to a threshold and . . .
- values exceeding threshold **cannot occur**.

Censoring

Can be seen as censored if:

- **limited** to a threshold and ...
- values exceeding threshold **cannot occur**.

Examples

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

Censoring

Can be seen as censored if:

- **limited** to a threshold and ...
- values exceeding threshold **cannot occur**.

Examples

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

Censoring

Can be seen as censored if:

- **limited** to a threshold and ...
- values exceeding threshold **cannot occur**.

Examples

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

Model assumption

$$\underbrace{y \sim \mathcal{N}(\mu, \sigma)}_{\text{latent}}, \quad \underbrace{\text{precipitation} = \max(0, y)^p}_{\text{}} \quad (1)$$

Censoring

Can be seen as censored if:

- **limited** to a threshold and ...
- values exceeding threshold **cannot occur**.

Examples

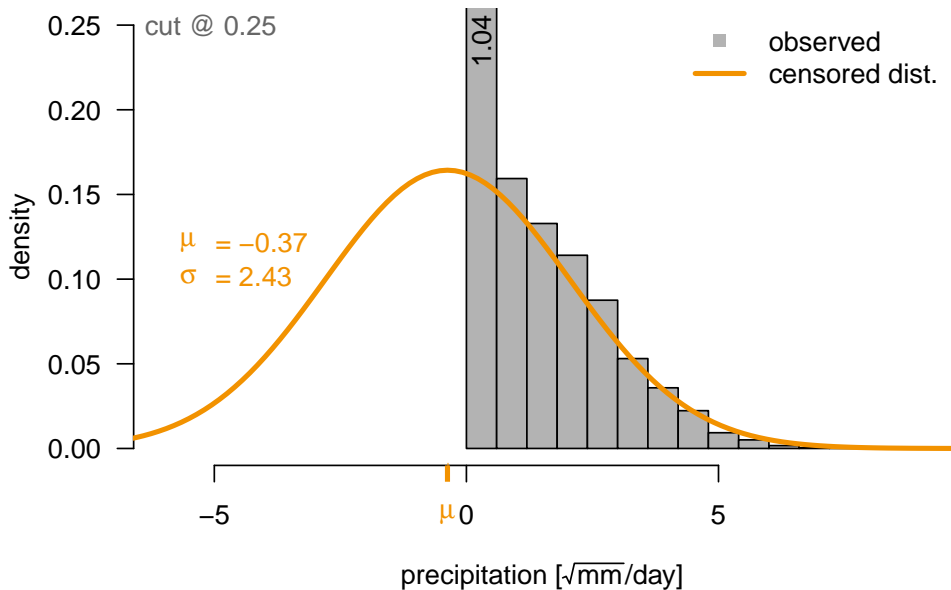
- survival in 5 yr clinical study: **right** $y_i \leq 5$
- hours worked this week: **two sided** $0 \leq y_i \leq 168$
- precipitation: **left** $0 \leq y_i$

Model assumption

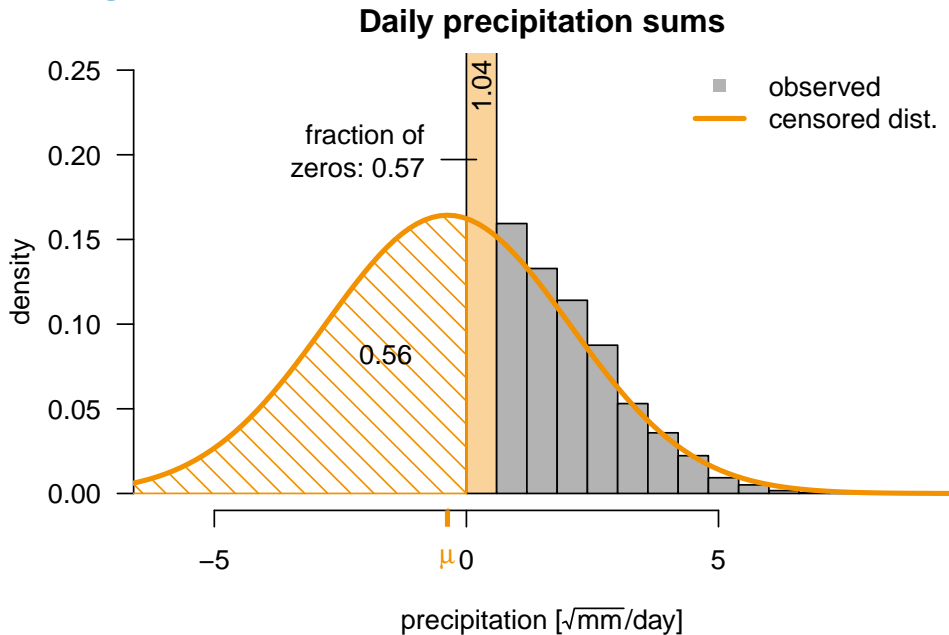
$$\underbrace{y \sim \mathcal{N}(\mu, \sigma)}_{\text{latent}}, \quad \underbrace{\text{precipitation} = \max(0, y)^p}_{\text{observable}} \quad (1)$$

Censoring

Daily precipitation sums



Censoring



Required Effects

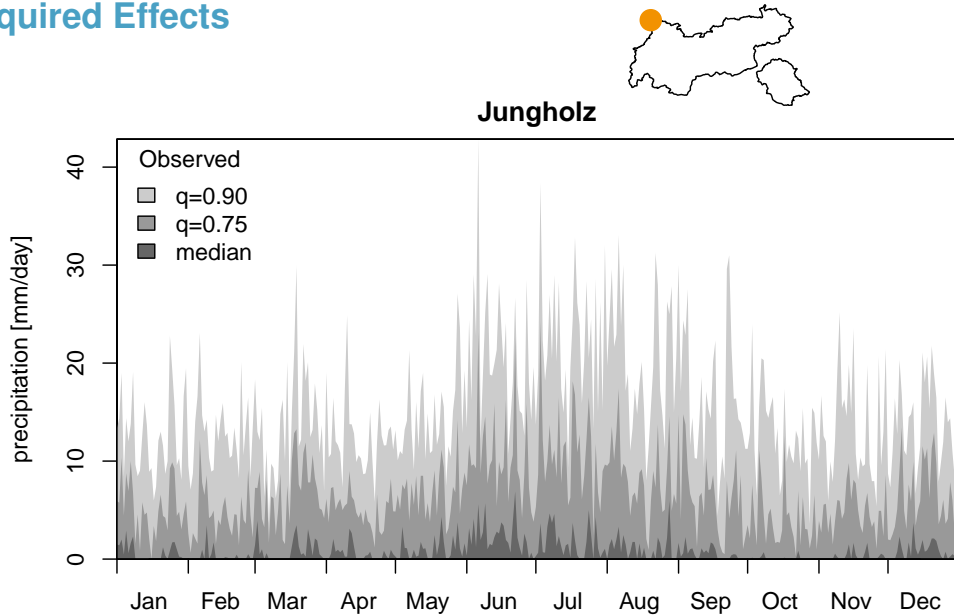


Figure 2: Empirical quantiles on a daily basis (1980-2012).

Required Effects

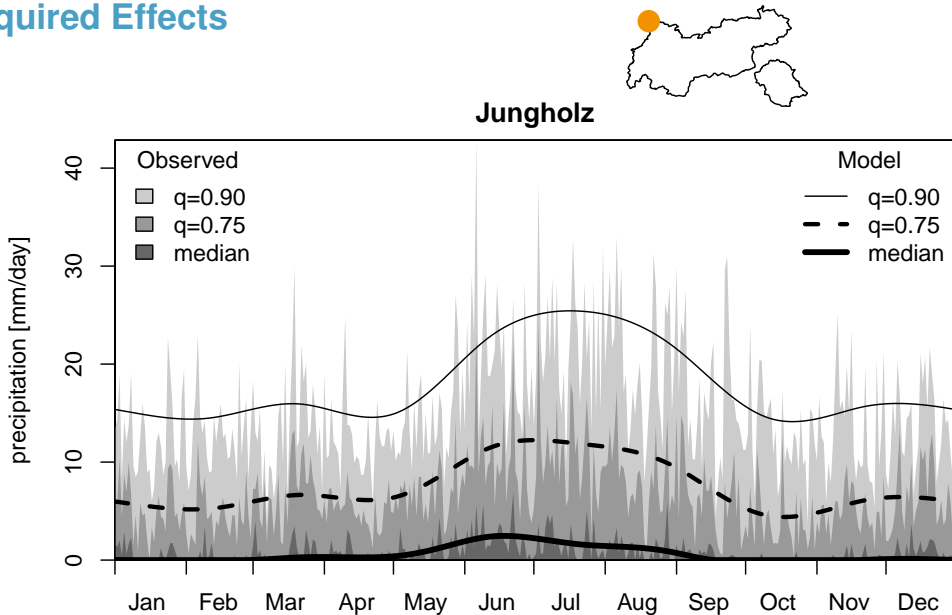


Figure 2: Empirical quantiles and climatological estimate.

Required Effects

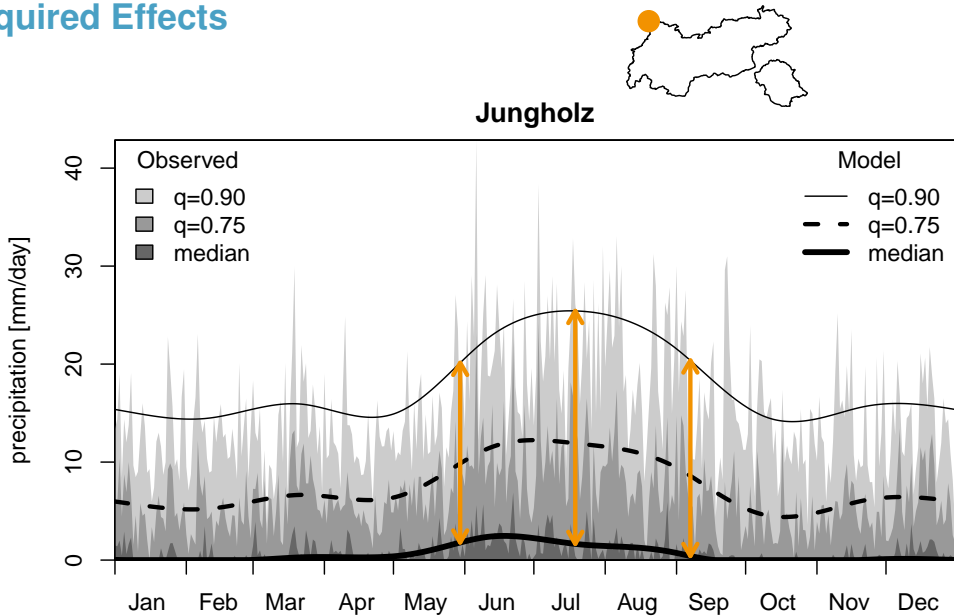


Figure 2: Empirical quantiles and climatological estimate.

Required Effects

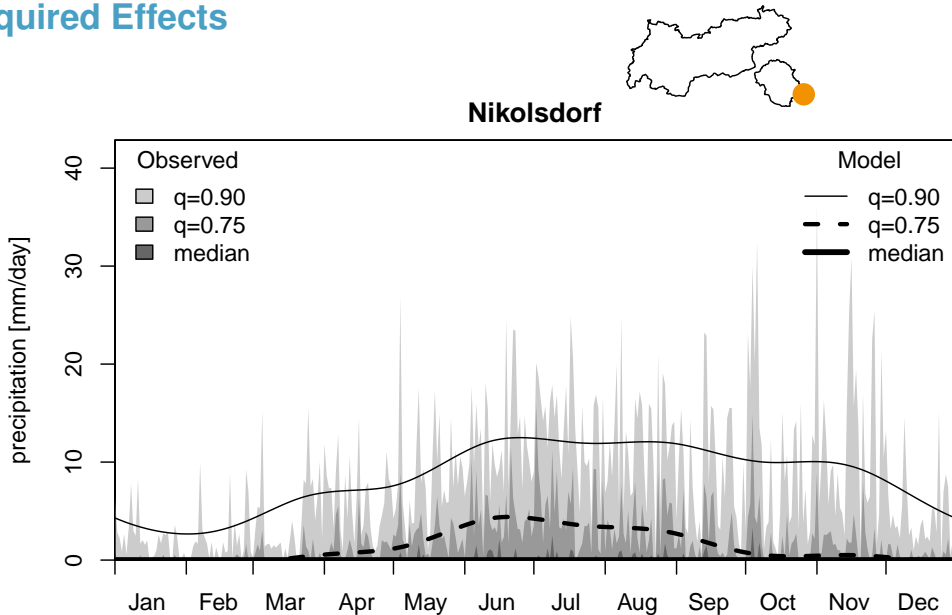


Figure 2: Empirical quantiles and climatological estimate.

