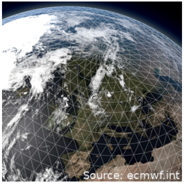
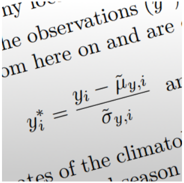
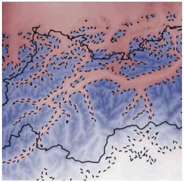
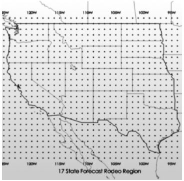
Standardized Anomaly Model Output Statistics Over Complex Terrain

Reto.Stauffer@uibk.ac.at

Outline





- statistical ensemble postprocessing
- introduction to **SAMOS**
- new snow amount forecasts in Tyrol
- sub-seasonal anomaly prediction over the U.S.

Ensemble Postprocessing

Numerical Weather Prediction

1. analysis: $\rightarrow$ current state
2. forecast: $\rightarrow$ 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

Ensemble Postprocessing

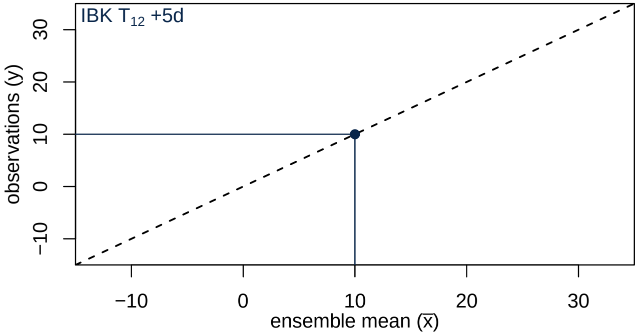
Forecast Error

- *total error = noise + systematic errors*
- *noise*: unsystematic error
- *systematic errors*: correction possible

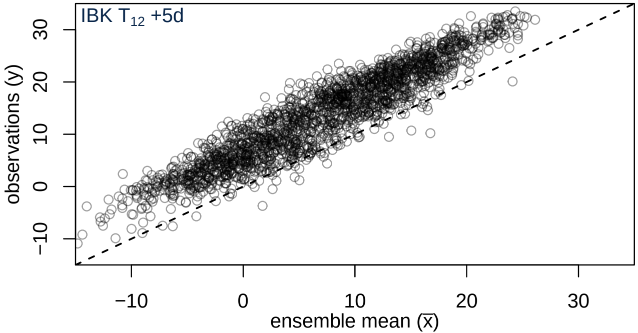
Ensemble Postprocessing

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

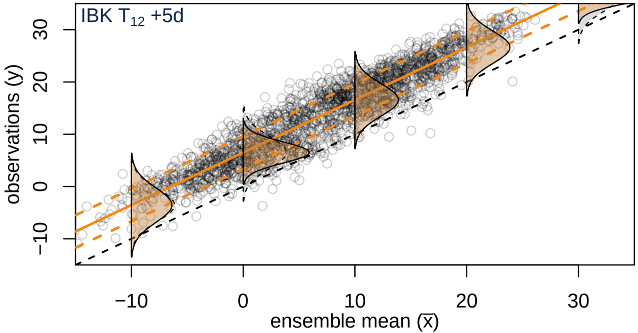
Ensemble Postprocessing



Ensemble Postprocessing



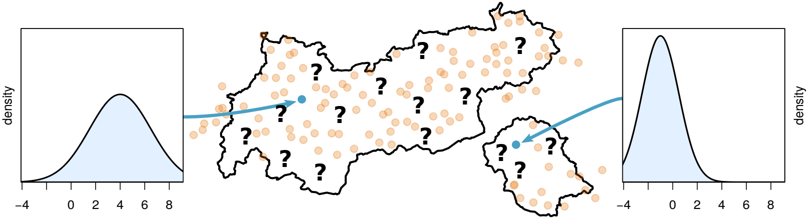
Ensemble Postprocessing



Non-homogeneous Gaussian Regression (NGR, EMOS)

$$y \sim \mathcal{N}(\mu, \sigma)$$
$$\mu = \beta_0 + \beta_1 \cdot \bar{x}$$
$$\log(\sigma) = \gamma_0 + \gamma_1 \cdot \log(s_x)$$

Spatial Postprocessing



NGR

- allows **station-wise** corrections
- **not suitable** for **spatial** predictions

Alternative approach required.

SAMOS

What is SAMOS?

SAMOS

What is SAMOS?

“A **probabilistic spatio-temporal ensemble postprocessing method** using climatological background information to **remove site specific characteristics**, which allows to **estimate one** simple **regression model** for all stations and forecast lead times **at once**.”

