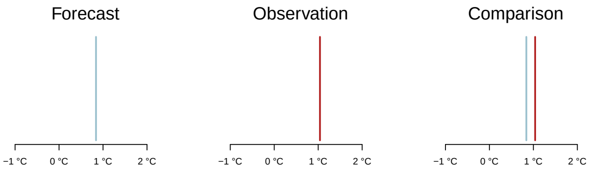
universität innsbruck

Remember the past: A comparison of time-adaptive training schemes for non-homogeneous regression

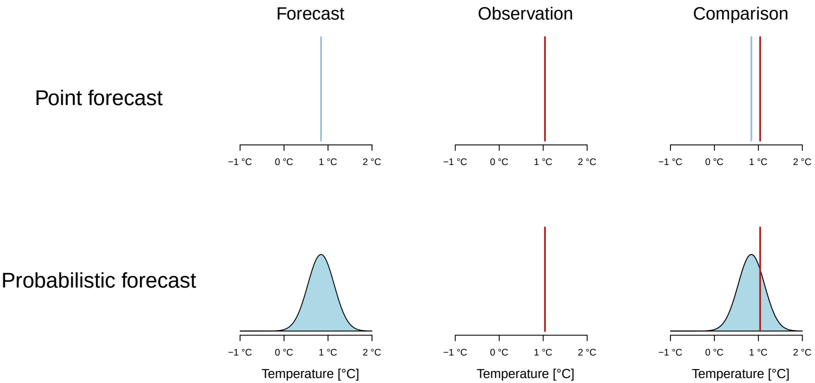
Moritz N. Lang, Sebastian Lerch, Georg J. Mayr, Thorsten Simon, Reto Stauffer, and Achim Zeileis

Probabilistic forecasting

Point forecast

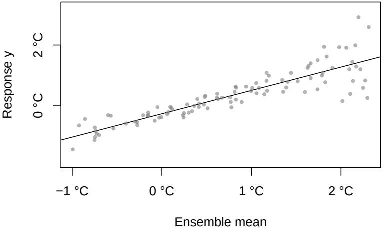


Probabilistic forecasting



- Probabilistic forecasts quantify the uncertainty of the predictions.
- Aims: Calibration and sharpness.

Non-homogeneous or distributional regression



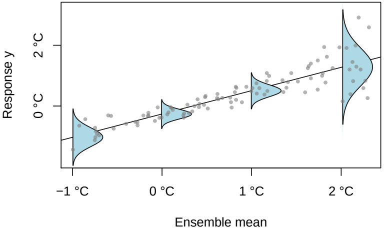
- Response y is assumed to follow a predefined parametric distribution, e.g., normal distribution.
- The two distribution parameters location and scale are expressed by a **linear function of covariates** $x_1 \dots x_{(k+l)}$:

$$y \sim \mathcal{N}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot x_1 + \dots + \beta_k \cdot x_k$$

$$\log(\sigma) = \gamma_0 + \gamma_1 \cdot x_{(k+1)} + \dots + \gamma_l \cdot x_{(k+l)}$$

Non-homogeneous or distributional regression



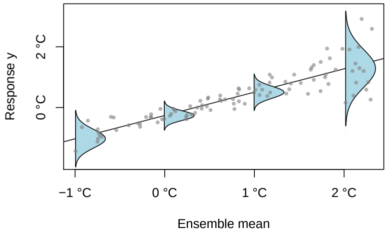
- Response y is assumed to follow a predefined parametric distribution, e.g., normal distribution.
- The two distribution parameters location and scale are expressed by a **linear function of covariates** $x_1 \dots x_{(k+l)}$:

$$y \sim \mathcal{N}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot x_1 + \dots + \beta_k \cdot x_k$$

$$\log(\sigma) = \gamma_0 + \gamma_1 \cdot x_{(k+1)} + \dots + \gamma_l \cdot x_{(k+l)}$$