Spatio-Temporal Model

Reminder

$$y \sim \mathcal{N}(\mu, \sigma), \quad \text{precipitation} = \max(0, y)^p \quad (2)$$

- y : latent response
- $\mathcal{N}(\dots)$: Gaussian distribution
- μ, σ : latent location and scale
- p : power parameter

Spatio-Temporal Model

Reminder

$$y \sim \mathcal{N}(\mu, \sigma), \quad \text{precipitation} = \max(0, y)^p \quad (2)$$

Linear predictors

$$\mu = \beta_0 + s_1(\text{altitude}) + s_2(\text{season}) + s_3(\text{long, lat}) + s_4(\text{season, long, lat}) \quad (3)$$

- y : latent response
- $\mathcal{N}(\dots)$: Gaussian distribution
- μ, σ : latent location and scale
- p : power parameter
- β_0 : global intercepts
- s_1 : altitudinal effect
- s_2 : cyclic seasonal effect
- s_3 : spatial effect
- s_4 : mixed effect

Spatio-Temporal Model

Reminder

$$y \sim \mathcal{N}(\mu, \sigma), \quad \text{precipitation} = \max(0, y)^p \quad (2)$$

Linear predictors

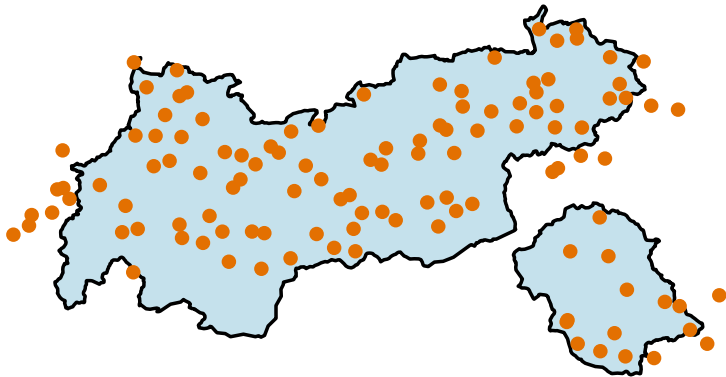
$$\mu = \beta_0 + s_1(\text{altitude}) + s_2(\text{season}) + s_3(\text{long, lat}) + s_4(\text{season, long, lat}) \quad (3)$$

$$\sigma = \gamma_0 + t_1(\text{altitude}) + t_2(\text{season}) + t_3(\text{long, lat}) + t_4(\text{season, long, lat})$$

- y : latent response
- $\mathcal{N}(\dots)$: Gaussian distribution
- μ, σ : latent location and scale
- p : power parameter
- β_0, γ_0 : global intercepts
- s_1, t_1 : altitudinal effect
- s_2, t_2 : cyclic seasonal effect
- s_3, t_3 : spatial effect
- s_4, t_4 : mixed effect

Data Set

- **118 stations**³
- Tyrol and surrounding
- **daily** observations
- 1971 – 2012
- more than **1.5 million** unique observations

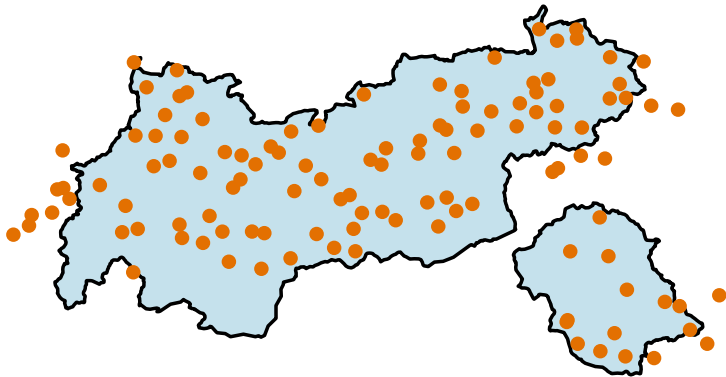


³Bundesministerium für Forstwirtschaft, Umwelt und Wasserwirtschaft, Abteilung IV/4 – Wasserhaushalt.

⁴Umlauf et al. 2016: bamls: Bayesian Additive Models for Location Scale and Shape (and Beyond).

Data Set

- **118 stations**³
- Tyrol and surrounding
- **daily** observations
- 1971 – 2012
- more than **1.5 million** unique observations
- R package **bamlss**⁴



³Bundesministerium für Forstwirtschaft, Umwelt und Wasserwirtschaft, Abteilung IV/4 – Wasserhaushalt.

⁴Umlauf et al. 2016: bamlss: Bayesian Additive Models for Location Scale and Shape (and Beyond).

Results

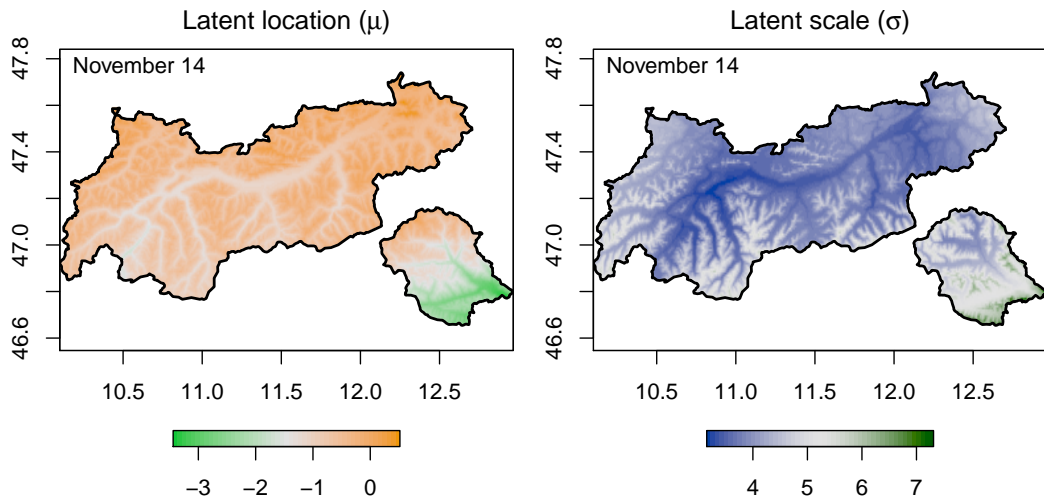


Figure 3: Climatological estimate. Location μ and scale σ on power-transformed latent scale.

Results

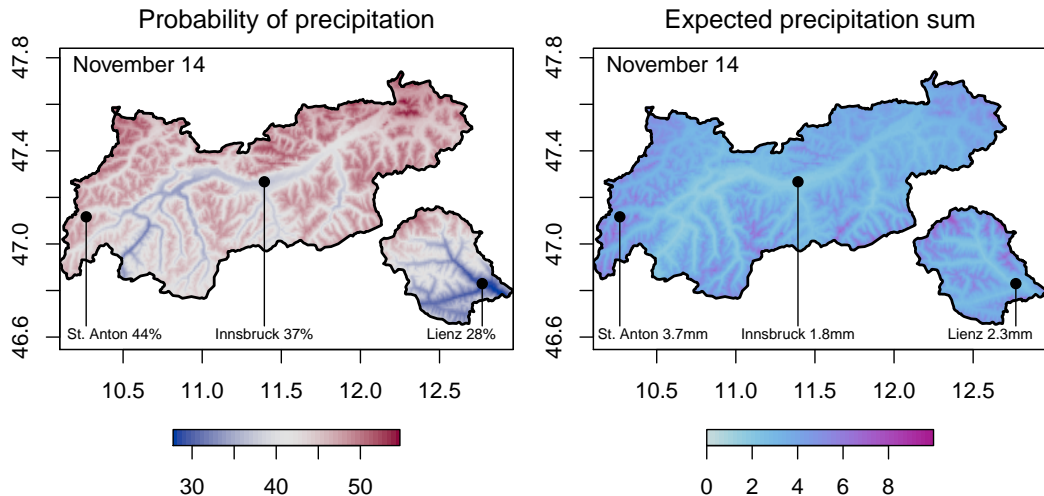


Figure 4: Climatological estimate. Probability of precipitation [%] and expectation [$mm\ day^{-1}$].

Results

Observed & estimated PDF

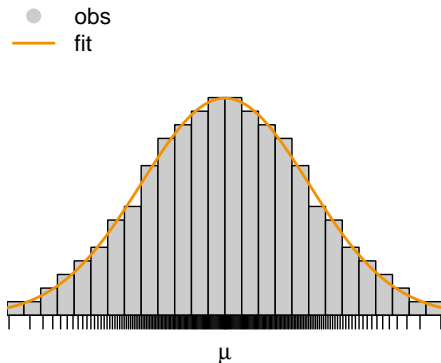


Figure 5: Probability integral transform (PIT; Gneiting et al. 2007) histogram.

Results

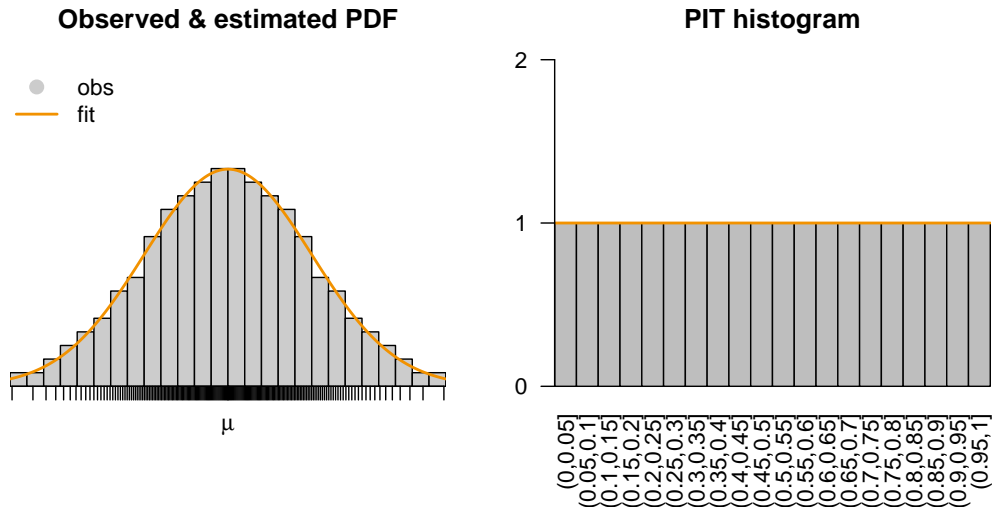


Figure 5: Probability integral transform (PIT; Gneiting et al. 2007) histogram.

Results

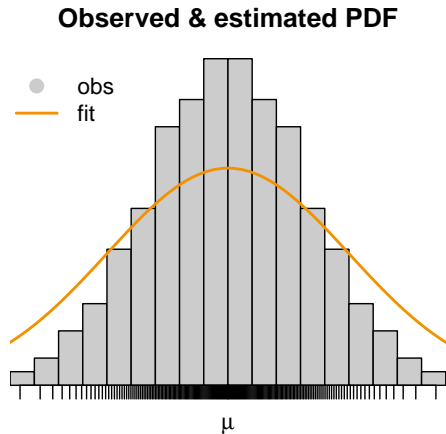


Figure 5: Probability integral transform (PIT; Gneiting et al. 2007) histogram.

Results

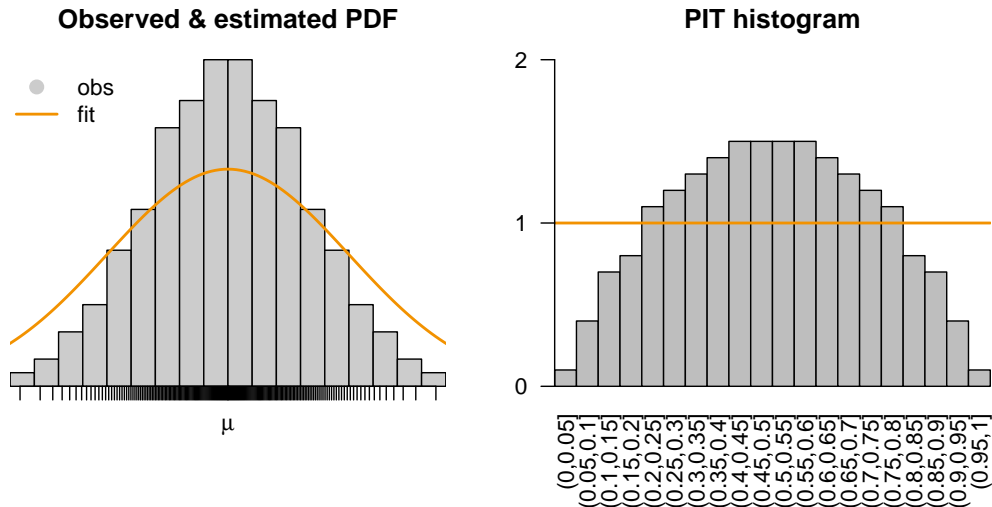


Figure 5: Probability integral transform (PIT; Gneiting et al. 2007) histogram.

Results

Observed & estimated PDF

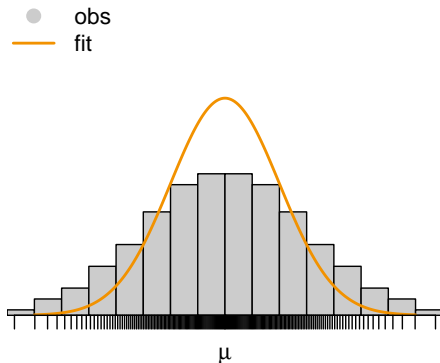


Figure 5: Probability integral transform (PIT; Gneiting et al. 2007) histogram.

