What is the optimal length of the calibration period?

Ilja van Meerveld, Marc Vis, Yuko Asano, and Jan Seibert

ilja.vanmeerveld@geo.uzh.ch

Universität Zürich[UZH]

The University of Tokyo Forests,
Graduate School of Agricultural and Life Sciences,
The University of Tokyo

What calibration length to choose for your model study?

As long as possible  **Not longer than necessary**

- Include as much of the climatic variation as possible
  - Dry and wet periods
  - Extreme events

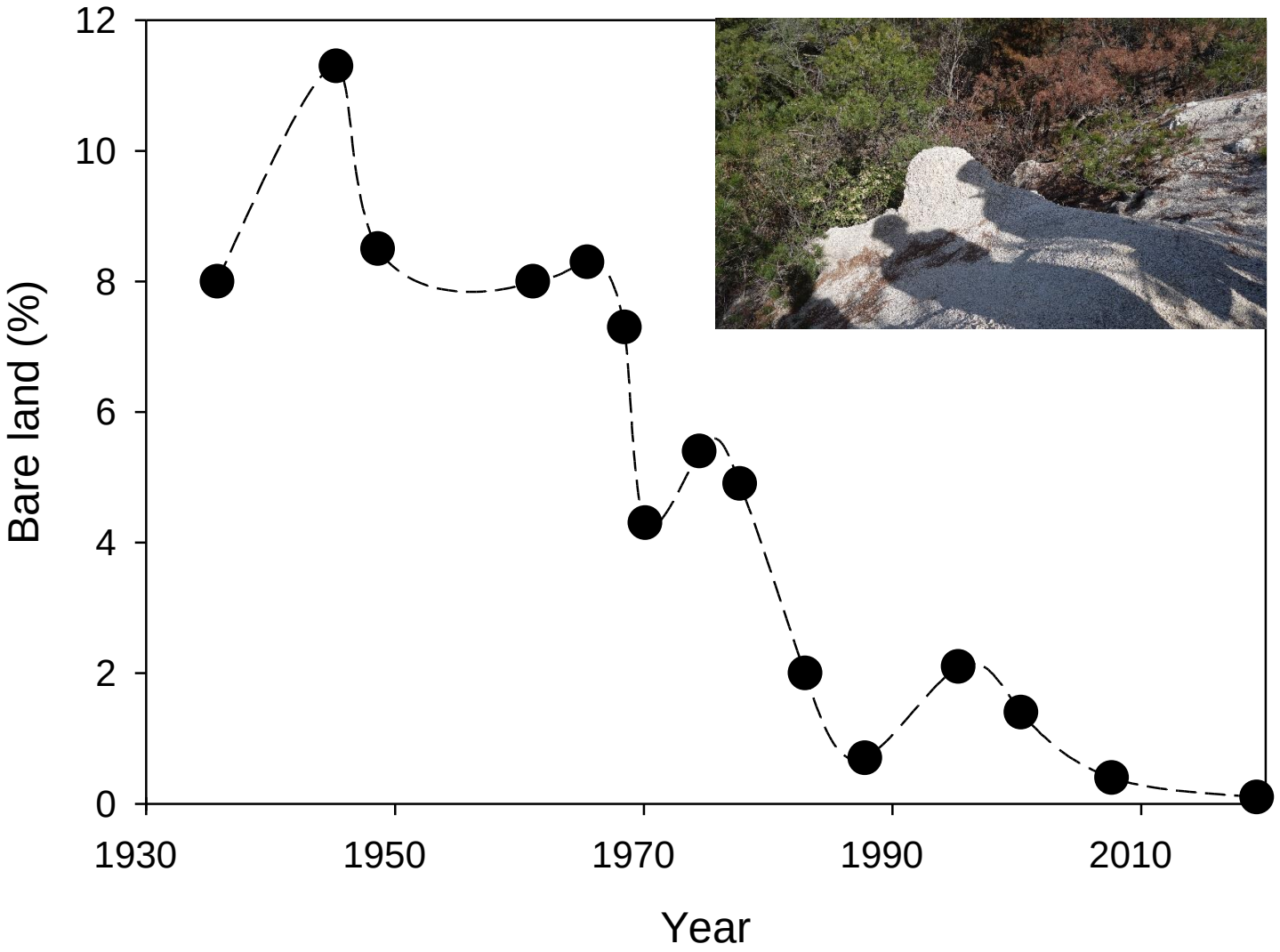
- No additional information after a few years of data
- Catchments change and past conditions may not reflect the current conditions