Non-homogeneous or distributional regression



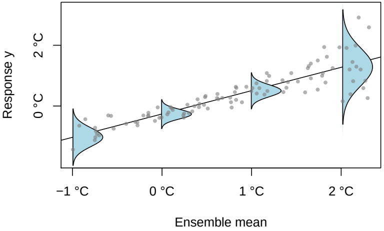
- Response y is assumed to follow a predefined parametric distribution, e.g., normal distribution.
- The two distribution parameters location and scale are expressed by the **ensemble mean** *m* and **ensemble standard deviation** *s* of the corresponding NWP of the response:

$$y \sim \mathcal{N}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot m$$

$$\log(\sigma) = \gamma_0 + \gamma_1 \cdot s$$

Non-homogeneous or distributional regression



- Response y is assumed to follow a predefined parametric distribution, e.g., normal distribution.
- The two distribution parameters location and scale are expressed by the **ensemble mean** *m* and **ensemble standard deviation** *s* of the corresponding NWP of the response:

$$y \sim \mathcal{N}(\mu, \sigma)$$

$$\mu = \beta_0 + \beta_1 \cdot m$$

$$\log(\sigma) = \gamma_0 + \gamma_1 \cdot s$$

$\Rightarrow$ **How to account for seasonally varying regression coefficients?**

Time adaptive training schemes

To adjust for seasonally varying error characteristics between covariates (ensemble forecasts) and response:

Time adaptive training schemes

To adjust for seasonally varying error characteristics between covariates (ensemble forecasts) and response:

- **Sliding-window** uses previous *n* days for model training.

Time adaptive training schemes

To adjust for seasonally varying error characteristics between covariates (ensemble forecasts) and response:

- **Sliding-window** uses previous *n* days for model training.
- **Regularized sliding-window** adds regularization in model estimation.

Time adaptive training schemes

To adjust for seasonally varying error characteristics between covariates (ensemble forecasts) and response:

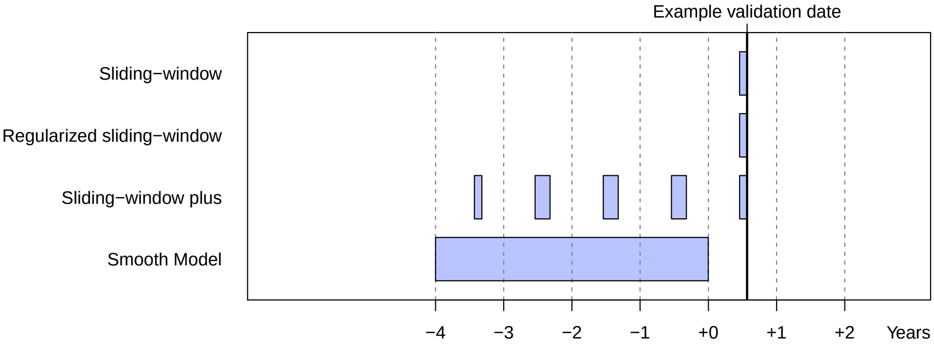
- **Sliding-window** uses previous n days for model training.
- **Regularized sliding-window** adds regularization in model estimation.
- **Sliding-window plus** uses $2n$ days centered around the same calendar day over all previous years for model training.

Time adaptive training schemes

To adjust for seasonally varying error characteristics between covariates (ensemble forecasts) and response:

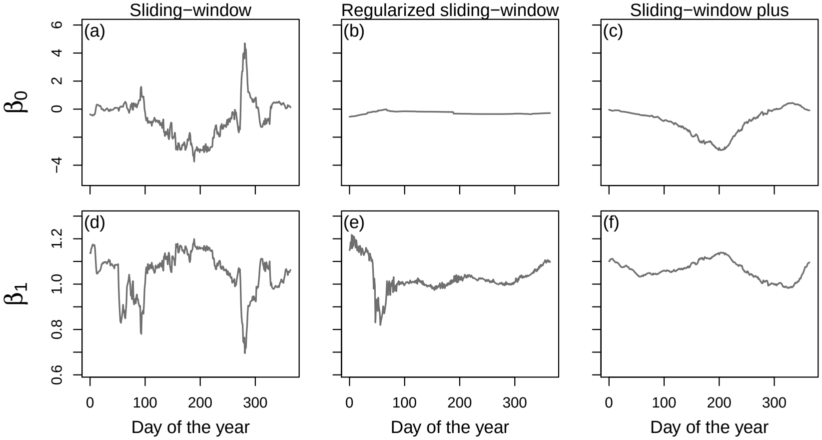
- **Sliding-window** uses previous n days for model training.
- **Regularized sliding-window** adds regularization in model estimation.
- **Sliding-window plus** uses $2n$ days centered around the same calendar day over all previous years for model training.
- **Smooth model** uses all available days for model training by allowing coefficients to smoothly evolve over the year.