Results

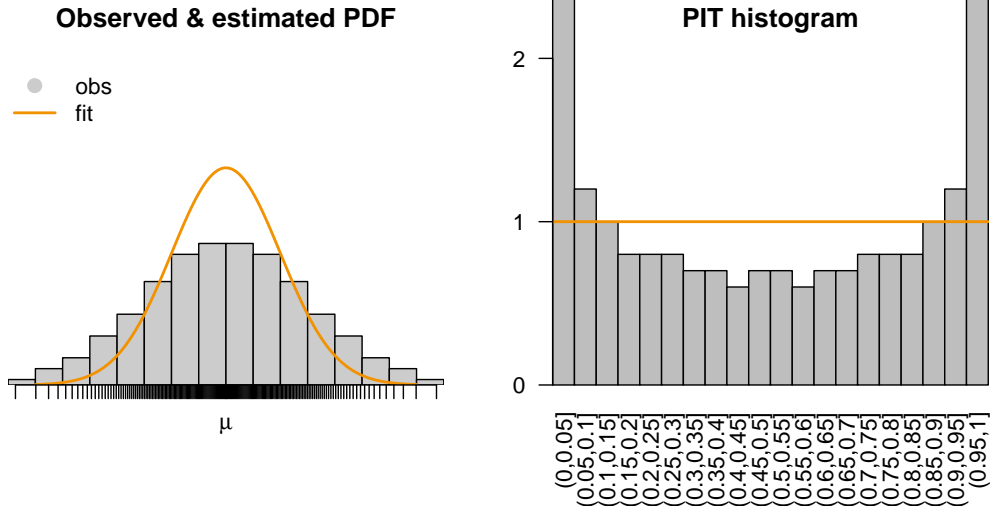
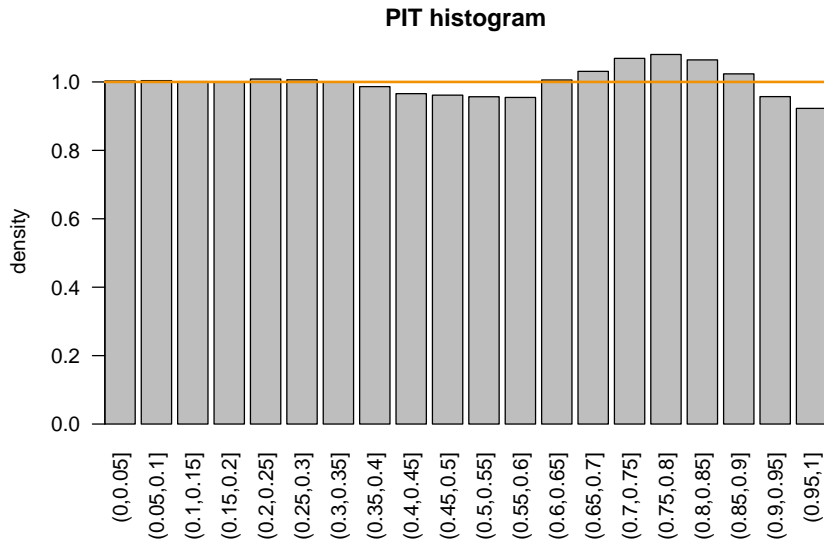
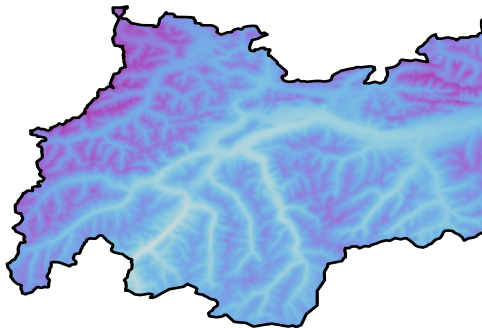


Figure 5: Probability integral transform (PIT; Gneiting et al. 2007) histogram.



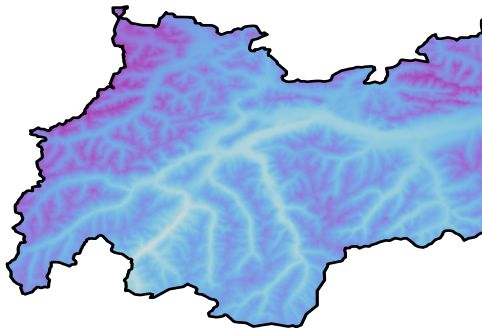
Summary I

- **power transformation** to remove skewness
- censoring to handle **zero-observations**
- **full** climatological **distribution**



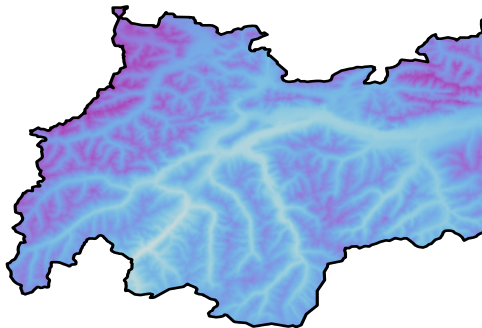
Summary I

- **power transformation** to remove skewness
- censoring to handle **zero-observations**
- **full** climatological **distribution**
- “**simple**” generalized setup
- **spatial/temporal** resolution arbitrary scalable
- accurate estimate at station level



Summary I

- **power transformation** to remove skewness
- censoring to handle **zero-observations**
- **full** climatological **distribution**
- “**simple**” generalized setup
- **spatial/temporal** resolution arbitrary scalable
- accurate estimate at station level
- **probabilistic** reference method
- **background information** for statistical post-processing methods



Best Student Poster Toyota Corporation Award

International Workshop of Statistical Modelling 2016

Ensemble Post-Processing of Daily Precipitation Sums over Complex Terrain Using Censored High-Resolution Standardized Anomalies

RETO STAUFFER* AND NIKOLAUS UMLAUF† AND JAKOB W. MESSNER‡ AND GEORG J. MAYR§ AND ALEXANDER ZEILEIS¶

Department of Statistics & Institute of Atmospheric and Cryospheric Sciences, University of Innsbruck, Austria

ABSTRACT

Predictable features provided by numerical ensemble prediction systems for extreme events and wet weather events are often used to improve the results. The lack of high-resolution standardized anomalies for precipitation sums over complex terrain leads to the use of standardized anomalies. The lack of high-resolution standardized anomalies for precipitation sums over complex terrain leads to the use of standardized anomalies. The lack of high-resolution standardized anomalies for precipitation sums over complex terrain leads to the use of standardized anomalies.

probabilistic information can be used for e.g. statistical planning or decision making. An ensemble consists of several independent forecast runs with highly correlated errors. The goal of an EPS is to not only provide one but many forecasts that provide additional information about the weather situation dependent forecast uncertainty. Although EPS are adequate to provide forecasts and are not able to provide reliable forecasts and are not able to provide reliable forecasts and are not able to provide reliable forecasts.

to correct for systematic errors and to correct the uncertainty provided by the EPS, post-processing methods for post-processing are available. The lack of high-resolution standardized anomalies for precipitation sums over complex terrain leads to the use of standardized anomalies.

Publication II

Stauffer, R., G. J. Mayr, J. W. Messner, N. Umlauf, and A. Zeileis (2016): Ensemble Post-Processing of Daily Precipitation Sums over Complex Terrain Using Censored High-Resolution Standardized Anomalies. *Monthly Weather Review*, in proof.

Introduction

Numerical weather prediction (NWP)

1. analysis: → current state
2. prognosis: → future state

Introduction

Numerical weather prediction (NWP)

1. analysis: → current state
2. prognosis: → future state

Error sources

- observations
- **simplified** model world
- numerical **approximation**
- “**unknown**” atmospheric **processes**

Introduction

Numerical weather prediction (NWP)

1. analysis: → current state
2. prognosis: → future state

Error sources

- observations
- **simplified** model world
- numerical **approximation**
- “**unknown**” atmospheric **processes**

Ensemble prediction systems (EPS)

- to **quantify** the **uncertainty**

Introduction

Numerical weather prediction (NWP)

1. analysis: → current state
2. prognosis: → future state

Error sources

- observations
- **simplified** model world
- numerical **approximation**
- “**unknown**” atmospheric **processes**

Ensemble prediction systems (EPS)

- to **quantify** the **uncertainty**
- **number of members** restricted
- typically **underdispersive**

Introduction

Forecast error

- *total error* = *noise* + *systematic errors*

Introduction

Forecast error

- *total error* = *noise* + *systematic errors*
- *noise*: unexplainable (signal-free)
- *systematic errors*: correction possible

Introduction

Forecast error

- *total error* = *noise* + *systematic errors*
- *noise*: unexplainable (signal-free)
- *systematic errors*: correction possible

Post-processing

- correct **bias**
- correct **uncertainty**

Introduction

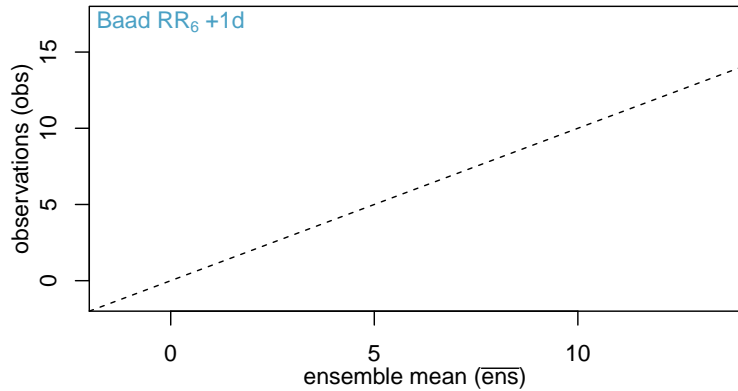
Forecast error

- *total error* = *noise* + *systematic errors*
- *noise*: unexplainable (signal-free)
- *systematic errors*: correction possible

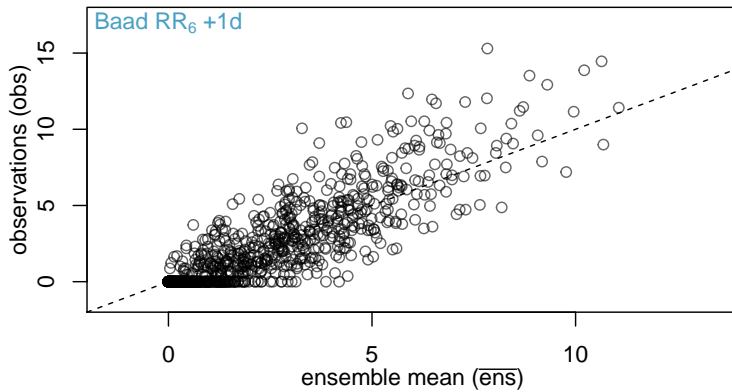
Post-processing

- correct **bias**
- correct **uncertainty**
- discrete → **full distribution**
- probabilities, quantiles, expectation

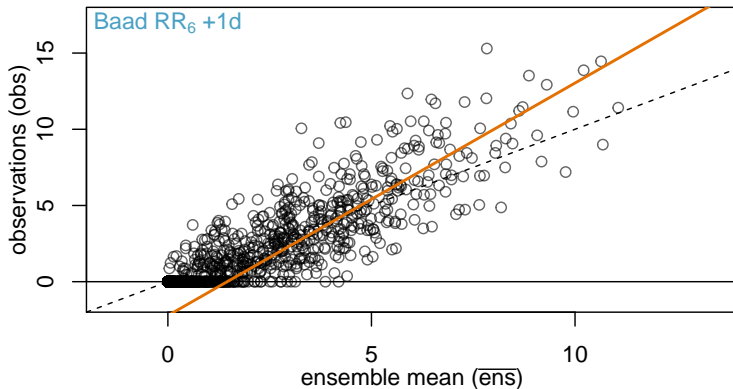
Methodology



Methodology



Methodology



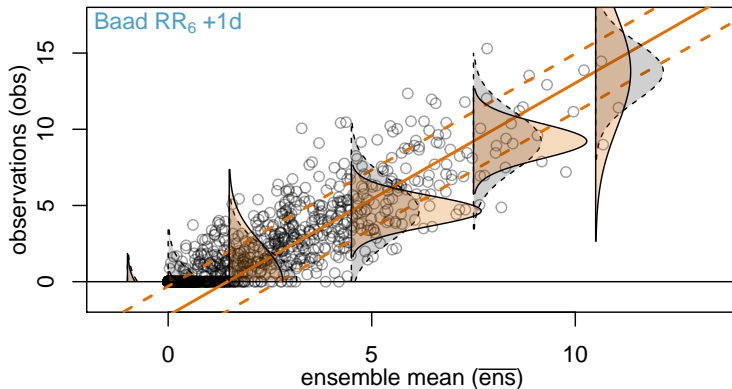
Censored logistic regression

precipitation = $\max(0, y)^p$ with $y \sim \mathcal{L}(\mu, \sigma)$

$$\mu = \beta_0 + \beta_1 \cdot \overline{ens} \quad (4)$$

$$\sigma = \gamma_0$$

Methodology



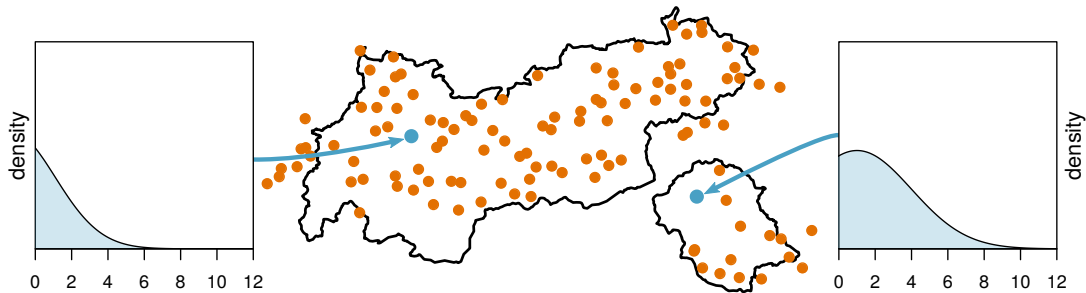
Censored **non-homogeneous** logistic regression

precipitation = $\max(0, y)^p$ with $y \sim \mathcal{L}(\mu, \sigma)$

$$\mu = \beta_0 + \beta_1 \cdot \overline{ens} \quad (4)$$

$$\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv}(\text{ens})$$

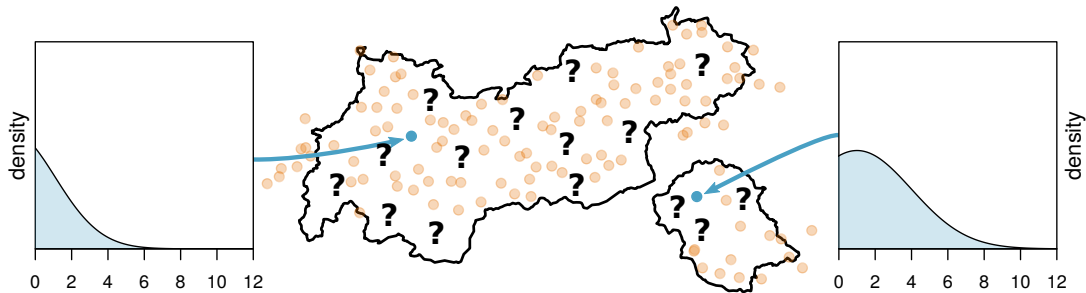
Pointwise Post-Processing



Pointwise models

- “relatively” simple
- interpolate ensemble
- apply censored non-homogeneous model
- **one model** for **each** station

