Spatial Post-Processing over Complex Terrain using Standardized Anomalies

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

- Introduction
- Motivation: Latest Forecast
- Methodology
- The SAMOS Approach
- Results
Introduction

Numerical Weather Prediction (NWP)

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

Numerical Weather Prediction (NWP)

1. analysis: $\rightarrow$ current state
2. prognosis: $\rightarrow$ future state

Error Sources

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

Ensemble Prediction Systems

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

- 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

• to quantify the uncertainty
• number of members restricted
• typically underdispersive
Introduction

Forecast Error

- total error = noise + systematic errors
Introduction

Forecast Error

- \( \text{total error} = \text{noise} + \text{systematic errors} \)
- \( \text{noise} \): unexplainable (signal-free)
- \( \text{systematic errors} \): 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

- \textit{total error} = \textit{noise} + \textit{systematic errors}
- \textit{noise}: unexplainable (signal-free)
- \textit{systematic errors}: correction possible

Post-Processing

- correct bias
- correct uncertainty
- discrete $\rightarrow$ full distribution
- probabilities, quantiles, extremes
Final Product

To motivate

Live beta – impression of the “outcome” of the approach.
Introduction to Methodology
Methodology

Introduce you to ... Left-Censored Non-Homogeneous Generalized Spatio-Temporal Additive Regression Model for Daily Precipitation Sums Using High-Resolution Standardized Anomalies
Methodology

Introduce you to...

Left-Censored Non-Homogeneous Generalized Spatio-Temporal Additive Regression Model for Daily Precipitation Sums Using High-Resolution Standardized Anomalies
Methodology

IBK $T_{12} + 5d$

-10 0 10 20 30
-10 0 10 20 30
ensemble mean (ens)
observations (obs)

Gaussian Regression

$$\hat{\text{obs}} = \beta_0 + \beta_1 \cdot \text{ens}$$

$$\text{obs} \sim N(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot \text{ens}$$
Methodology

\[ \hat{\text{obs}} = \beta_0 + \beta_1 \cdot \text{ens} \]  

\[ \text{obs} \sim N(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \text{ens} \]

IBK T_{12} +5d
Methodology

\[ \hat{\text{obs}} = \beta_0 + \beta_1 \cdot \text{ens} \] (1)

\[ \text{obs} \sim \mathcal{N}(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \text{ens} \] (2)
Methodology

Gaussian Regression

\[
\hat{\text{obs}} = \beta_0 + \beta_1 \cdot \text{ens} \quad (1)
\]
Gaussian Regression

\[ \hat{\text{obs}} = \beta_0 + \beta_1 \cdot \overline{\text{ens}} \]  
\[ \text{obs} \sim \mathcal{N}(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \overline{\text{ens}} \]
Methodology

Gaussian Regression

\[
\hat{\text{obs}} = \beta_0 + \beta_1 \cdot \overline{\text{ens}} \quad (1)
\]

\[
\text{obs} \sim \mathcal{N}(\mu, \sigma)
\]

\[
\mu = \beta_0 + \beta_1 \cdot \overline{\text{ens}} \quad (2)
\]

\[
\sigma = \gamma_0
\]
Methodology

$$\text{Baad RR}_6 + 1d$$

Observations ($\text{obs}$) vs. ensemble mean ($\text{ens}$)

Gaussian Regression

$$\text{obs} \sim N(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot \text{ens}$$
Methodology

Baad RR₆ +1d

observations (obs)
ensemble mean (ens)

Gaussian Regression

\[ \text{observations (obs)} \sim \mathcal{N}(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \text{ens} \]
Methodology

Gaussian Regression

\[ \text{obs} \sim \mathcal{N}(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \text{ens} \]

\[ \sigma = \gamma_0 \]

(3)

Fraction:

0.48
0.34
Methodology

Gaussian Regression

\[ \text{obs} \sim \mathcal{N}(\mu, \sigma) \]
\[ \mu = \beta_0 + \beta_1 \cdot \text{ens} \]
\[ \sigma = \gamma_0 \]
Censored Gaussian Regression

\[ \text{obs} = \max(0, y) \quad \text{with} \quad y \sim \mathcal{N}(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \overline{\text{ens}} \]

\[ \sigma = \gamma_0 \] (3)
Censored Gaussian Regression

$$\text{obs} = \max(0, y) \quad \text{with} \quad y \sim \mathcal{N}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot \text{ens}$$

$$\sigma = \gamma_0$$
Methodology

Censored Non-Homogeneous Gaussian Regression

\[ \text{obs} = \max(0, y) \quad \text{with} \quad y \sim \mathcal{N}(\mu, \sigma) \]
\[ \mu = \beta_0 + \beta_1 \cdot \overline{\text{ens}} \]
\[ \sigma = \gamma_0 + \gamma_1 \cdot \text{stdv(ens)} \]
Methodology

Generalized Additive Model for Location, Shape, and Scale (GAMLSS)

\[
\text{obs} = \max(0, y) \text{ with } y \sim \mathcal{N}(\mu, \sigma)
\]

(4)
Methodology

Generalized Additive Model for Location, Shape, and Scale (GAMLSS)

\[
\text{obs} = \max(0, y) \text{ with } y \sim \mathcal{N}(\mu, \sigma)
\]

\[
\mu = \beta_0 + s_1(\text{alt}) + s_2(\text{lon, lat}) + \ldots
\]
Methodology

Generalized Additive Model for Location, Shape, and Scale (GAMLSS)

\[ \text{obs} = \max(0, y) \text{ with } y \sim \mathcal{N}(\mu, \sigma) \]
\[ \mu = \beta_0 + s_1(\text{alt}) + s_2(\text{lon, lat}) + \ldots \]
\[ \sigma = \gamma_0 + t_1(\text{doy, lon, lat}) + \ldots \]
Methodology

Generalized Additive Model for Location, Shape, and Scale (GAMLSS)

\[ \text{obs} = \max(0, y) \quad \text{with} \quad y \sim \mathcal{N}(\mu, \sigma) \]
\[ \mu = \beta_0 + s_1(\text{alt}) + s_2(\text{lon, lat}) + \ldots \]
\[ \sigma = \gamma_0 + t_1(\text{doy, lon, lat}) + \ldots \]
Methodology

To Summarize

- left-censored
- non-homogenous
- generalized
- spatio-temporal
- additive model
- anomalies
Methodology

To Summarize

- account for **physical limit** and large **fractions of** 0 observations
Methodology

To Summarize

- account for physical limit and large fractions of 0 observations
- standard deviation as a function of covariates

left-censored
non-homogenous
generalized
spatio-temporal
additive model
anomalies
Methodology

To Summarize

- account for physical limit and large fractions of 0 observations
- standard deviation as a function of covariates
- including all kind of effects (linear, multidim. splines, \ldots)
Methodology

To Summarize

- account for **physical limit** and large fractions of 0 observations
- **standard deviation** as a **function** of covariates
- **including** all kind of **effects** (linear, multidim. splines, ...)
- **if geographical & date/time covariates** included
Methodology

To Summarize

- account for **physical limit** and large fractions of 0 observations
- **standard deviation** as a **function** of covariates
- including all kind of **effects** (linear, multidim. splines, …)
- if **geographical & date/time covariates** included
- **simple** linear additive framework
Methodology

To Summarize

- account for physical limit and large fractions of 0 observations
- standard deviation as a function of covariates
- including all kind of effects (linear, multidim. splines, ...)
- if geographical & date/time covariates included
- simple linear additive framework
- stay tuned for the anomalies!
Spatial Ensemble Post-Processing: The SAMOS Approach
Data