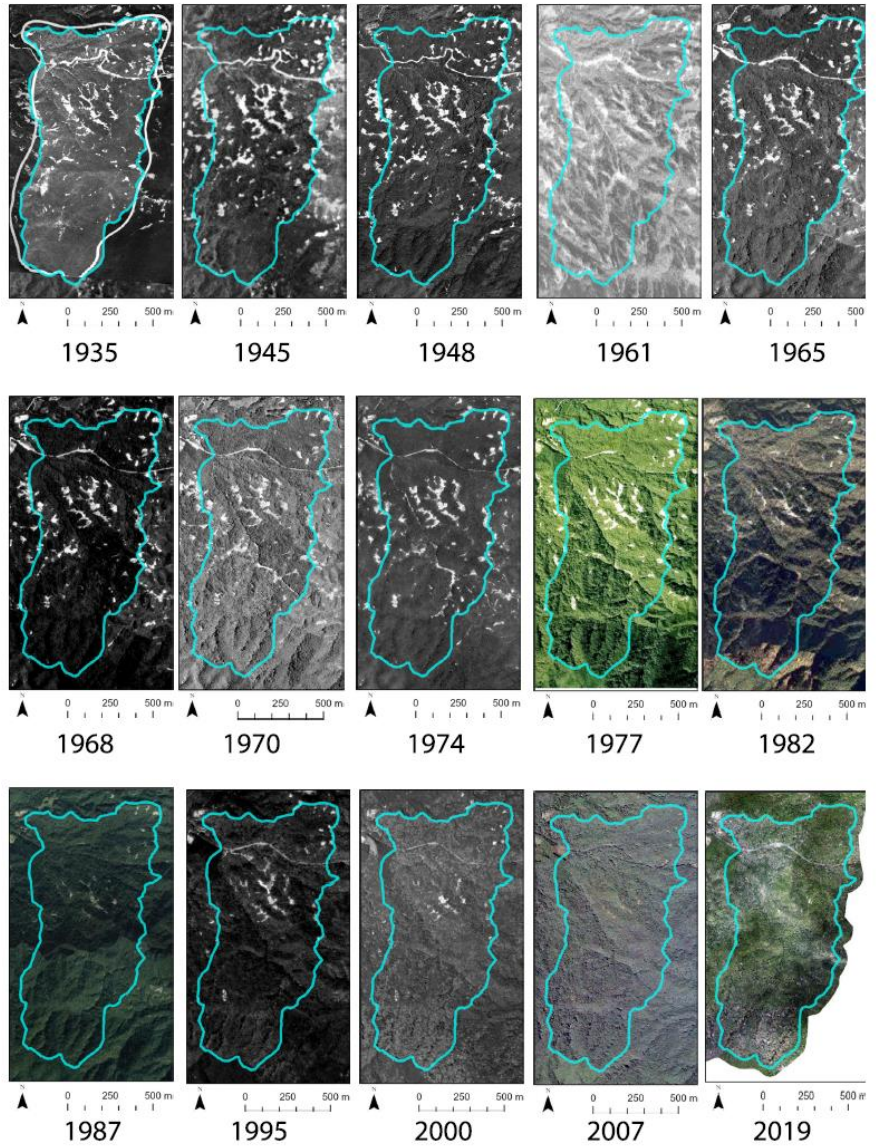
Check for catchment with long data series