SAMOS

Reminder: NGR

$$y \sim \mathcal{N}(\mu, \sigma)$$
$$\mu = \beta_0 + \beta_1 \cdot \bar{x}$$
$$\log(\sigma) = \gamma_0 + \gamma_1 \cdot \log(s_x)$$

NGR $\rightarrow$ SAMOS

Transform all quantities (y, x) into standardized anomalies:

$$y^* = \frac{y - \tilde{\mu}_y}{\tilde{\sigma}_y}, \quad x^* = \frac{x - \tilde{\mu}_x}{\tilde{\sigma}_x}$$

y^*, x^*: standardized anomalies

$\tilde{\mu}_\bullet$, $\tilde{\sigma}_\bullet$: climatological properties of y, x

SAMOS

Gaussian SAMOS

$$y^* \sim \mathcal{N}(\mu^*, \sigma^*)$$
$$\mu^* = \beta_0 + \beta_1 \cdot \bar{x}^*$$
$$\log(\sigma^*) = \gamma_0 + \gamma_1 \cdot \log(s_x^*)$$

NGR $\rightarrow$ SAMOS

Transform all quantities (y, x) into standardized anomalies:

$$y^* = \frac{y - \tilde{\mu}_y}{\tilde{\sigma}_y}, \quad x^* = \frac{x - \tilde{\mu}_x}{\tilde{\sigma}_x}$$

y^*, x^*: standardized anomalies

$\tilde{\mu}_\bullet$, $\tilde{\sigma}_\bullet$: climatological properties of y, x

SAMOS

Gaussian SAMOS

$$y^* \sim \mathcal{N}(\mu^*, \sigma^*)$$
$$\mu^* = \beta_0 + \beta_1 \cdot \bar{x}^*$$
$$\log(\sigma^*) = \gamma_0 + \gamma_1 \cdot \log(s_x^*)$$

NGR $\rightarrow$ SAMOS

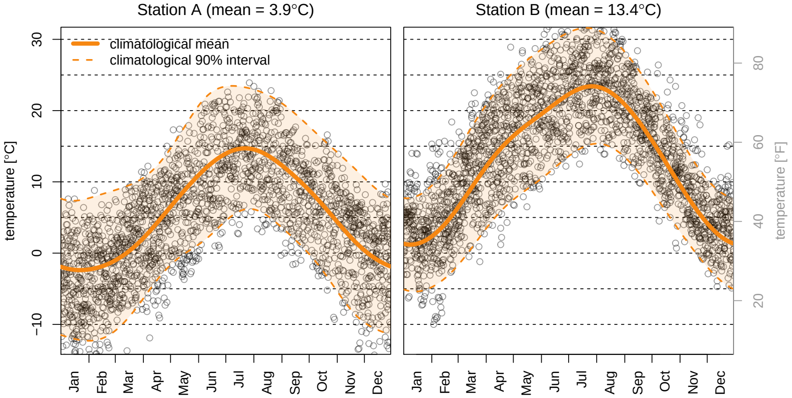
Transform all quantities (y, x) into standardized anomalies:

$$y^* = \frac{y - \tilde{\mu}_y}{\tilde{\sigma}_y}, \quad x^* = \frac{x - \tilde{\mu}_x}{\tilde{\sigma}_x}$$

y^*, x^*: standardized anomalies

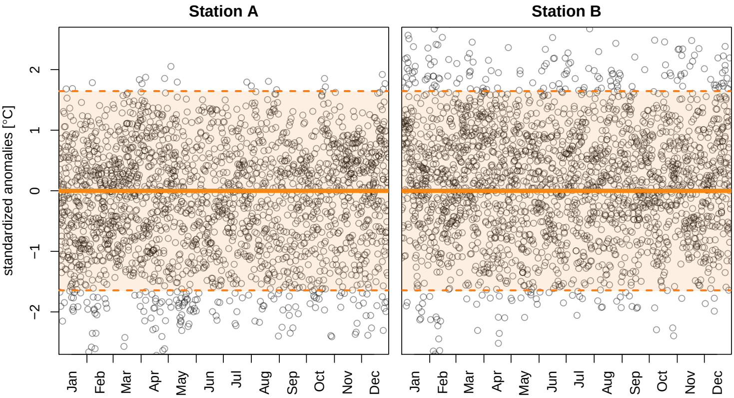
$\tilde{\mu}_\bullet$, $\tilde{\sigma}_\bullet$: climatological properties of y, x

SAMOS



Convert y to y^* using $y^* = \frac{y - \tilde{\mu}_y}{\tilde{\sigma}_y}$.

SAMOS


Station A
Station B
standardized anomalies [°C]
2
1
0
−1
−2
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

SAMOS Summary

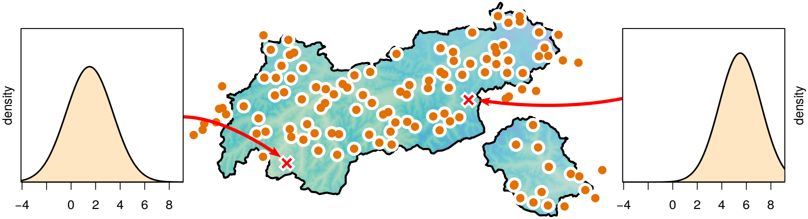
Spatio-Temporal Climatologies

- Account for ...
- **seasonal** and **diurnal** patterns,
- **spatial** differences (longitude, latitude, altitude),
- and possible interactions.