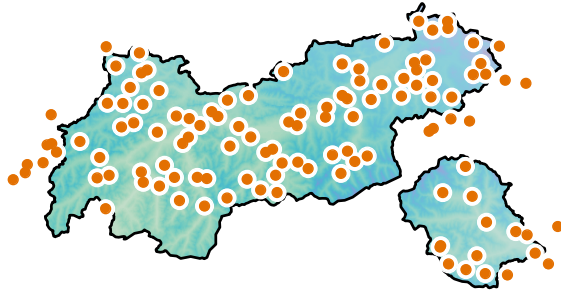
Pointwise Post-Processing



Pointwise models

- “relatively” simple
- interpolate ensemble
- apply censored non-homogeneous model
- **one model** for **each** station

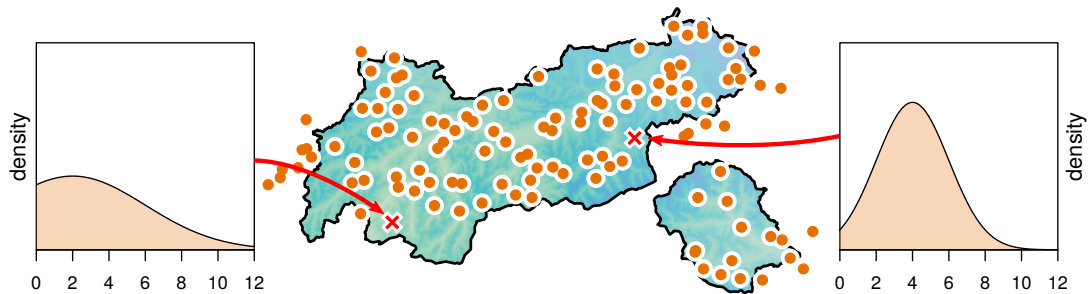
Spatial Post-Processing



Spatial model

- **one model** for **all** stations
- model: **station independent**
- full **spatial** prediction

Spatial Post-Processing



Spatial model

- **one model** for **all** stations
- model: **station independent**
- full **spatial** prediction

Spatial Post-Processing



Spatial Post-Processing



Spatial Post-Processing



Standardized Anomaly Model Output Statistics (SAMOS⁵)

- Use climatology as **background information**
- ... to **remove location-dependent features**
- ... and to bring stations to a **comparable scale**.

⁵Dabernig et al. (2016): Spatial Ensemble Post-Processing with Standardized Anomalies.

Spatial Post-Processing

Standardized Anomaly Model Output Statistics (SAMOS)

$$\frac{y - obs_{\mu}}{obs_{\sigma}} \sim \mathcal{L}(\mu, \sigma)$$

(5)

- y : observations^{1/p}
- obs_{μ}, obs_{σ} : observed climatology
- $\mathcal{L}(\dots)$: logistic distribution
- μ, σ : latent location and scale

Spatial Post-Processing

Standardized Anomaly Model Output Statistics (SAMOS)

$$\frac{y - obs_{\mu}}{obs_{\sigma}} \sim \mathcal{L}(\mu, \sigma)$$
$$\mu = \beta_0 + \beta_1 \cdot \text{mean}\left(\frac{ens - ens_{\mu}}{ens_{\sigma}}\right) \quad (5)$$

- y : observations^{1/p}
- obs_{μ}, obs_{σ} : observed climatology
- $\mathcal{L}(\dots)$: logistic distribution
- μ, σ : latent location and scale
- ens : latest EPS forecast^{1/p}
- ens_{μ}, ens_{σ} : EPS climatology
- β_0 : global intercept
- β_1 : steepness coefficient

Spatial Post-Processing

Standardized Anomaly Model Output Statistics (SAMOS)

$$\begin{aligned}\frac{y - obs_{\mu}}{obs_{\sigma}} &\sim \mathcal{L}(\mu, \sigma) \\ \mu &= \beta_0 + \beta_1 \cdot \text{mean}\left(\frac{ens - ens_{\mu}}{ens_{\sigma}}\right) \\ \sigma &= \gamma_0 + \gamma_1 \cdot \text{stdv}\left(\frac{ens - ens_{\mu}}{ens_{\sigma}}\right)\end{aligned}\tag{5}$$

- y : observations^{1/p}
- obs_{μ}, obs_{σ} : observed climatology
- $\mathcal{L}(\dots)$: logistic distribution
- μ, σ : latent location and scale
- ens : latest EPS forecast^{1/p}
- ens_{μ}, ens_{σ} : EPS climatology
- β_0, γ_0 : global intercepts
- β_1, γ_1 : steepness coefficients

Spatial Post-Processing

Standardized Anomaly Model Output Statistics (SAMOS)

$$\begin{aligned}\frac{y - \text{obs}_{\mu}}{\text{obs}_{\sigma}} &\sim \mathcal{L}(\mu, \sigma) \\ \mu &= \beta_0 + \beta_1 \cdot \text{mean}\left(\frac{\text{ens} - \text{ens}_{\mu}}{\text{ens}_{\sigma}}\right) \\ \sigma &= \gamma_0 + \gamma_1 \cdot \text{stdv}\left(\frac{\text{ens} - \text{ens}_{\mu}}{\text{ens}_{\sigma}}\right)\end{aligned}\tag{5}$$

- y : observations^{1/p}
- $\text{obs}_{\mu}, \text{obs}_{\sigma}$: observed climatology
- $\mathcal{L}(\dots)$: logistic distribution
- μ, σ : latent location and scale
- ens : latest EPS forecast^{1/p}
- $\text{ens}_{\mu}, \text{ens}_{\sigma}$: EPS climatology
- β_0, γ_0 : global intercepts
- β_1, γ_1 : steepness coefficients

Spatial Post-Processing

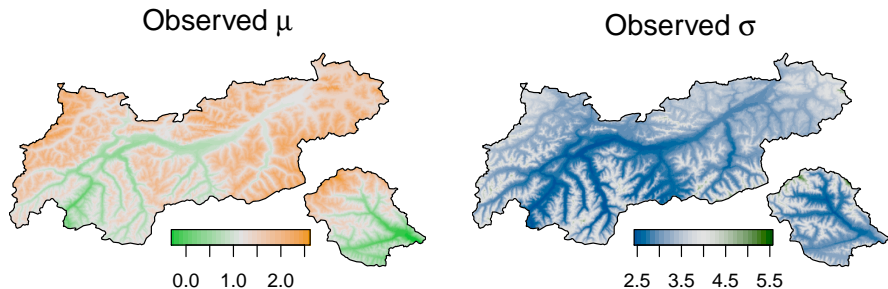


Figure 7: Spatio-Temporal Climatology, Stauffer et al. 2016.

Standardized Anomaly Model Output Statistics (SAMOS)

$$\frac{y - \text{obs}_{\mu}}{\text{obs}_{\sigma}} \sim \mathcal{L}(\mu, \sigma)$$
$$\mu = \beta_0 + \beta_1 \cdot \text{mean}\left(\frac{\text{ens} - \text{ens}_{\mu}}{\text{ens}_{\sigma}}\right)$$
$$\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv}\left(\frac{\text{ens} - \text{ens}_{\mu}}{\text{ens}_{\sigma}}\right)$$
(6)

Spatial Post-Processing

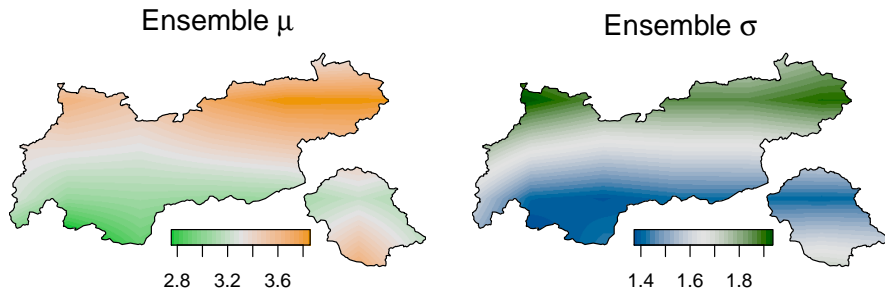


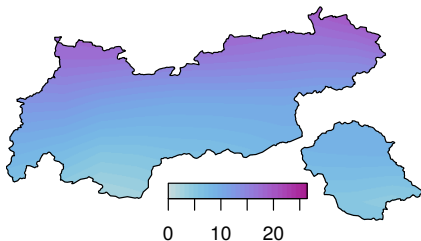
Figure 7: ECMWF ENS climatology: ECMWF reforecasts.

Standardized Anomaly Model Output Statistics (SAMOS)

$$\frac{y - obs_{\mu}}{obs_{\sigma}} \sim \mathcal{L}(\mu, \sigma)$$
$$\mu = \beta_0 + \beta_1 \cdot \text{mean}\left(\frac{ens - ens_{\mu}}{ens_{\sigma}}\right)$$
$$\sigma = \gamma_0 + \gamma_1 \cdot \text{mean}\left(\frac{ens - ens_{\mu}}{ens_{\sigma}}\right)$$
(6)

Spatial Post-Processing

Ensemble Mean



Standardized Anomaly Model Output Statistics (SAMOS)

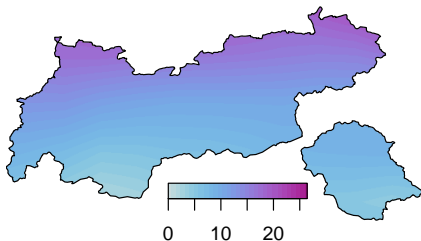
$$\frac{y - obs_{\mu}}{obs_{\sigma}} \sim \mathcal{L}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot \text{mean}\left(\frac{ens - ens_{\mu}}{ens_{\sigma}}\right) \quad (6)$$

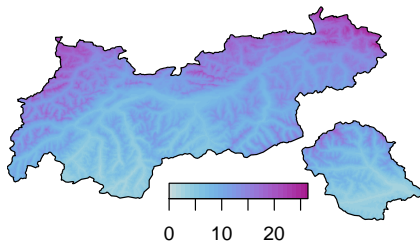
$$\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv}\left(\frac{ens - ens_{\mu}}{ens_{\sigma}}\right)$$

Spatial Post-Processing

Ensemble Mean



SAMOS Mean



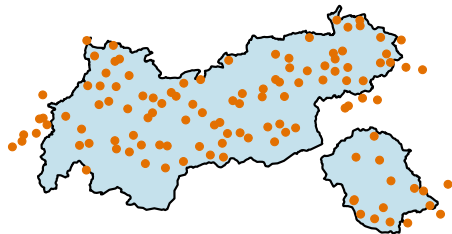
Standardized Anomaly Model Output Statistics (SAMOS)

$$\frac{y - obs_{\mu}}{obs_{\sigma}} \sim \mathcal{L}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot \text{mean}\left(\frac{ens_{\mu}}{ens_{\sigma}}\right) \quad (6)$$

$$\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv}\left(\frac{ens_{\mu}}{ens_{\sigma}}\right)$$

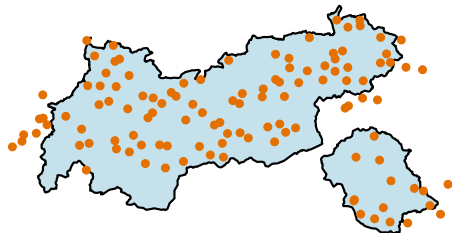
SAMOS Data & Results



Observations

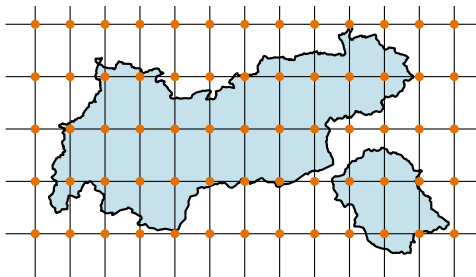
- 118 stations
- daily observations
- 1971 – 2009; 2010 – 2012

SAMOS Data & Results



Observations

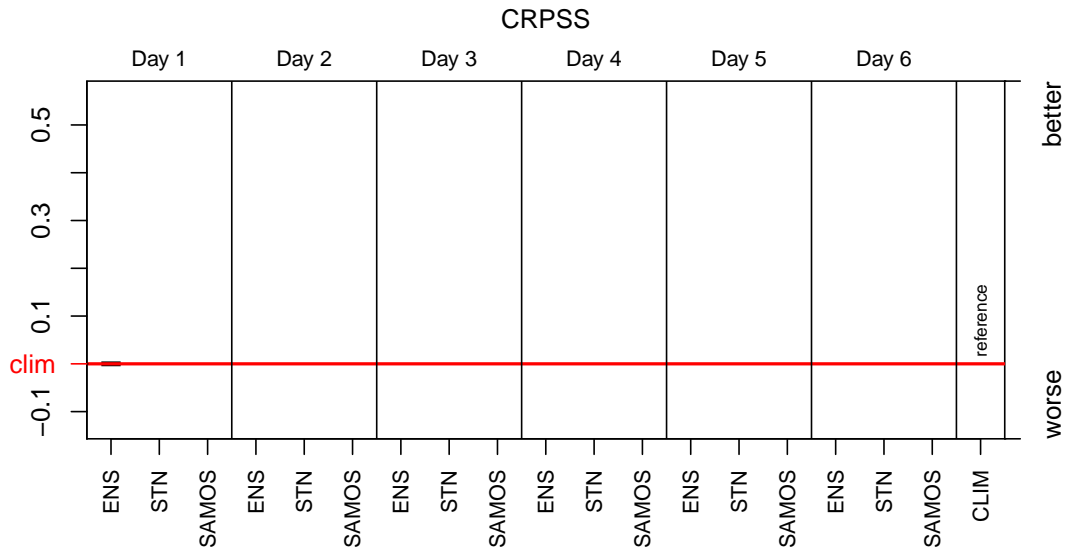
- 118 stations
- daily observations
- 1971 – 2009; 2010 – 2012



