universität innsbruck

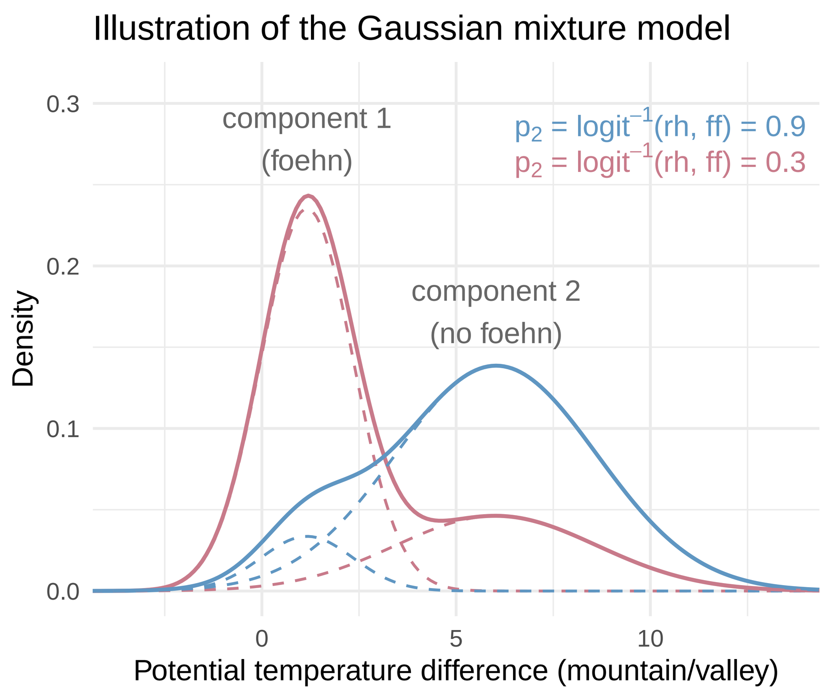
Long-term foehn reconstruction combining unsupervised and supervised learning

Reto Stauffer (*Reto.Stauffer@uibk.ac.at*), Georg J. Mayr and Achim Zeileis

Foehn "observations" — Last 14–22 years

- Foehn is a wind on the leeward side of a mountain range
- Characterized by a sharp increase in wind speed and changes in temperature and relative humidity
- Strong effect on local climate and can cause severe damage
- **Cannot be measured directly**
- **Classification based on mixture model with concomitants**

Classification: Two-component mixture model with concomitants to classify 'foehn' and 'no foehn' (Stauffer 2023).



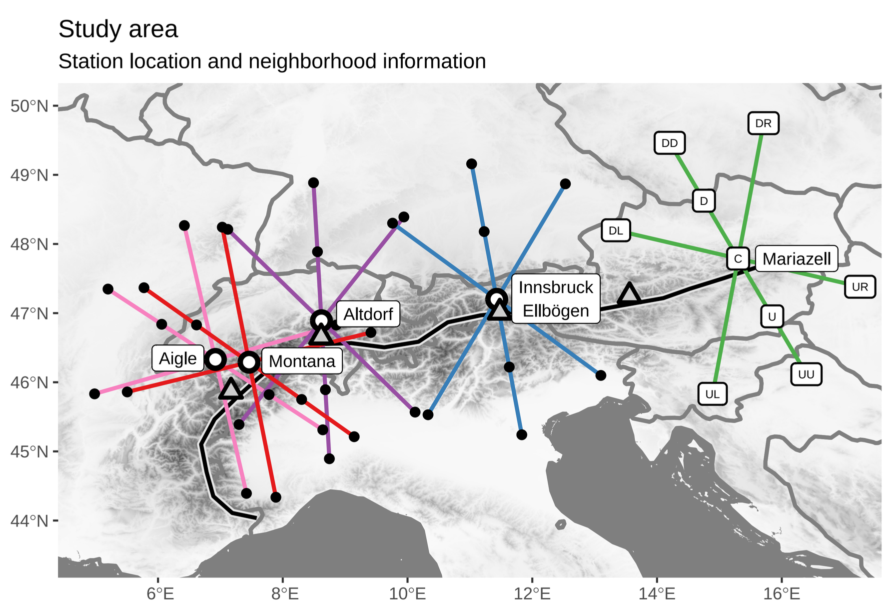
Mixture model with:

- Two Gaussian components
- Potential temperature differences to separate the main components
- Concomitant model: logit model using relative humidity (rh) and wind speed (ff)
- Estimate based on 14–22 years of data, 10 min resolution (Fig 3)

Figure 1: Illustrative example of the combined density for two different probabilities $p_2 = \{0.9, 0.3\}$ as returned by the concomitant model.

ERA5 reanalysis (Hersbach et al. 2023) — 1940–2022

- Physical global atmospheric reanalysis data, hourly resolution
- **Provides detailed long-term weather characterisation, but no direct information about foehn**



Examples of ERA5 variables
C.t500
temperature 500 hPa at 'C'
C.tpsum6h
6 h precipitation sum at 'C'
C.lt700_800
level thickness 700-800 hPa at 'C'
diffmsl_m3hp0h_ULDR
difference mean sea level pressure change 'UL'/'DR' over past 3 hours
. . . ca. 500 variables in total

Figure 2: Location of the six stations in the European Alps (left). Triangle: Mountain stations for classification. 'Star': Station location (C; center) and additional locations upstream/downstream (U*/D*) used to calculate derived ERA5 variables (right).

Supervised learning

- Based on 14–22 years with foehn classification ('yes'/'no')
- Estimate binary response model with ERA5 data as covariates
- **Allows to reconstruct foehn occurrence from 1940–2022 on an hourly scale**

Binary classifier: Here, a logit model with lasso regularization (Friedman et al. 2023) is employed using ≈500 different covariates from ERA5.

$$\Pr(\text{foehn} = \text{yes}) = f(ERA5)$$

Reconstruction: Obtain fitted foehn probabilities based on the estimated classifier using

$$\hat{\pi}_{1h} = \hat{f}(ERA5).$$

As an alternative to logistic regression, any supervised learner for a binary response could be used.