Standardized Anomalies

- **location** or **station** independent
- independent from **season** and **time**

SAMOS Summary



On the Anomaly Scale

- combine data from **all stations** and **lead times**
- estimate **one "simple"** model for the area of interest
- **correct** current ensemble forecast $\rightarrow \mu^*$, σ^*
- de-standardize:

$$\mu = \mu^* \cdot \tilde{\sigma_y} + \tilde{\mu}_y$$

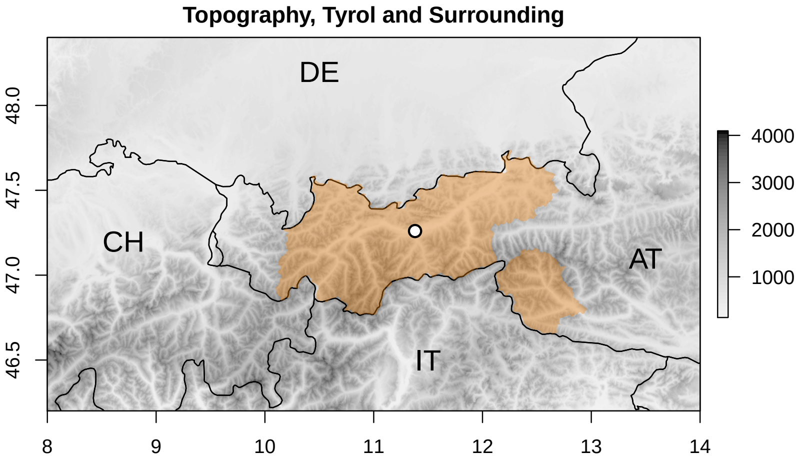
$$\sigma = \sigma^* \cdot \tilde{\sigma_y}$$

Hourly Probabilistic Snow Forecasts Over Complex Terrain

A Hybrid Ensemble Postprocessing Approach

R Stauffer, GJ Mayr, JW Messner, A Zeileis

Snow Forecasts



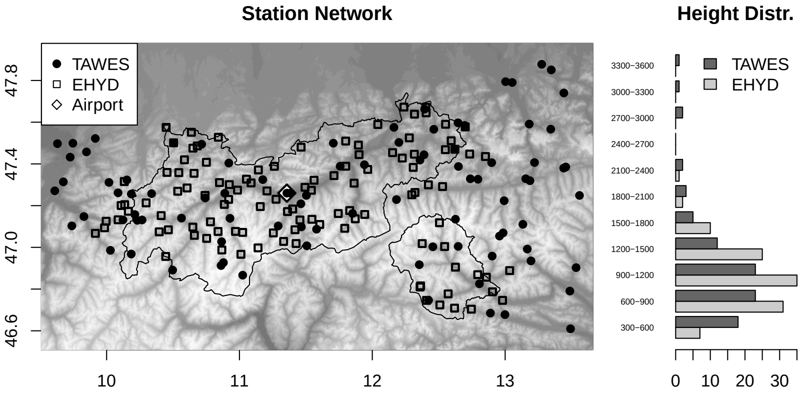
Problem: lack of reliable fresh snow observations.

Snow Forecasts

Idea: Hybrid Approach

- forecast temperature & precipitation instead
- SAMOS
  - hourly temperature forecasts
  - daily precipitation forecasts
- ensemble copula coupling
  - restore spatio-temporal structure
  - convert temperature & precipitation into snow

Snow Forecasts



Observation Data

- hourly 2 m air temperature: 90 stations, up to 10+ years
- daily precipitation sums: 110 stations, up to 40+ years

Snow Forecasts

Numerical Weather Forecast Data

ECMWF hindcast

- $10 + 1$ member ensemble
- 6-hourly output
- initialized twice a week (0000 UTC, 20 years)

ECMWF ensemble

- $50 + 1$ member ensemble
- hourly output
- initialized 0000 UTC

Snow Forecasts

SAMOS Model Assumptions

Temperature

$$T^* \sim \mathcal{N}(\mu^*, \sigma^*)$$
$$\mu^* = \beta_0 + \beta_1 \cdot \bar{x}^*$$
$$\log(\sigma^*) = \gamma_0 + \gamma_1 \cdot \log(x^*)$$

x^*: std. anomalies of the 2m temperature.