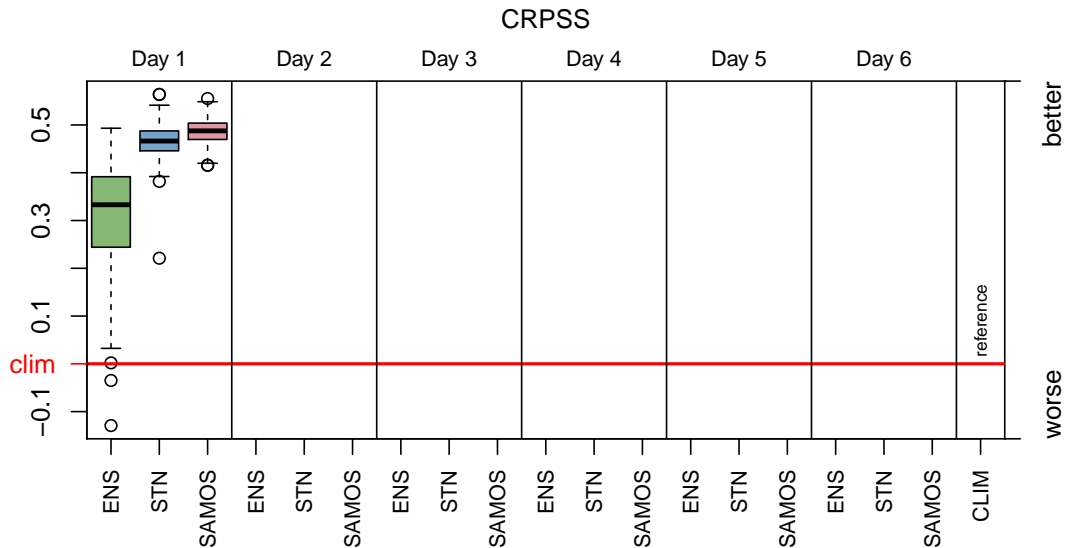
NWP data

- ECMWF ENS
- ECMWF reforecasts
- February 2010–2012
- $\Delta x/\Delta y$: 25km

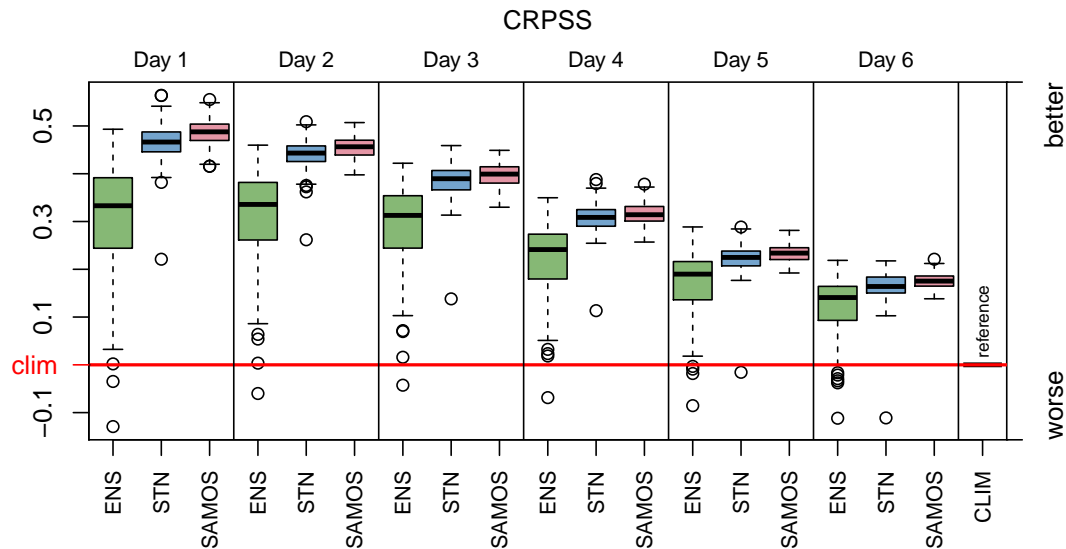
SAMOS Results: CRPSS



SAMOS Results: CRPSS



SAMOS Results: CRPSS



SAMOS Data & Results

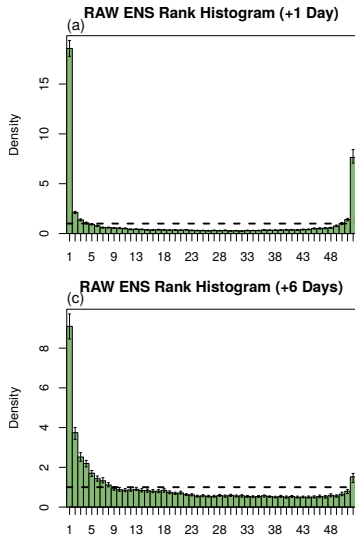


Figure 8: Rank histogram (left) and probability integral transform histograms (right) for one-day-ahead and six-day-ahead forecasts.

SAMOS Data & Results

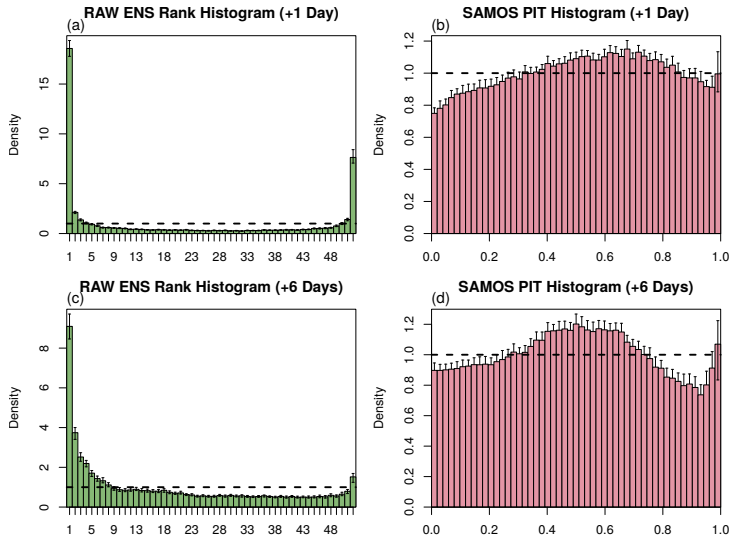
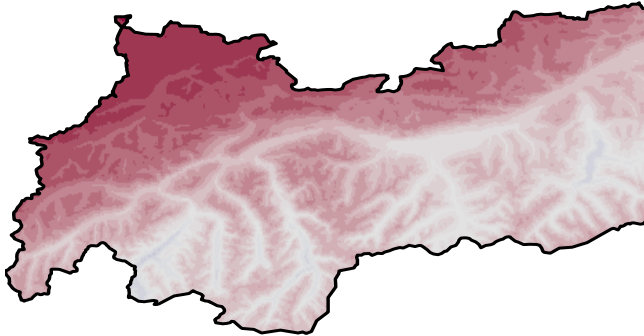


Figure 8: Rank histogram (left) and probability integral transform histograms (right) for one-day-ahead and six-day-ahead forecasts.

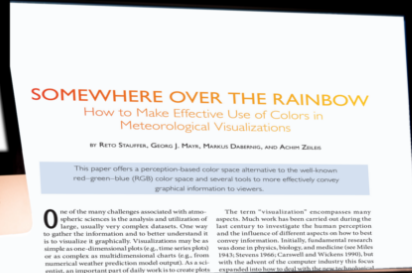
Summary II

- **concept proofed** for daily precipitation
- **accurately** predicts **full distribution**
- **outperforms station-wise** estimates



Best Student Paper Award 2015

Leopold-Franzens Universität



Publication III

Stauffer, R., G. J. Mayr, M. Dabernig, and A. Zeileis (2015): Somewhere Over the Rainbow: How to Make Effective Use of Colors in Meteorological Visualizations. *Bulletin of the American Meteorological Society*, **96**(2), 203–216.

Introduction

Color

- **integral element** in graphical displays
- **easily available** in most common software languages
- **omnipresent**: publications, presentation slides, . . .

Introduction

Color

- **integral element** in graphical displays
- **easily available** in most common software languages
- **omnipresent**: publications, presentation slides, . . .

The Problem: Only little guidance about “how to choose appropriate colors” for a particular task.

Introduction

Goal of our work

- **raise** the **awareness**
- introduce **Hue-Chroma-Luminance** (HCL) model
 - based on human perception
 - better control for choosing color palettes
- **provide information** on why and how to use
- **provide** convenient **software** for everyone

The RGB Color Space

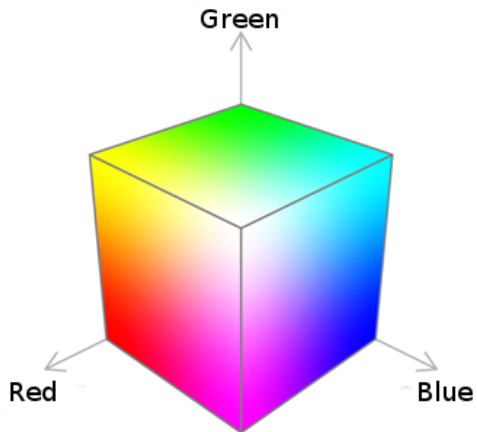


Figure 9: The Red-Green-Blue RGB color space.

The RGB Color Space

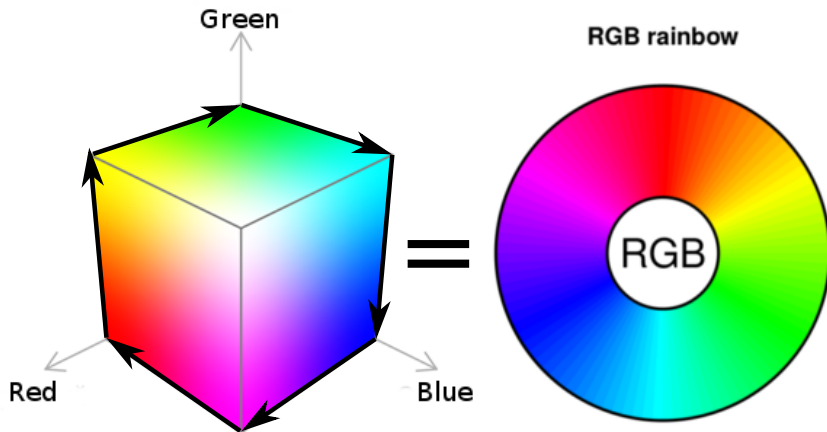
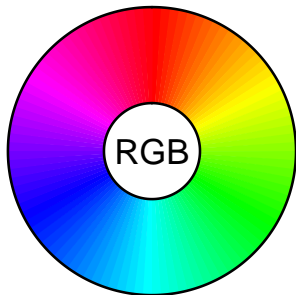


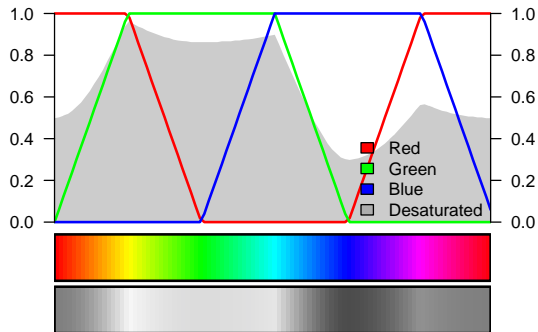
Figure 9: The (in)famous Red-Green-Blue RGB rainbow palette.

The RGB Color Space

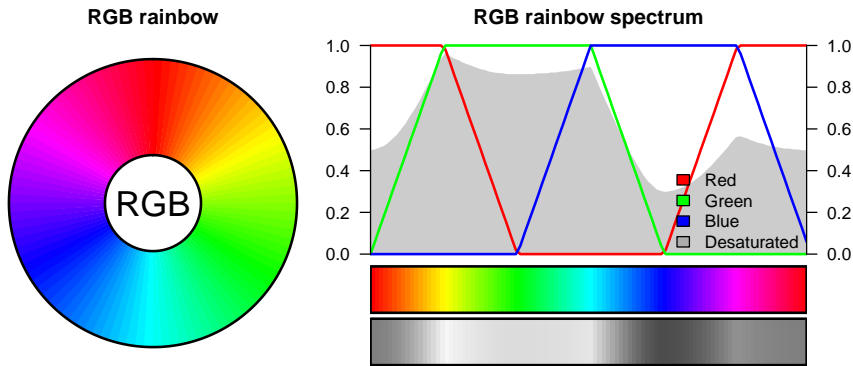
RGB rainbow



RGB rainbow spectrum

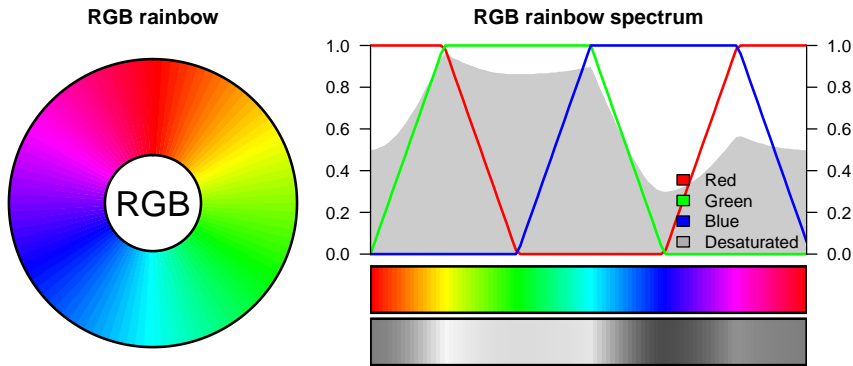


The RGB Color Space



- **default color map** in many software packages
- **conveniently used** by many practitioners

The RGB Color Space

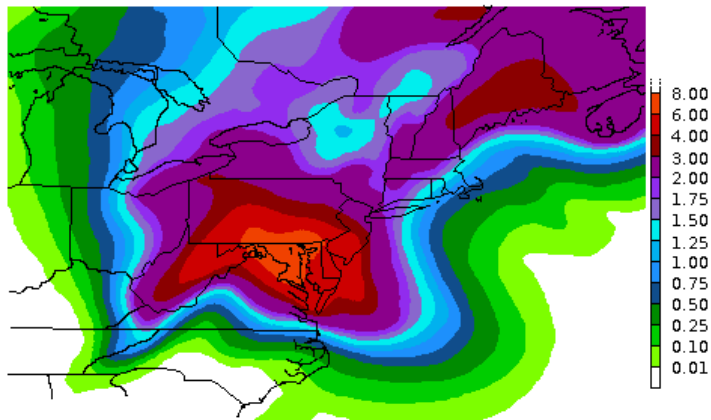


- **default color map** in many software packages
- **conveniently used** by many practitioners

Question

Everybody does it – why should it be wrong?