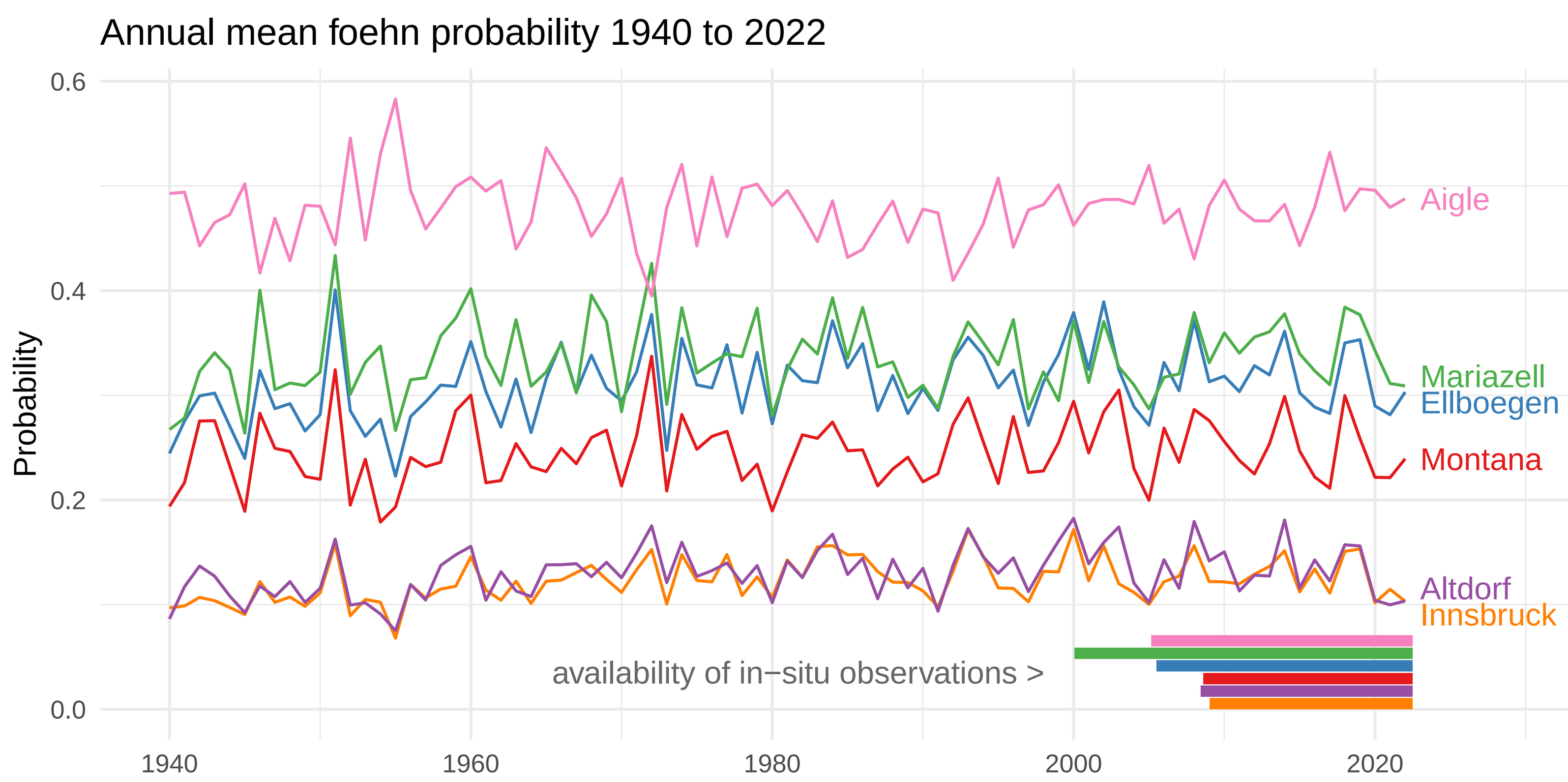


Figure 3: Annual mean of daily maxima for all six stations (see Fig 2).

References:

Dokumentov A and Hyndman RJ (2022). *stR: STR Decomposition.* R package version 0.5, https://cran.r-project.org/package=stR.

Friedman J et al. (2023). *glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models*, R package version 4.1-7, https://cran.r-project.org/package=glmnet.

Hersbach H, et al. (2023). "ERA5 Hourly data on Pressure Levels/Single Levels from 1940 to present". *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, (Accessed on 12-07-2023).

Stauffer R et al. (2023). *foehnix: Objective Foehn Diagnosis.* R package version 0.1.6, https://github.com/retostauffer/Rfoehnix.

Acknowledgements:

The computational results presented here have been achieved (in part) using the LEO HPC infrastructure of the University of Innsbruck.

Poster

Data (43MB)

Final result

- High-resolution (hourly) long-term time series of foehn probabilities (≈720 000 hourly probabilities per station)
- Serves as input for additional applications

Example application

- Season-trend analysis on monthly basis for Ellbögen
- Identify changes/trends over the past eight decades

Season-trend decomposition: Following Dokumentov et al. (2022), a model of the following form is employed.

$$\hat{\pi}_{1m,t} = T_t + S_t + R_t$$

where $\hat{\pi}_{1m,t}$ is the monthly mean of daily maxima for time t, T_t a smooth trend, S_t a smoothly changing seasonality, and R_t the remainder. The results (Fig 4) show a slight upward trend over the second half of the last century, but a stable seasonal pattern.