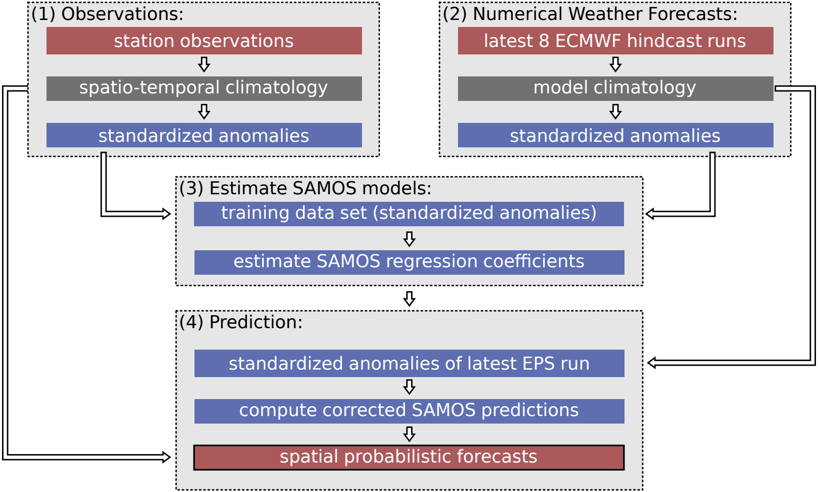
Precipitation

$$\text{precip}^{p*} \sim \mathcal{L}_{cens}(\mu^*, \sigma^*)$$
$$\mu^* = \beta_0 + \beta_1 \cdot \bar{x}^* \cdot (1 - z) + \beta_2 \cdot z$$
$$\log(\sigma^*) = \gamma_0 + \gamma_1 \cdot \log(s_x^*) \cdot (1 - z)$$

x^*: std. anomalies of power-transformed total precipitation.

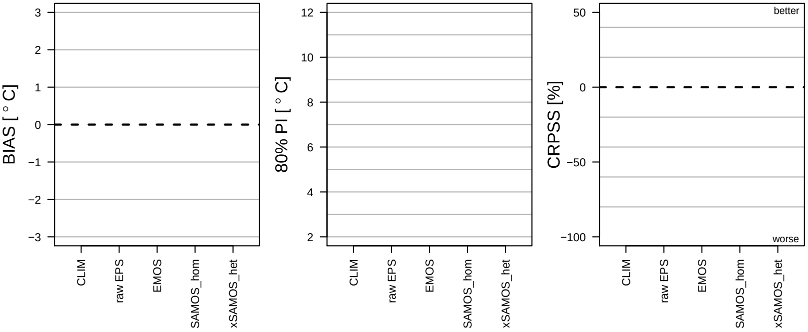
z: split-variable (binary) to handle unanimous predictions.

Snow Forecasts



Snow Forecasts

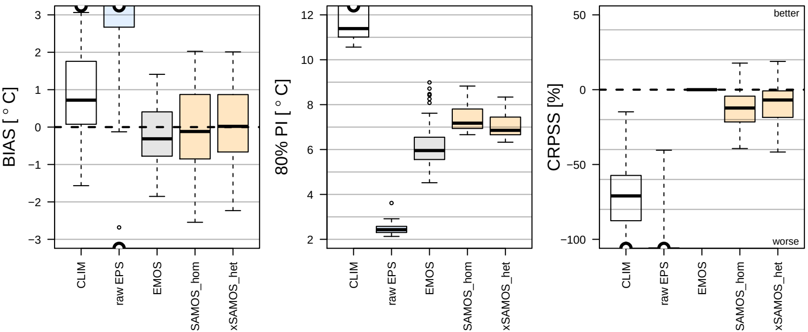
Temperature Forecasts



Verification: Dec 2016 - mid April 2017.

Snow Forecasts

Temperature Forecasts

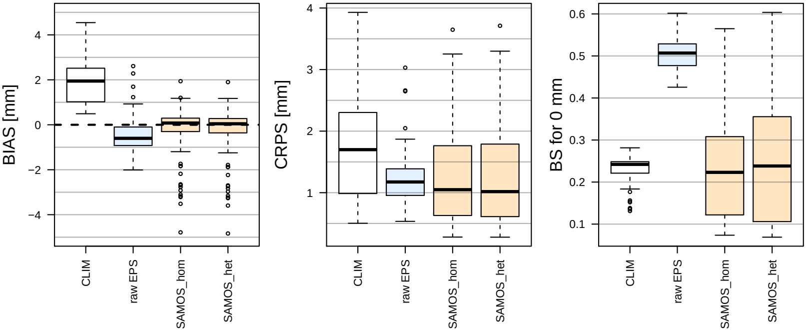


SAMOS

- Slightly outperformed by EMOS, ...
- but allows for spatial predictions.
- Ensemble spread: barely any additional information.

Snow Forecasts

24 h Precipitation Forecasts



SAMOS

- Outperforms raw EPS, less skillful than for temperature.
- Ensemble variance barely any additional information.