What is Wrong?

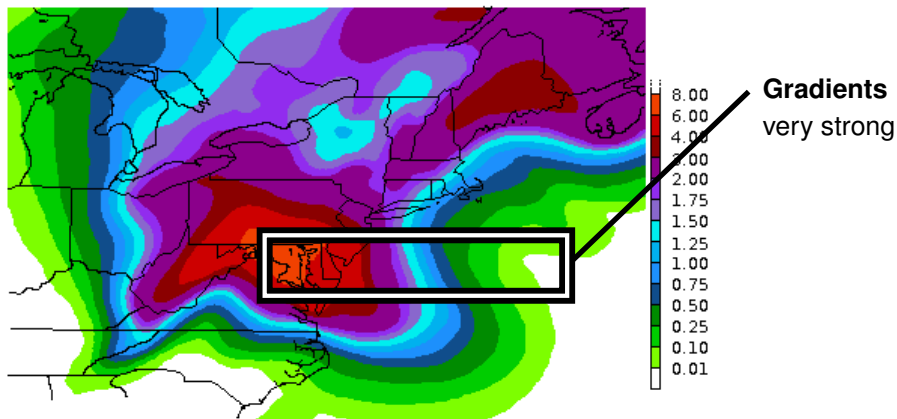


**Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast**

Original figure as published by NOAA.

NOAA forecast, www.noaa.gov, 2012-10-27.

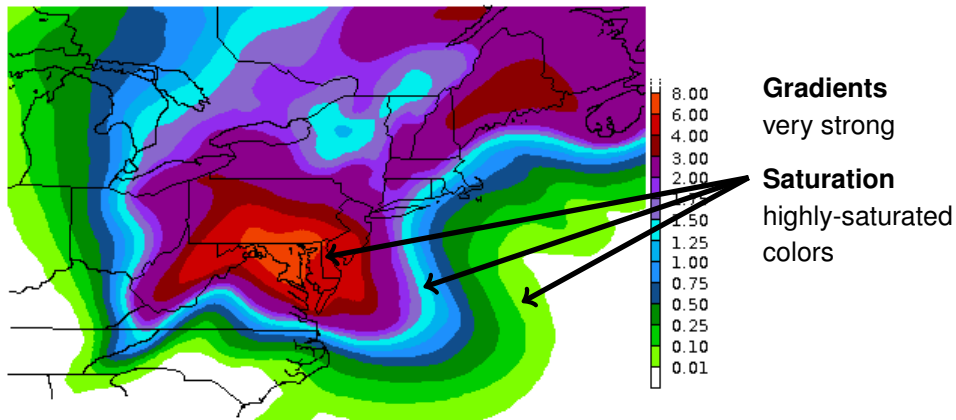
What is Wrong?



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Original figure as published by NOAA.

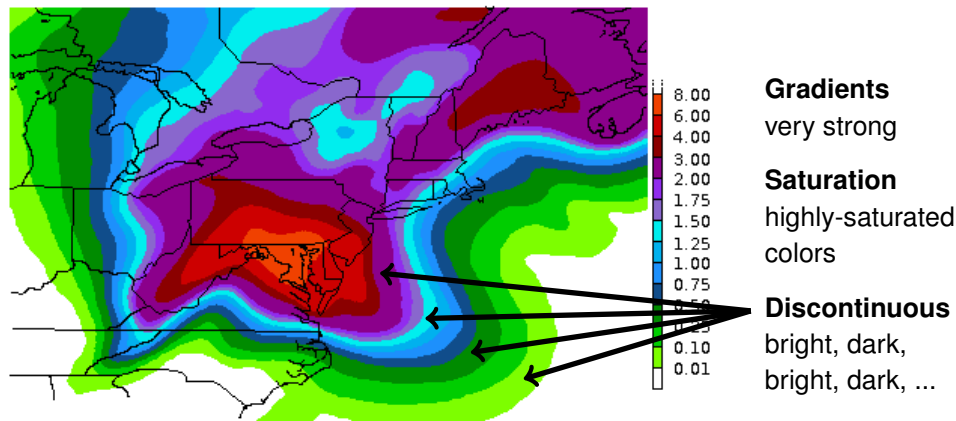
What is Wrong?



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Original figure as published by NOAA.

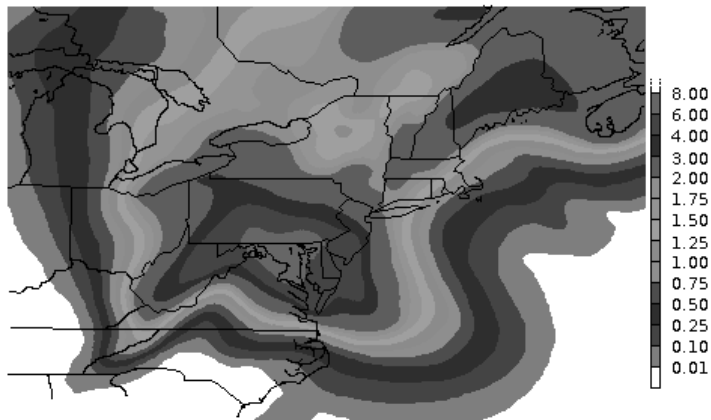
What is Wrong?



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Original figure as published by NOAA.

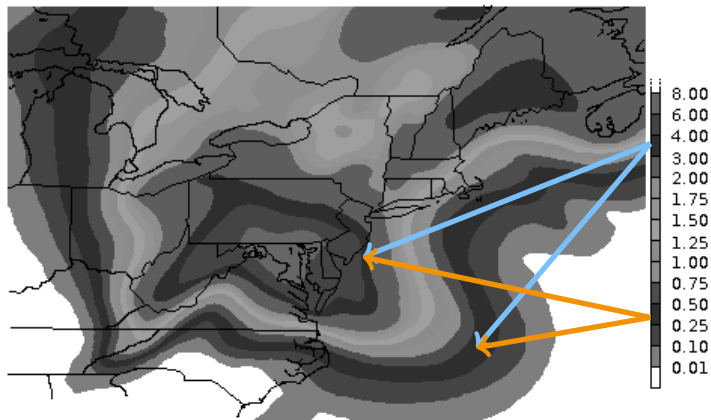
What is Wrong?



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Desaturated version of the original figure.

What is Wrong?

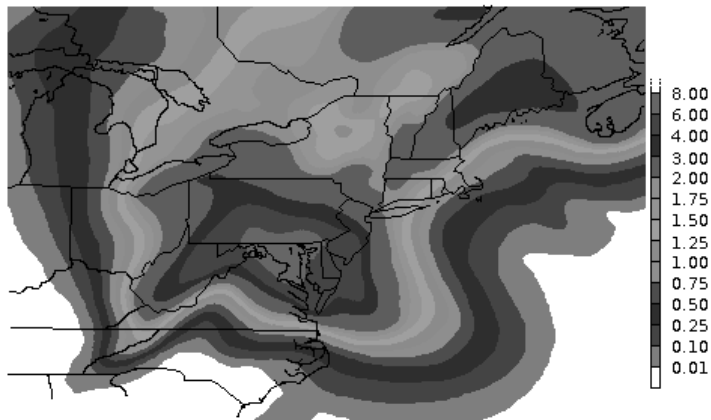


Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Desaturated version of the original figure.

Assignment
no longer unique

What is Wrong?



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Desaturated version of the original figure.

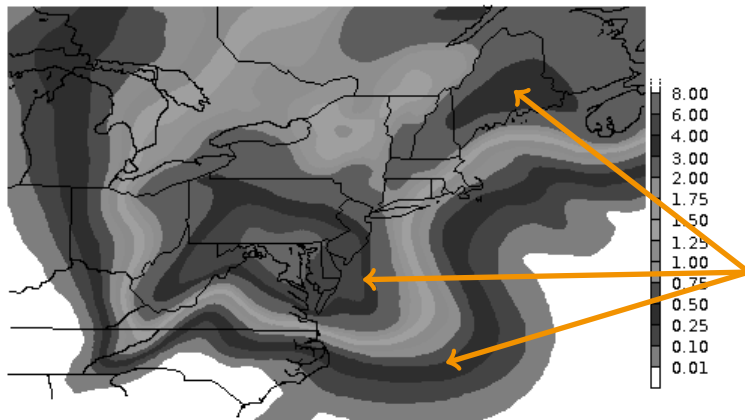
Assignment

no longer unique

Interpretation

where is the maximum?

What is Wrong?



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Desaturated version of the original figure.

Assignment

no longer unique

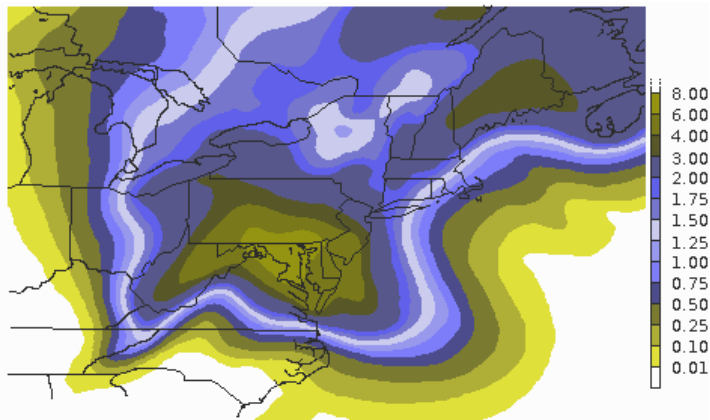
Interpretation

where is the maximum?

Focus

on dark artifacts

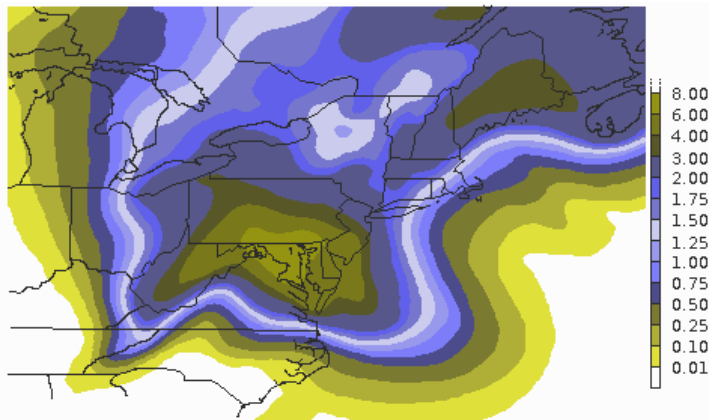
What's wrong?



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Protanope sight (red-green weakness).
About 5% of all Europeans are affected.

What's wrong?

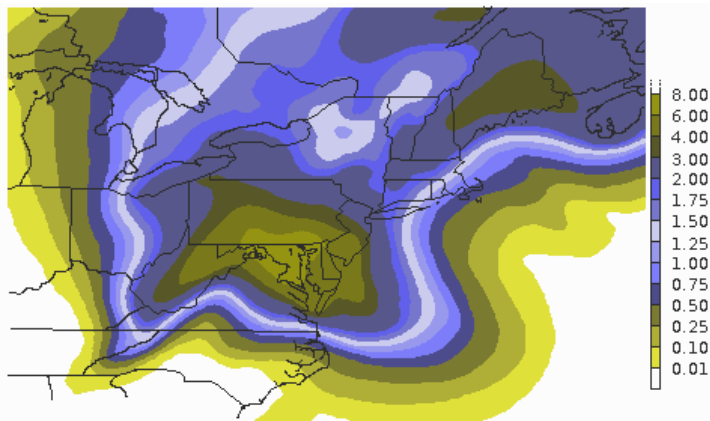


Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

End-user
who is it?

Protanope sight (red-green weakness).
About 5% of all Europeans are affected.

What's wrong?



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

*Protanopia sight (red-green weakness).
About 5% of all Europeans are affected.*

End-user
who is it?

To regard
visual constraints?

Challenges

Summary: Colors in a palette should ...

- be **simple** and **natural**,
- not be un**appealing**,
- **highlight** the **important** information,
- **not mislead** the reader,
- **work everywhere** and for everyone.

Challenges

Summary: Colors in a palette should ...

- be **simple** and **natural**,
- not be un**appealing**,
- **highlight** the **important** information,
- **not mislead** the reader,
- **work everywhere** and for everyone.

In practice

- People often **do not think** about it at all
- ... and simply **use default** colors.

Challenges

Summary: Colors in a palette should ...

- be **simple** and **natural**,
- not be un**appealing**,
- **highlight** the **important** information,
- **not mislead** the reader,
- **work everywhere** and for everyone.

In practice

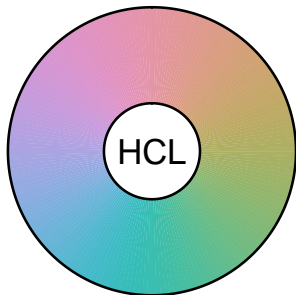
- People often **do not think** about it at all
- ... and simply **use default** colors.