Time adaptive training schemes



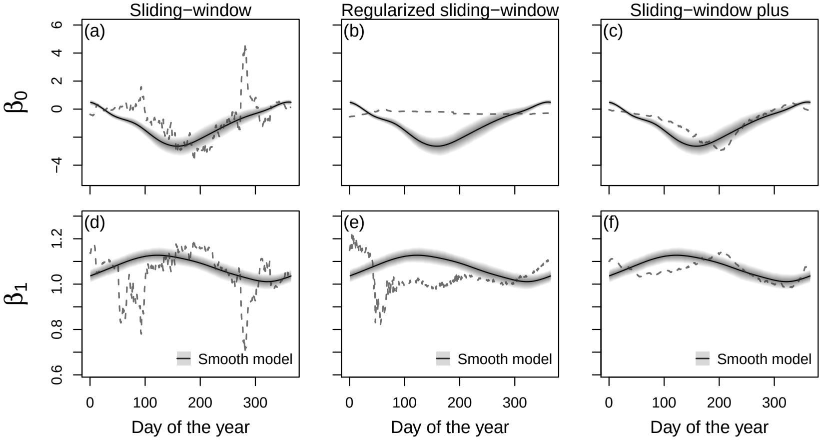
Surface temperature forecasting

Location parameter



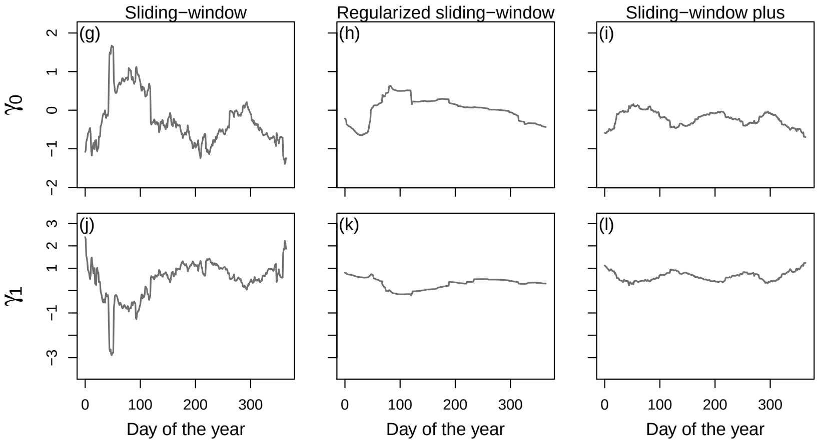
Surface temperature forecasting

Location parameter



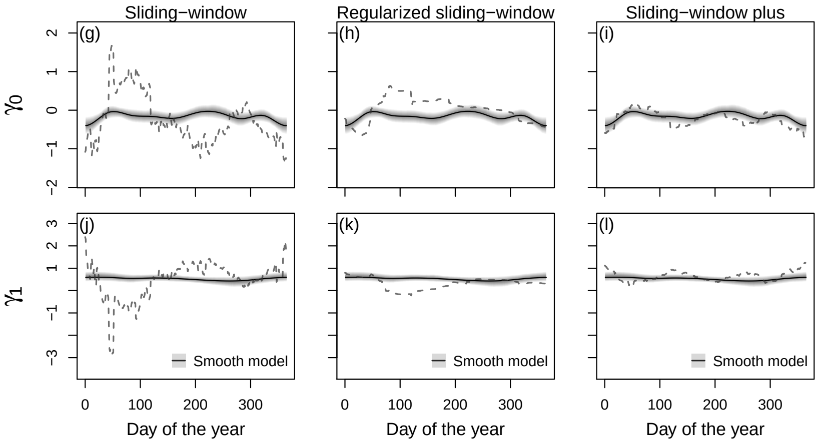
Surface temperature forecasting

Scale parameter



Surface temperature forecasting

Scale parameter



Surface temperature forecasting

Validation setup

Response:

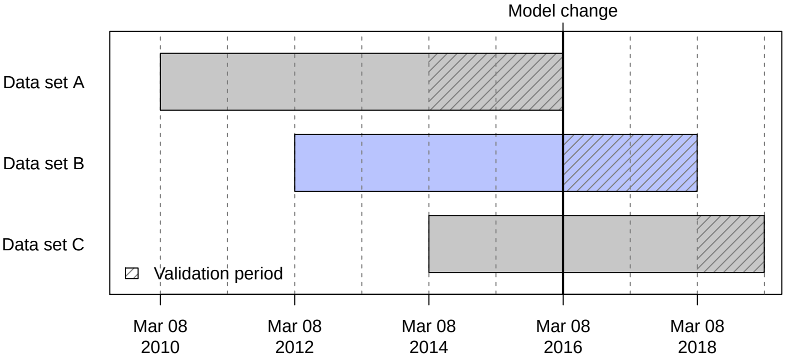
- 2 m temperature forecasts at five weather stations located in the plain of northern Germany.

Covariates:

- Ensemble mean m and the ensemble standard deviation s of bilinearly interpolated 2 m temperature forecasts issued by the ECMWF.
- Considered forecast steps from $+12$ to $+72$ h, at a 12-hourly temporal resolution (00:00 UTC run).

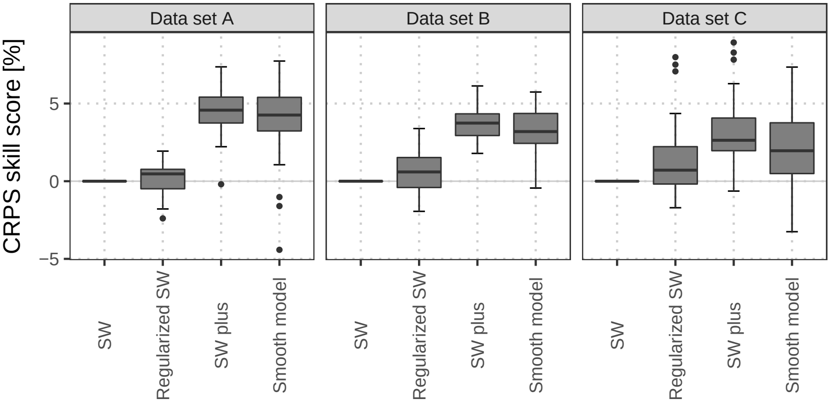
Surface temperature forecasting

Validation setup



Surface temperature forecasting

Validation results



Summary

- “Remembering the past” from multiple years of training data stabilizes and improves the calibration.

Summary

- “Remembering the past” from multiple years of training data stabilizes and improves the calibration.
- In case of certain ensemble model changes, using multiple years of training data is still superior to the classical sliding-window approach.

Summary

- "Remembering the past" from multiple years of training data stabilizes and improves the calibration.
- In case of certain ensemble model changes, using multiple years of training data is still superior to the classical sliding-window approach.
- Reducing the variance of the regression estimates appears to be more important than adjusting rapidly for changing forecast biases.

References

Gneiting, T., Raftery, A. E., Westveld III, A. H., and Goldman, T.: Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation, *Mon. Weather Rev.*, **133**, 1098 – 1118, doi:10.1175/MWR2904.1, 2005.

Möller, A., Spazzini, L., Kraus, D., Nagler, T., and Czado, C.: Vine copula based post-processing of ensemble forecasts for temperature, arXiv 1811.02255, *arXiv.org E-Print Archive*, in review, 2018.

Scheuerer, M.: Probabilistic quantitative precipitation forecasting using ensemble model output statistics, *Q. J. Roy. Meteor. Soc.*, **140**, 1086 – 1096, doi:10.1002/qj.2183, 2014.

Vogel, P., Knippertz, P., Fink, A. H., Schlueter, A., and Gneiting, T.: Skill of global raw and postprocessed ensemble predictions of rainfall over northern tropical Africa, *Weather Forecast.*, **33**, 369 – 388, doi:10.1175/waf-d-17-0127.1, 2018.

Lang, M. N., Lerch, S., Mayr, G. J., Simon, T., Stauffer, R., and Zeileis, A.:
Remember the past: A comparison of time-adaptive training schemes for non-homogeneous regression, *Nonlin. Processes Geophys.*, **27**, 23 – 34, doi:10.5194/npg-27-23-2020, 2020.

moritz.lang@uibk.ac.at MoritzNLang