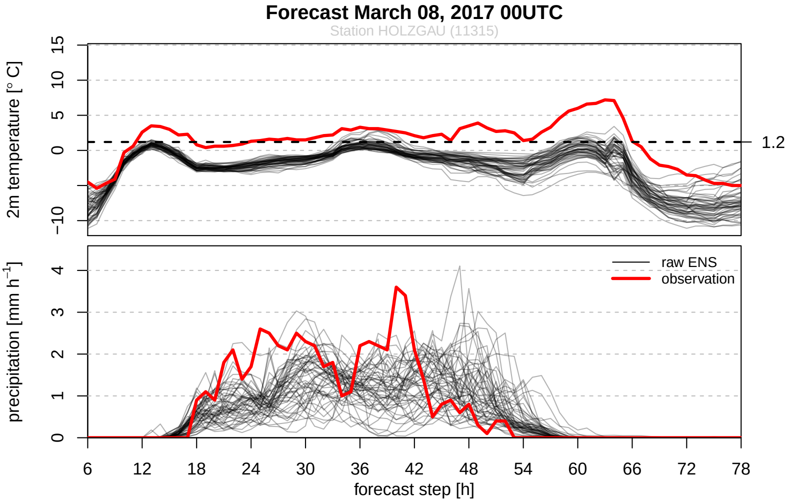
Snow Forecasts

How to Get Hourly Snow Predictions?

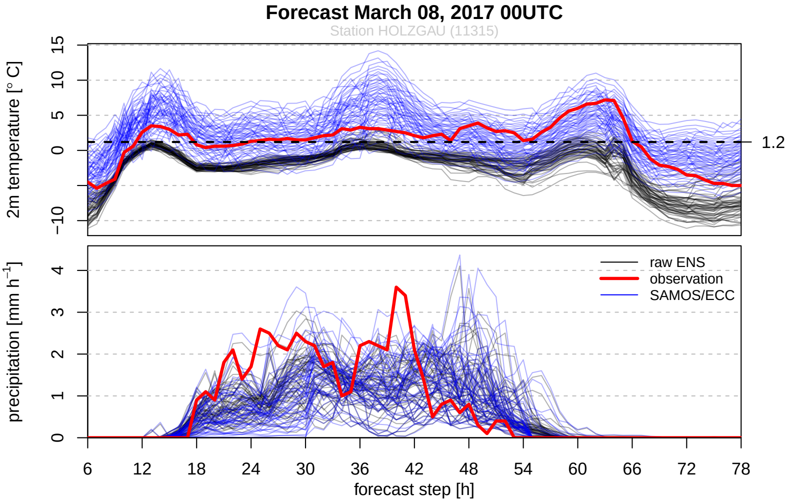
- combine temperature/precipitation using ECC
- temperature:
  - draw 51 member ensemble from corrected hourly $\mathcal{N}$
  - restore rank order structure
- 24 h precipitation:
  - draw 51 member ensemble from corrected $\mathcal{L}_0$
  - restore rank order structure
- hourly precipitation:
  - re-weight raw hourly ensemble forecasts using:

    $\omega_{ms} = \frac{\hat{tp}_{copula,ms}}{tp_{EPS,ms}}$

Snow Forecasts



Snow Forecasts



Snow Forecasts

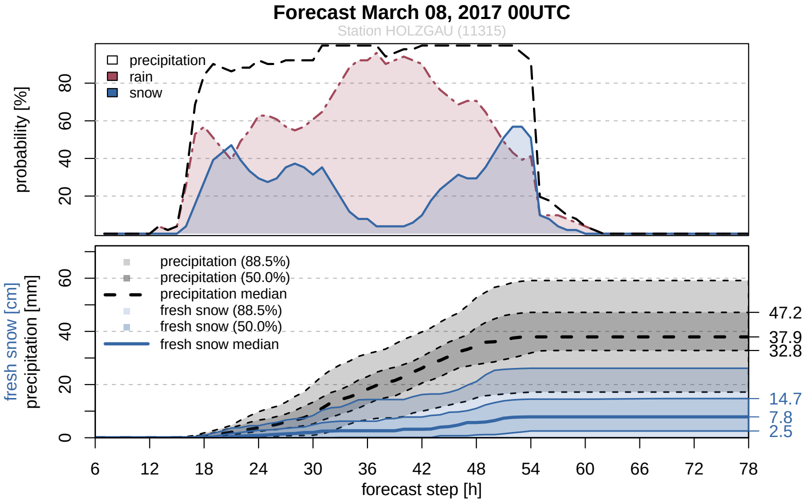
Hourly Snow Forecasts

$$
PI_{ms} = \begin{cases} \text{“dry” if:} \\ \qquad \text{precipitation}_{ms} \leq 0.05 \, \frac{mm}{h} \\ \\ \text{“rain” if:} \\ \qquad \text{precipitation}_{ms} > 0.05 \, \frac{mm}{h} \wedge T_{2m,ms} > 1.2^\circ C \\ \text{“snow” if:} \\ \qquad \text{precipitation}_{ms} > 0.05 \, \frac{mm}{h} \wedge T_{2m,ms} \leq 1.2^\circ C \end{cases}
$$