Observations

- 118 stations
- Daily observations
- 1971 – 2012
Data

Observations

- 118 stations
- daily observations
- 1971 – 2012

NWP Model

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

Single Station

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

Single Station

- “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

Ensemble Mean

Spatial Model: Naive Assumption

\[ \text{obs} = \max(0, y) \text{ with } y \sim \mathcal{N}(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \overline{\text{ens}} \]

\[ \sigma = \gamma_0 + \gamma_1 \cdot \text{stdv}(\text{ens}) \]
Spatial Post-Processing

Spatial Model: Naive Assumption

\[
\text{obs} = \max(0, y) \quad \text{with} \quad y \sim \mathcal{N}(\mu, \sigma)
\]

\[
\mu = \beta_0 + \beta_1 \cdot \overline{\text{ens}}
\]

\[
\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv(ens)}
\]
Spatial Model: Naive Assumption

\[ \text{obs} = \max(0, y) \quad \text{with} \quad y \sim \mathcal{N}(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \bar{\text{ens}} \]  \hspace{1cm} (5)

\[ \sigma = \gamma_0 + \gamma_1 \cdot \text{stdv}(\text{ens}) \]
Spatial Post-Processing: SAMOS

Naive Assumption

- not suitable
- local features can’t be depict

1 Markus Dabernig: seminar in 3 weeks.
Spatial Post-Processing: SAMOS

Naive Assumption

• not suitable
• local features can’t be depict
⇒ different approach required
Spatial Post-Processing: SAMOS

Naive Assumption

- not suitable
- local features can’t be depict

⇒ different approach required

**SAMOS:** Post-Processing Using *Standardized Anomalies*¹

¹Markus Dabernig: seminar in 3 weeks.
Spatial Post-Processing: SAMOS

Naive Assumption

• not suitable
• local features can’t be depict
⇒ different approach required

**SAMOS**: Post-Processing Using **Standardized Anomalies**

- climatology as **background knowledge**
- local variations described by climatology

---

\(^1\)Markus Dabernig: seminar in 3 weeks.
Spatial Post-Processing: SAMOS

Naive Assumption

- not suitable
- local features can’t be depict

⇒ different approach required

**SAMOS**: Post-Processing Using **Standardized Anomalies**

- climatology as **background knowledge**
- local variations described by climatology
- climatology to **remove location-dependent features**
- bring stations to **compareable scale**

---

\(^1\)Markus Dabernig: seminar in 3 weeks.
Spatial Post-Processing: SAMOS

Reminder: Naive Assumption

$$\text{obs} = \max(0, y) \text{ with } y \sim \mathcal{N}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot \overline{\text{ens}}$$

$$\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv(ens)}$$
Reminder: Naive Assumption

\[
\text{obs} = \max(0, y) \quad \text{with} \quad y \sim \mathcal{N}(\mu, \sigma)
\]

\[
\mu = \beta_0 + \beta_1 \cdot \text{ens}
\]

\[
\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv(ens)}
\]
**Spatial Post-Processing: SAMOS**

### Reminder: Naive Assumption

\[
\text{obs} = \max(0, y) \quad \text{with} \quad y \sim \mathcal{N}(\mu, \sigma)
\]

\[
\mu = \beta_0 + \beta_1 \cdot \text{ens}
\]

\[
\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv(ens)}
\]  \hspace{1cm} (6)

### Standardized Anomaly Model Output Statistics (SAMOS)

\[
\text{obs} = \max(0, y) \quad \text{with} \quad \frac{y - \text{obs}_\mu}{\text{obs}_\sigma} \sim \mathcal{N}(\mu, \sigma)
\]

\[
\mu = \beta_0 + \beta_1 \cdot \text{mean}\left(\frac{\text{ens} - \text{ens}_\mu}{\text{ens}_\sigma}\right)
\]

\[
\sigma = \gamma_0 + \gamma_1 \cdot \text{stdv}\left(\frac{\text{ens} - \text{ens}_\mu}{\text{ens}_\sigma}\right)
\]  \hspace{1cm} (7)
SAMOS Results

Observed $\mu$

Observed $\sigma$

Figure: Climatology of observations: Stauffer et al. 2016.