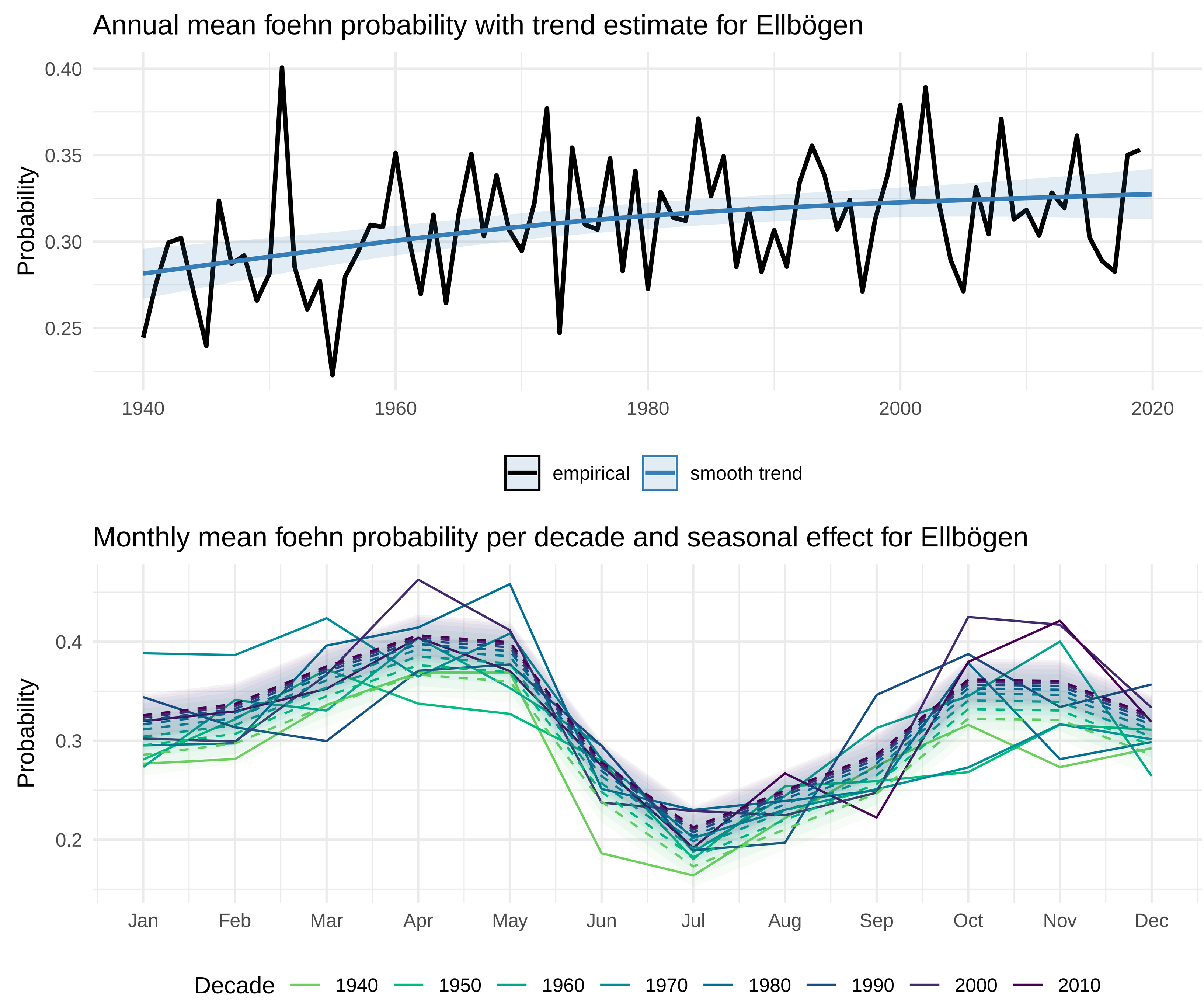


Figure 4: Season-trend decomposition for Ellbögen based on monthly means of daily maxima showing the temporal trend (top) as well as the underlying seasonality (bottom) for the most recent eight decades.