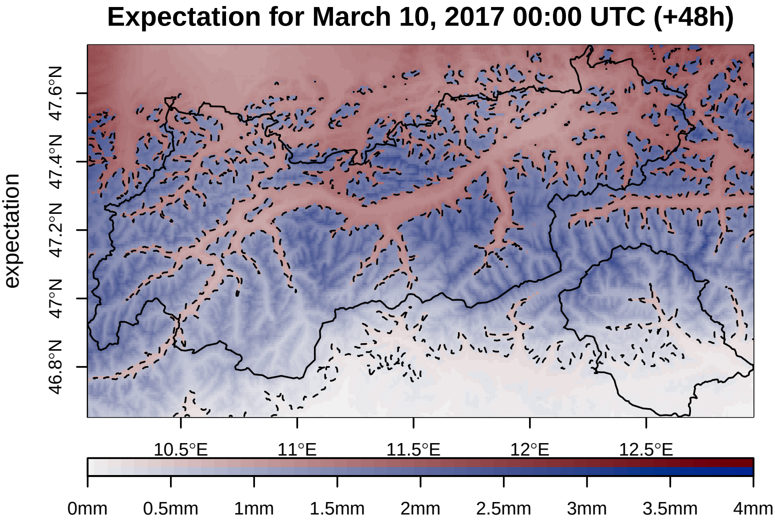
PI**:** precipitation type indicator {dry, rain, snow}

m/s**:** copula member and forecast step

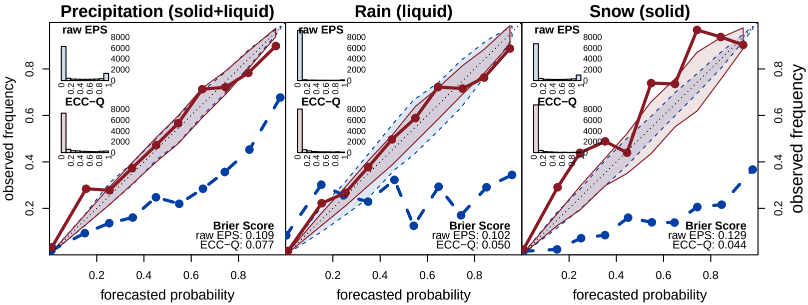
Snow Forecasts



Snow Forecasts



Snow Forecasts



Empirical frequencies:
precipitation 15.8 %, rain 9.8 %, snow 7.5 %.

Snow Forecasts Summary

- **ensemble spread**: barely any information
- **large spread** for postprocessed **temperature**
- **improvements**: more pronounced for **temperature**
- **combine** different **data sources**
- **use data sets** with different **temporal** resolution
- **reliable** hourly probabilistic spatial snow forecasts

Sub-Seasonal Climate Forecast Rodeo

Improve Existing Sub-Seasonal Forecasts, Bureau of Reclamation

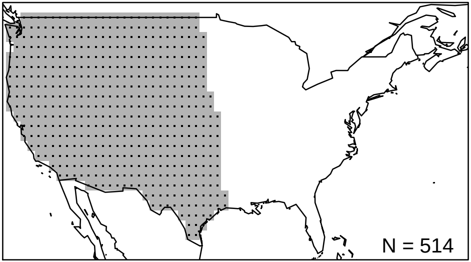

Source: www.retro-housewife.com

The Rodeo

The Challenge

- predict anomalies
- mean temperature and accumulated precipitation
- week 3-4 and 5-6

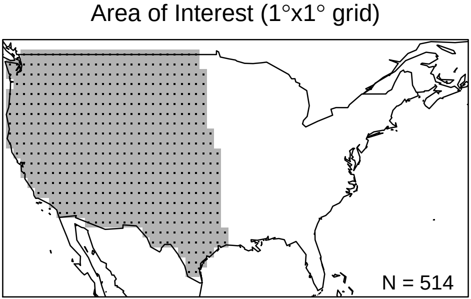
Area of Interest (1°x1° grid)



The Rodeo

The Price

- 800 000 USD in total
- iff. anomaly forecasts are significantly better than:
  - a damped persistence
  - the CFSv2 itself



The Rodeo

The Truth

- Climate Prediction Center's gridded data set
- gridded gauge data set
- gridded temperature data set
- climatology: 2-week mean 1981–2010

The Forecasts

- 4 member CFSv2
- mean over 8 runs ($8 \times 4 = 32$ member mean)

The Measure

$$ACC = \frac{\sum \left(f' \cdot o'\right)}{\sqrt{\sum o'^2 \cdot \sum f'^2}}$$

The Rodeo

The Idea

- CPC: provides observation climatology ($\tilde{\mu}_y$, $\tilde{\sigma}_y$)
- CVSv2 reforecasts: provide ensemble climatology ($\tilde{\mu}_x$, $\tilde{\sigma}_x$)
- apply “complex” homoscedastic AMOS/SAMOS