Standardized Anomaly Model Output Statistics (SAMOS)

$$\text{obs} = \max(0, y) \text{ with } \frac{y - \text{obs}_\mu}{\text{obs}_\sigma} \sim \mathcal{N}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot \text{mean}(\frac{\text{ens} - \text{ens}_\mu}{\text{ens}_\sigma})$$

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

(8)
**SAMOS Results**

**Ensemble $\mu$**

**Ensemble $\sigma$**

**Figure:** ECMWF ENS climatology: ECMWF reforecasts.

**Standardized Anomaly Model Output Statistics (SAMOS)**

$\text{obs} = \max(0, y)$ with $\frac{y - \text{obs}_\mu}{\text{obs}_\sigma} \sim \mathcal{N}(\mu, \sigma)$

$\mu = \beta_0 + \beta_1 \cdot \text{mean}(\frac{\text{ens} - \text{ens}_\mu}{\text{ens}_\sigma})$

$\sigma = \gamma_0 + \gamma_1 \cdot \text{mean}(\frac{\text{ens} - \text{ens}_\mu}{\text{ens}_\sigma})$  \(8\)
Standardized Anomaly Model Output Statistics (SAMOS)

\[ \text{obs} = \max(0, y) \quad \text{with} \quad \frac{y - \text{obs}_\mu}{\text{obs}_\sigma} \sim \mathcal{N}(\mu, \sigma) \]

\[ \mu = \beta_0 + \beta_1 \cdot \text{mean}(\frac{\text{ens} - \text{ens}_\mu}{\text{ens}_\sigma}) \]

\[ \sigma = \gamma_0 + \gamma_1 \cdot \text{stdv}(\frac{\text{ens} - \text{ens}_\mu}{\text{ens}_\sigma}) \]
SAMOS Results

Model Comparison

- **CRPS skill score**: full distribution
- **Mean Absolute Error**: deterministic score
- **Brier Score**: probability $\text{obs} > \tau$
SAMOS Results

Model Comparison

- **CRPS skill score**: full distribution
- **Mean Absolute Error**: deterministic score
- **Brier Score**: probability $obs > \tau$

Reference Methods

- **ENS**: uncorrected ECMWF ENS
- **STN**: stationwise regression model
- **SAMOS**: spatial regression model (loo)
- **CLIM**: climatological estimates
SAMOS Results

CRPS Skill Scores

-0.1 clim 0.1

ENS STN SAMOS ENS STN SAMOS ENS STN SAMOS ENS STN SAMOS
Note: ECEPS worse than the climatology in e.g., Axams, Hall, Imst, Oetz.
SAMOS Results

Mean Absolute Errors

- ENS
- STN
- SAMOS

+030 +054 +078 +102 +126 +150
SAMOS Results

Brier Score [>0mm]

Day 1
Day 2
Day 3
Day 4
Day 5
Day 6

better
worse
SAMOS Results

Brier Score [>10mm]

Day 1  Day 2  Day 3  Day 4  Day 5  Day 6

ENS  STN  SAMOS

ENS  STN  SAMOS

ENS  STN  SAMOS

ENS  STN  SAMOS

CLIM

better  worse
Summary & Outlook

SAMOS Approach

- concept proofed for daily precipitation
- accurately predicts full distribution
- outperforms station-wise estimates
- all historical observations included
- ECMWF reforecast: always on latest NWP cycle

Outlook

- e.g. wind direction dependent climatologies
- include additional predictors
- precipitation ⇒ new snow
Summary & Outlook

SAMOS Approach

- concept proofed for daily precipitation
- accurately predicts full distribution
- outperforms station-wise estimates
- all historical observations included
- ECMWF reforecast: always on latest NWP cycle

Outlook

- e.g. wind direction dependent climatologies
- include additional predictors
- precipitation $\Rightarrow$ new snow
Thank you for your attention!

Special thanks to the SWAT, my colleagues, and advisors!