Potential problems

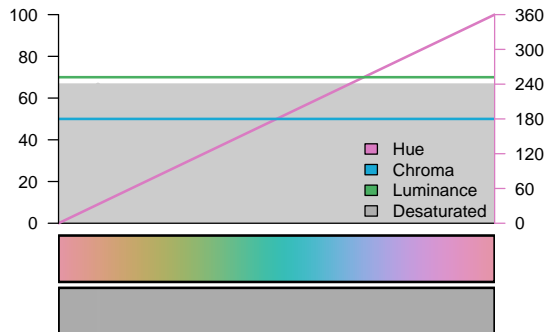
- for end-users – reviewer, supervisor, colleague, customer
- for your own day-to-day work

Perception-Based Way: HCL

A HCL rainbow

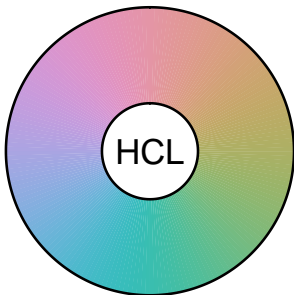


HCL rainbow spectrum

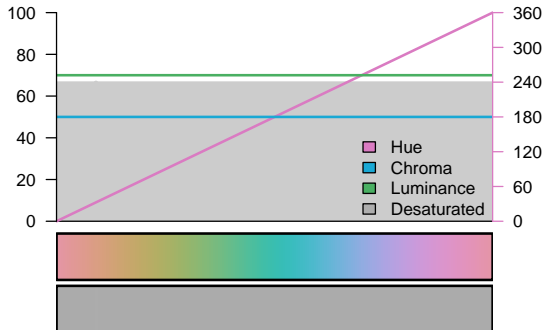


Perception-Based Way: HCL

A HCL rainbow



HCL rainbow spectrum

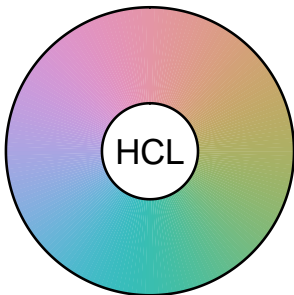


Triplet of:

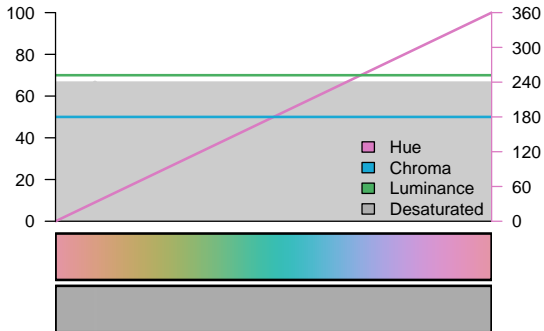
- **Hue** (*defines the color*)

Perception-Based Way: HCL

A HCL rainbow



HCL rainbow spectrum

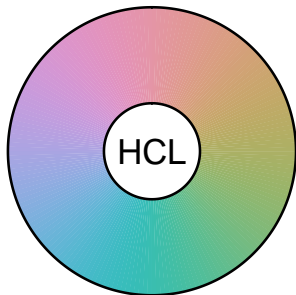


Triplet of:

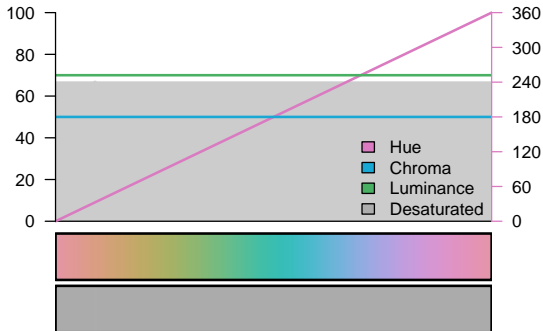
- **Hue** (*defines the color*)
- **Chroma** (*defines the colorness*) and

Perception-Based Way: HCL

A HCL rainbow



HCL rainbow spectrum



Triplet of:

- **Hue** (*defines the color*)
- **Chroma** (*defines the colorness*) and
- **Luminance** (*defines the brightness*)

Perception-Based Way: HCL

Advantages

- based on **human perception**

Perception-Based Way: HCL

Advantages

- based on **human perception**
- easy to control

Perception-Based Way: HCL

Advantages

- based on **human perception**
- easy to control
- simple to use

Perception-Based Way: HCL

Advantages

- based on **human perception**
- easy to control
- simple to use
- improving **readability, clarity**

Perception-Based Way: HCL

Advantages

- based on **human perception**
- easy to control
- simple to use
- improving **readability, clarity**
- helps transferring **complex issues**

Perception-Based Way: HCL

Advantages

- based on **human perception**
- easy to control
- simple to use
- improving **readability, clarity**
- helps transferring **complex issues**

Hue



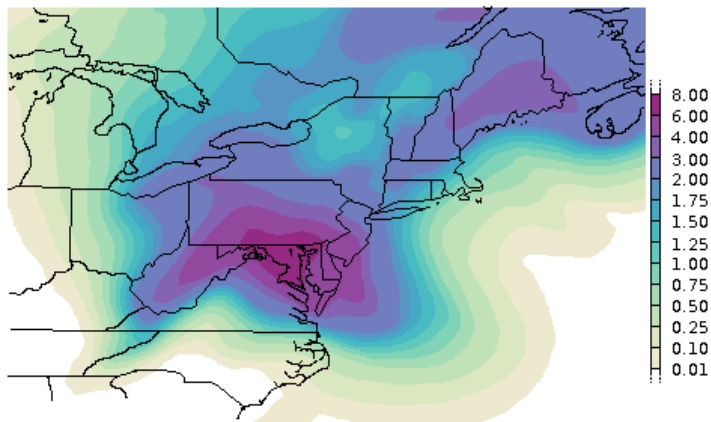
Chroma



Luminance



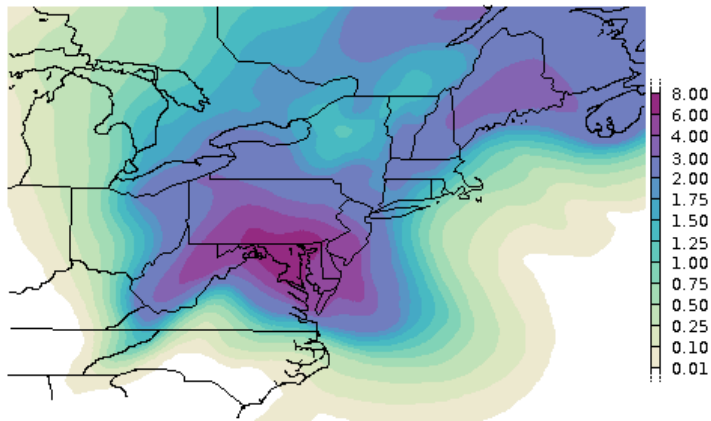
HCL Version



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Same information, changed color scheme.

HCL Version



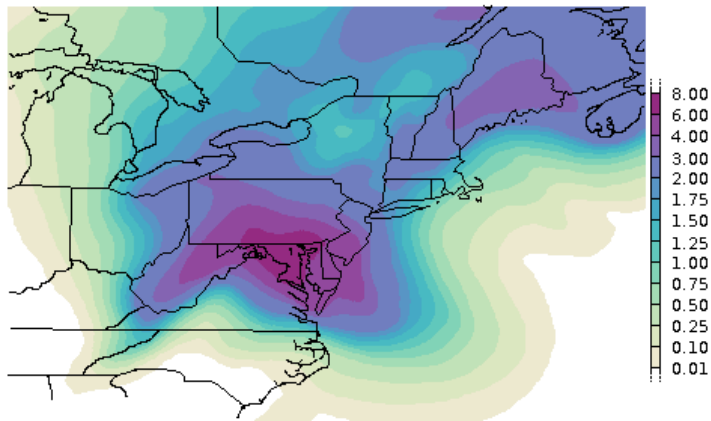
Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Same information, changed color scheme.

Colors

only two colors;
no irritating gradients

HCL Version



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Same information, changed color scheme.

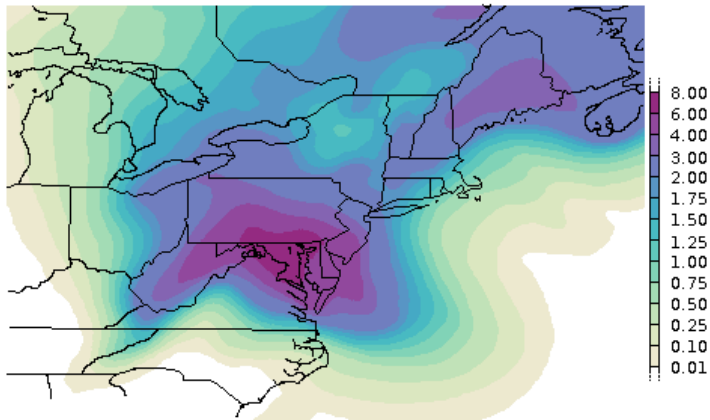
Colors

only two colors;
no irritating gradients

Information

guiding; no hidden
information

HCL Version



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Same information, changed color scheme.

Colors

only two colors;
no irritating gradients

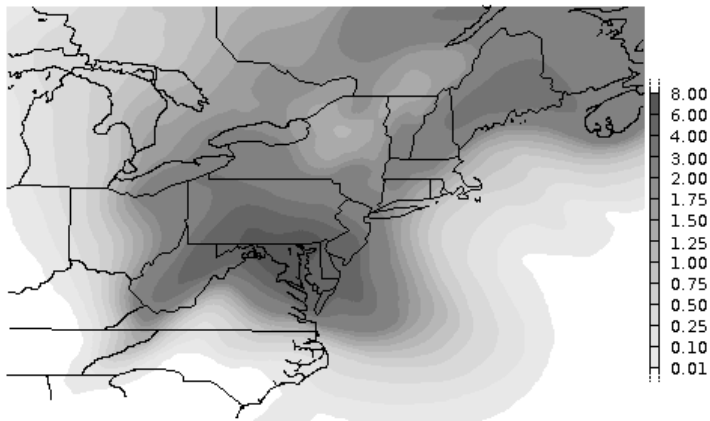
Information

guiding; no hidden
information

Works

screen; projector;
gray-scale device

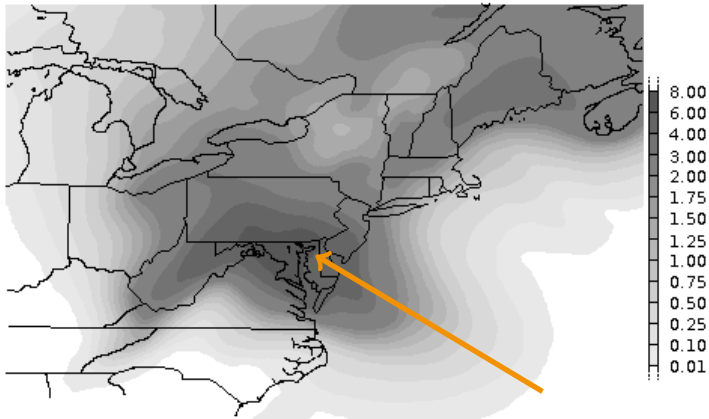
HCL Version



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Desaturated sight of the HCL-version.

HCL Version



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Desaturated sight of the HCL-version.

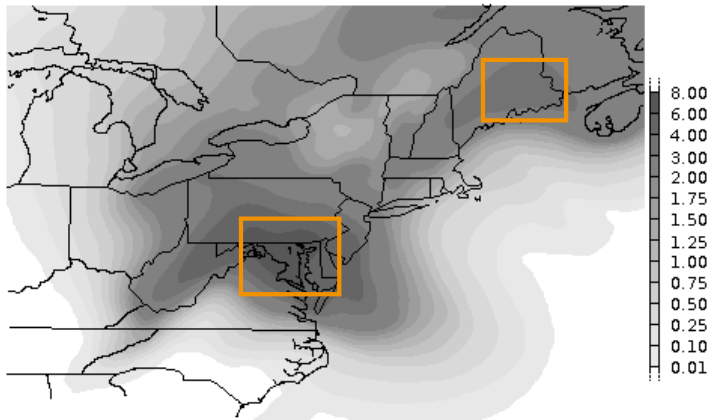
Assignment

higher values

(more precipitation)

→ lower luminance

HCL Version



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Desaturated sight of the HCL-version.

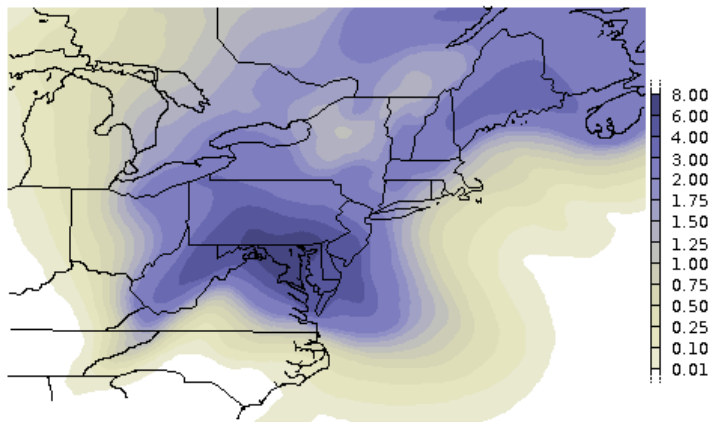
Assignment

higher values
(more precipitation)
→ lower luminance

Focus

leads readers to most important areas

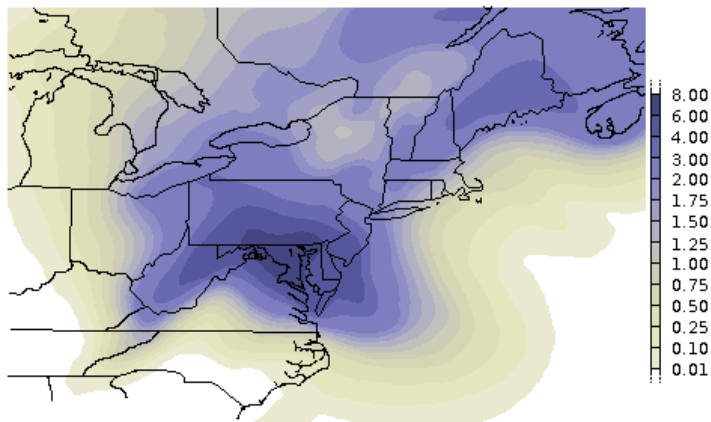
HCL Version



Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Protanope sight of the HCL-version.