Model Assumption

$$y^* \sim \mathcal{D}\big(\mu^*, \sigma^*\big)$$

Where μ^* may include:

- forecasted anomalies (2 m temperature, dewpoint, ...)
- spatial effects $f(\text{long}, \text{lat})$
- teleconnection indices (NAO, NA, PNA; CPC)
- snow cover data (NSIDC)

The Rodeo

ECMWF Based Approach

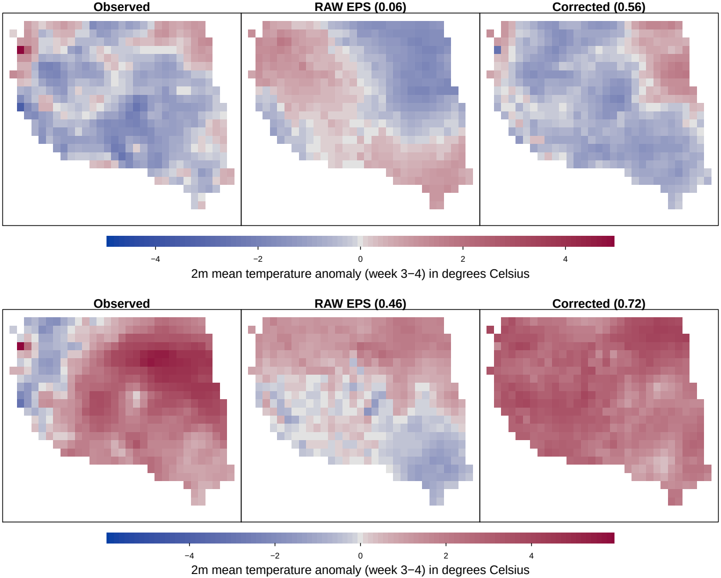
- temperature (week 3-4)
- Gaussian AMOS
- 70 covariates
- optimization:

  based on *R* package `bamlss`

  likelihood-based gradient boosting

  variable selection & parameter estimation

The Rodeo


Observed
RAW EPS (0.06)
Corrected (0.56)
−4
−2
0
2
4
2m mean temperature anomaly (week 3−4) in degrees Celsius
Observed
RAW EPS (0.46)
Corrected (0.72)
−4
−2
0
2
4
2m mean temperature anomaly (week 3−4) in degrees Celsius

The Rodeo Leader Board

Weeks 3&4 Temperature

Team	Newest Score	Average Score
bgzimmerman	-0.0994	0.2855
prxwx	-0.1821	0.2265
StillLearning	0.029	0.217
DampedPersistence	**-0.0794**	**0.1952**
CFSv2	**-0.3997**	**0.1589**
asanteko2000	-0.1117	0.0909
lupoa13	-0.2187	0.0895
Salient	0.05	-0.1365

Weeks 5&6 Temperature

Team	Newest Score	Average Score
bgzimmerman	-0.4472	0.2357
CFSv2	**0.5267**	**0.2192**
StillLearning	0.1436	0.2044
prxwx	0.3105	0.2026
lupoa13	-0.5854	0.1675
asanteko2000	-0.1046	0.0897
DampedPersistence	**0.1084**	**-0.0762**
Salient	-0.8229	-0.09

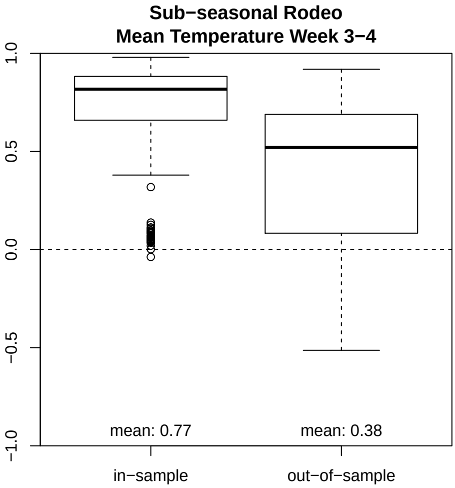
Weeks 3&4 Precipitation

Team	Newest Score	Average Score
Salient	0.7758	0.2144
prxwx	0.0921	0.1711
lupoa13	-0.1367	0.1246
CFSv2	**0.1837**	**0.0713**
StillLearning	0.7987	0.0227
bgzimmerman	0.1087	-0.0221
asanteko2000	-0.7981	-0.0612
DampedPersistence	**-0.7996**	**-0.1463**

Weeks 5&6 Precipitation

Team	Newest Score	Average Score
Salient	0.5897	0.2162
prxwx	0.0995	0.1208
StillLearning	0.5816	0.0941
lupoa13	0.0916	0.0931
bgzimmerman	0.303	0.0773
CFSv2	**0.0692**	**0.0227**
asanteko2000	-0.5561	-0.0879
DampedPersistence	**-0.4375**	**-0.1613**

The Rodeo



Sub−seasonal Rodeo
Mean Temperature Week 3−4
1.0
0.5
0.0
−0.5
−1.0
mean: 0.77
mean: 0.38
in−sample
out−of−sample

References I

Gneiting, T, AE Raftery, AH Westveld, and T Goldman, 2005: Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *Monthly Weather Review*, **133**, 1098–1118, doi:10.1175/MWR2904.1.

Dabernig, M, GJ Mayr, JW Messner, and A Zeileis, 2017: Spatial Ensemble Post-Processing with Standardized Anomalies. *Quarterly Journal of the Royal Meteorological Society*, **143**, 909–916, doi:10.1002/qj.2975.

Stauffer, R, N Umlauf, JW Messner, GJ Mayr, and A Zeileis, 2017: Ensemble Postprocessing of Daily Precipitation Sums over Complex Terrain Using Censored High-Resolution Standardized Anomalies. *Monthly Weather Review*, **145**, 955–969, 10.1175/MWR-D-16-0260.1.

Dabernig, M, GJ Mayr, JW Messner, and A Zeileis, 2017: Simultaneous Ensemble Postprocessing for Multiple Lead Times with Standardized Anomalies. *Monthly Weather Review*, **145**, 2523–2531, 10.1175/MWR-D-16-0413.1.

References II

Gebetsberger, M, JW Messner, GJ Mayr, and A Zeileis, 2016: Tricks for Improving Non-Homogeneous Regression for Probabilistic Precipitation Forecasts: Perfect Predictions, Heavy Tails, and Link Functions. Working Papers, Faculty of Economics and Statistics, University of Innsbruck.

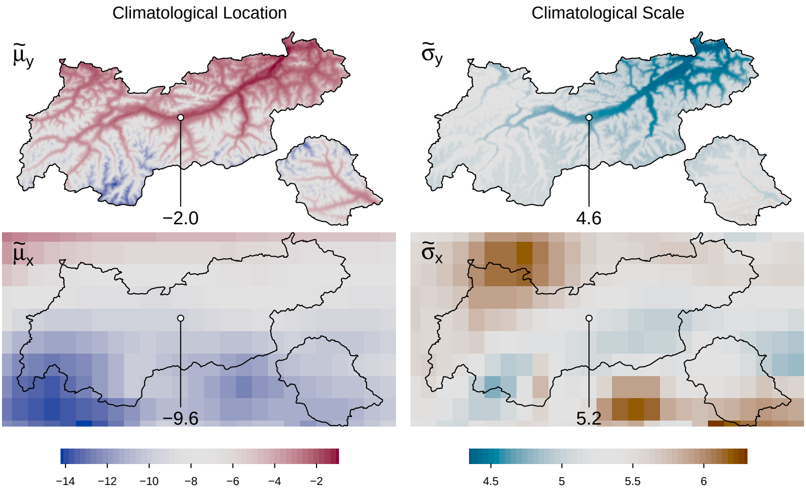
Stauffer R, GJ Mayr, JW Messner, and A Zeileis, 2018: Hourly Probabilistic Snow Forecasts over Complex Terrain: A Hybrid Ensemble Postprocessing Approach. Working papers, Faculty of Economics and Statistics, University of Innsbruck.

Thank you for your attention

and special thanks to Tom for the invitation!

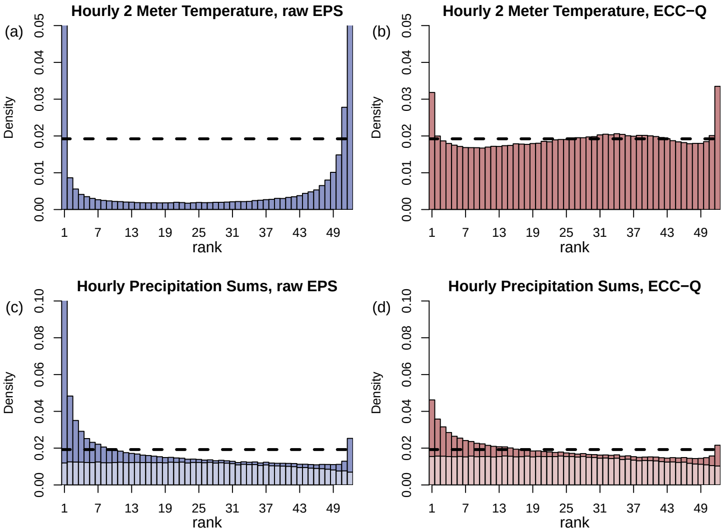
Appendix

SAMOS



2m air temperature, January 1, 0000 UTC (+24h forecast). Unit: °Celsius

Snow Forecast: Hourly Calibration



Snow Forecast: Hourly Verification