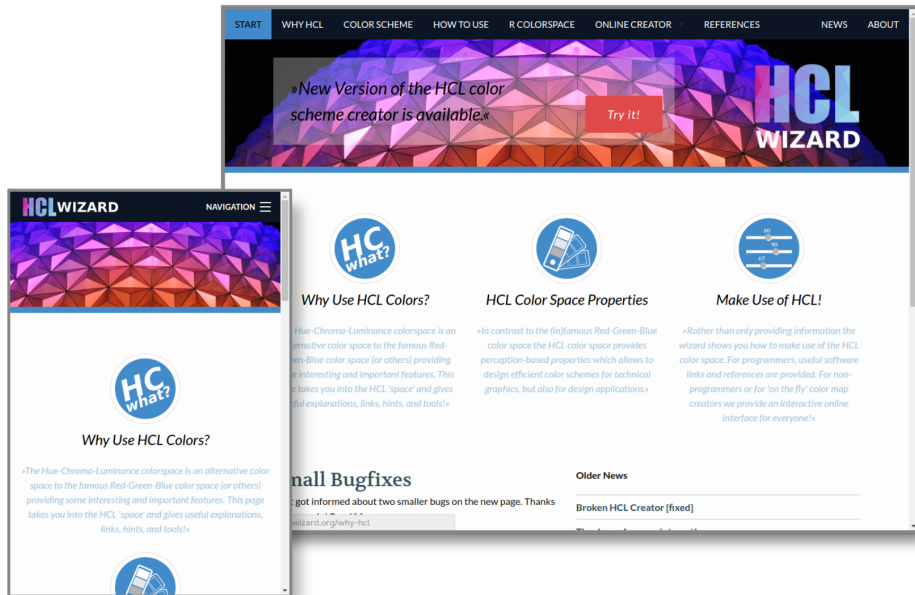
HCL Version

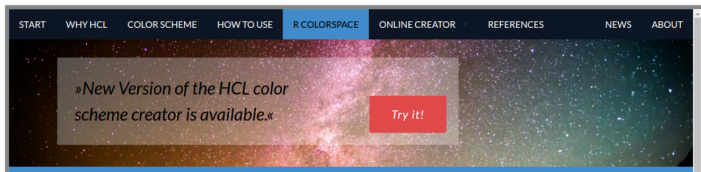


Hurricane Sandy
120-hour Day 1-5 Rainfall Forecast

Summary

Solved a lot of problems by changing the color palette.





colorspace Package

Using the colorspace package in R

The open source software R provides a package called [Ihaka, Murrell, Hornik, Fisher, & Zeileis, 2016] which uses Ihaka's colorspace library. The package offers some preset color palettes (rainbow_hcl, terrain_hcl, heat_hcl, ...) to compare to the default RGB color palettes. Furthermore there is a graphical user interface (GUI) where you can design your own color palette. The function therefore is called choose_palette() (needs some dependent packages) and use them for your own work.

The first thing you need is an R installation on your computer. The installation packages for all available operating systems (Windows, OSX, Linux) can be found on the <http://cran.r-project.org>. The installation is no adventure and the R base version just needs a few MB of space. For beginners the R-Studio GUI is recommended (for a similar R editor).

If R is successfully installed on your system you can install optional packages like the colorspace package. You can simply do that over your GUI interface or install them by using the following code line:

```
install.packages('colorspace')
```



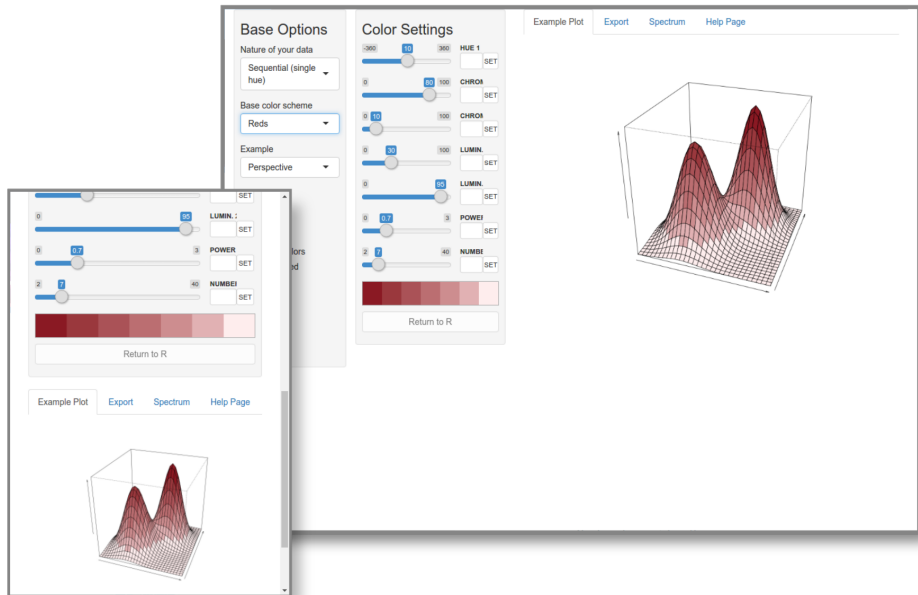
»I'm an advanced R-user, but I never heard about the colorspace package. How can I use this awesome feature?«



Download R Package

The current version of the R colorspace package can be downloaded from the cran website:
Visit: [cran package colorspace](http://cran.r-project.org/web/packages/colorspace/index.html)





Example Plot

Export

Spectrum

Help Page

RAW

GrADS

Python

matlab

RGB values [0-1]

RGB values [0-255]

Download

Download

HEX colors, no alpha

Color Map

Base Options

Nature of your data

Sequential (single hue)

Base color scheme

Reds

Example

Perspective

Color Settings

HUE 1

300

10

300

SET

CHROM

0

50

100

SET

CHROM

0

10

100

SET

LUMIN

0

50

100

SET

LUMIN

0

50

100

SET

POWER

0

0.7

3

SET

NUMBER

2

7

40

SET

Return to R

Example Plot

Export

Spectrum

Help Page

RAW

GrADS

Python

matlab

```

%% Define rgb matrix first (matrix size ncolors x 3)
colors = [0.541,0.098,0.137;
          0.588,0.192,0.216;
          0.635,0.271,0.290;
          0.682,0.349,0.365;
          0.733,0.431,0.443;
          0.784,0.522,0.529;
          0.839,0.620,0.627;
          0.902,0.737,0.741;
          0.996,0.929,0.929]

%% Create surface data
[X,Y] = meshgrid(-8:5:8);
R = sqrt(X.^2 + Y.^2) + eps;
Z = sin(R)./R;

%% Plotting surface
surf(X,Y,Z,'EdgeColor','black')

%% Adding your color palette and colorbar
colormap(colors)
colorbar()

```

Summary III

Choice of colors

- **use** color **with care**, do not underestimate power of color
- think about **who** the reader/users are
- avoid large areas of **flashy, highly-saturated** colors
- employ **monotonic luminance scale** for continuous data

Summary III

Choice of colors

- **use** color **with care**, do not underestimate power of color
- think about **who** the reader/users are
- avoid large areas of **flashy, highly-saturated** colors
- employ **monotonic luminance scale** for continuous data

Try it yourself

- <https://hclwizard.org>
 - interactive
 - several export options
 - information and guidance
- *colorspace* in *R*

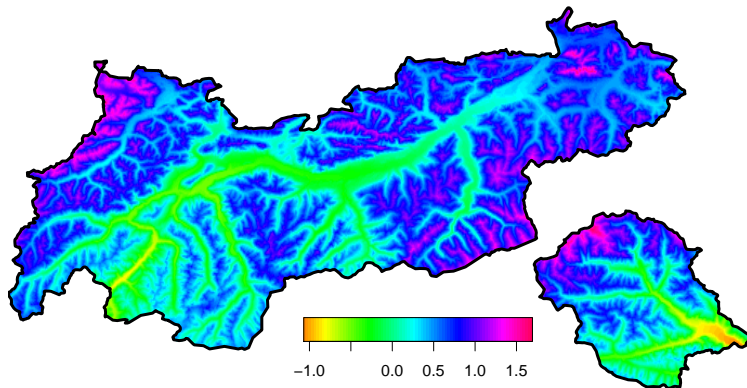


To Recapitulate



To Recapitulate

Climatology: latent location μ

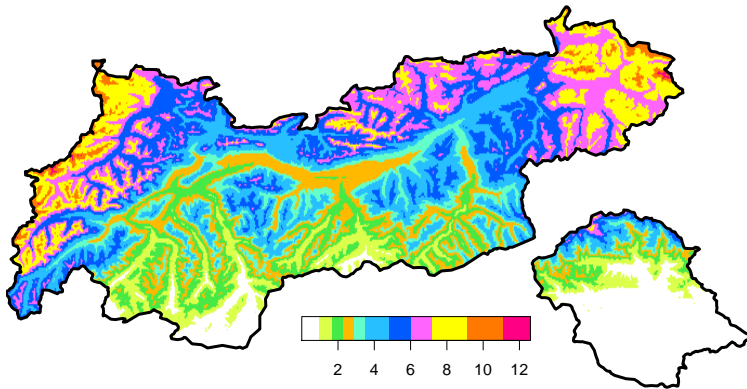


Part I

Spatio-Temporal Precipitation Climatology over Complex Terrain Using a Censored Additive Regression Model.

To Recapitulate

Precipitation forecast [mm/day]

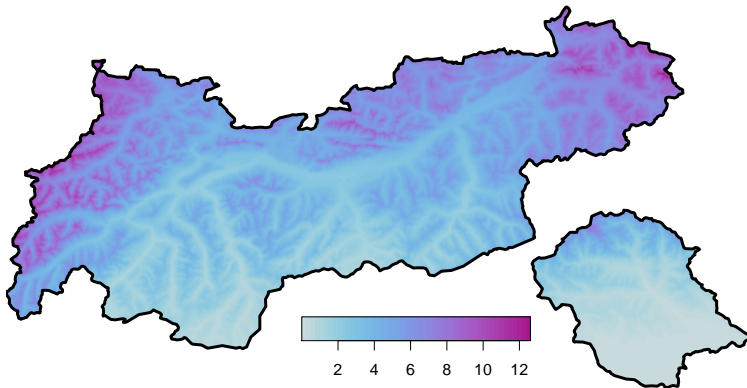


Part II

Ensemble Post-Processing of Daily Precipitation Sums over Complex Terrain Using Censored High-Resolution Standardized Anomalies.

To Recapitulate

Precipitation forecast [mm/day]



Part III

Somewhere Over the Rainbow: How to Make Effective Use of Colors in Meteorological Visualizations.










Thank you for your attention!



Special thanks to the *SWAT*, my colleagues and advisors, my family and the FWF which made this possible!

References I

-  Amt der Tiroler Landesregierung, 2014: Statistisches Jahrbuch Bundesland Tirol. Accessed: July 15 2016, Available at https://www.tirol.gv.at/fileadmin/themen/statistik-budget/statistik/downloads/Statistisches_Handbuch_2014.pdf.
-  BMLFUW, 2016: Bundesministerium für Land und Forstwirtschaft, Umwelt und Wasserwirtschaft (BMLFUW), Abteilung IV/4 - Wasserhaushalt. Accessed: February 29 2016, Available at <http://ehyd.gv.at>.
-  Cohen, A. C. J., 1959: Simplified estimators for the normal distribution when samples are singly censored or truncated. *Technometrics*, **1** (3), 217–237, doi:10.1080/00401706.1959.10489859.
-  Dabernig, M., G. J. Mayr, J. W. Messner, and A. Zeileis, 2016: Spatial ensemble post-processing with standardized anomalies. Working papers, Atmospheric and Cryospheric Institute, University of Innsbruck. URL <http://EconPapers.repec.org/RePEc:inn:wpaper:2016-08>.
-  Gneiting, T., F. Balabdaoui, and A. E. Raftery, 2007: Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **69** (2), 243–268, doi:10.1111/j.1467-9868.2007.00587.x.
-  Hagedorn, R., R. Buizza, T. M. Hamill, M. Leutbecher, and T. N. Palmer, 2012: Comparing TIGGE multimodel forecasts with reforecast-calibrated ECMWF ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, **138** (668), 1814–1827, doi:10.1002/qj.1895.
-  Lawinenwarndienst Tirol, 2009–2015: Winterberichte 2009/2010 bis 2014/2015. Accessed: July 15 2016, Available at <https://lawine.tirol.gv.at/archiv/winterberichte/>.

References II



Mullen, S. L. and R. Buizza, 2001: Quantitative precipitation forecasts over the United States by the ECMWF ensemble prediction system. *Monthly Weather Review*, **129** (4), 638–663, doi:10.1175/1520-0493(2001)129<0638:QPF0TU>2.0.CO;2.



Stauffer, R., G. J. Mayr, M. Dabernig, and A. Zeileis, 2015: Somewhere over the rainbow: How to make effective use of colors in meteorological visualizations. *Bulletin of the American Meteorological Society*, **96** (2), 203–216, doi:10.1175/BAMS-D-13-00155.1.



Stauffer, R., J. W. Messner, G. J. Mayr, N. Umlauf, and A. Zeileis, 2016a: Ensemble post-processing of daily precipitation sums over complex terrain using censored high-resolution standardized anomalies. *Monthly Weather Review*, accepted, URL <http://EconPapers.repec.org/RePEc:inn:wpaper:2016-21>.



Stauffer, R., J. W. Messner, G. J. Mayr, N. Umlauf, and A. Zeileis, 2016b: Spatio-temporal precipitation climatology over complex terrain using a censored additive regression model. *International Journal of Climatology*, in print, doi:10.1002/joc.4913.