- Example Shirasaka catchment (Aichi prefecture, Japan)
- Data since 1930

- 0.9 km^2
- 1860 mm/y (snow: 1%)
- Cambisols on granite

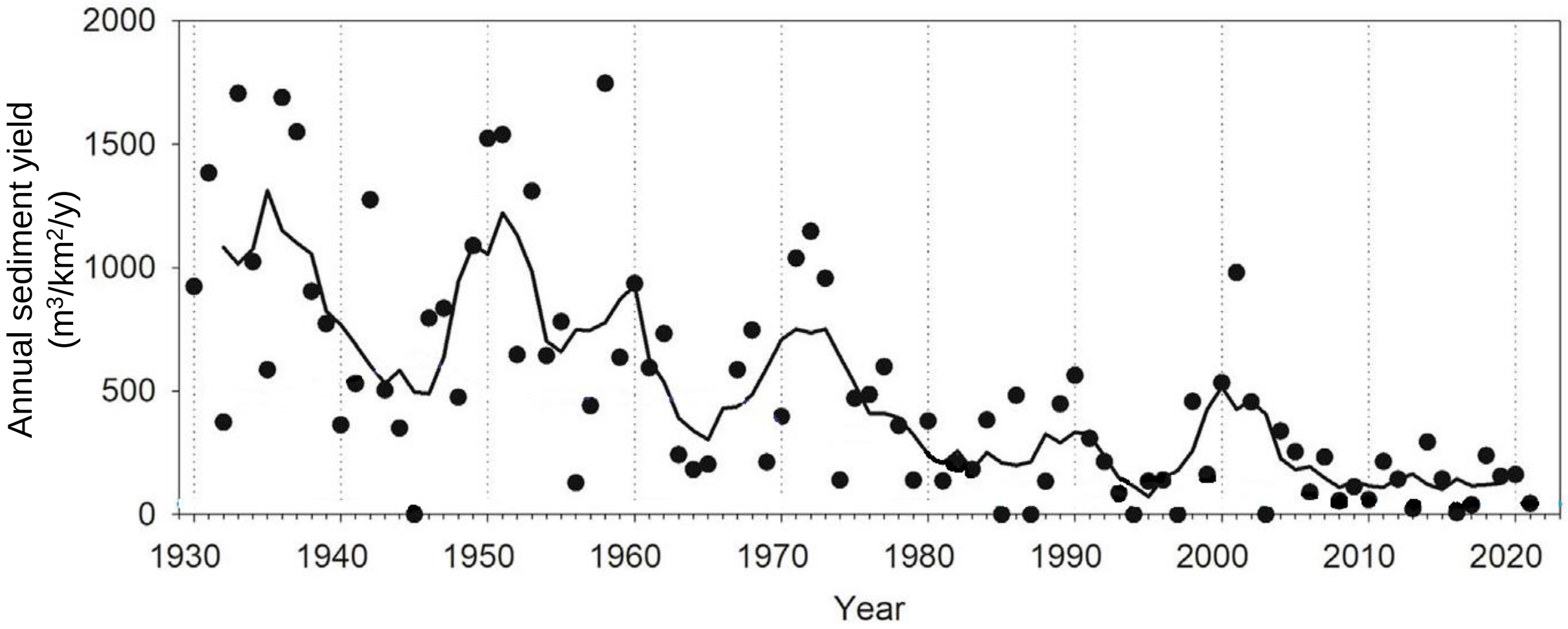
Change in degraded bare land area





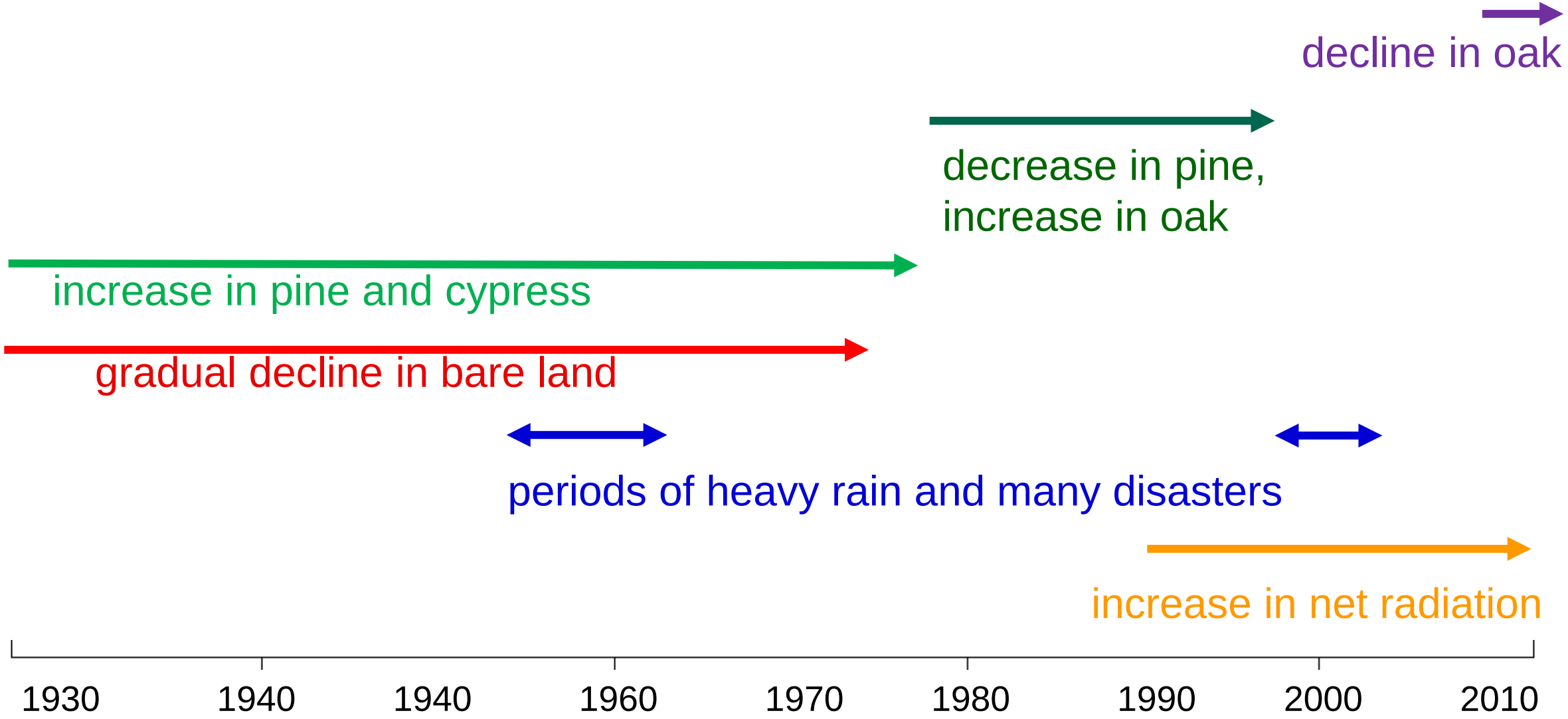
From: Asano et al, ESPL, 2025, https://doi.org/10.1002/esp.6071

Change in annual sediment yield



From: Asano et al, ESPL, 2025, https://doi.org/10.1002/esp.6071

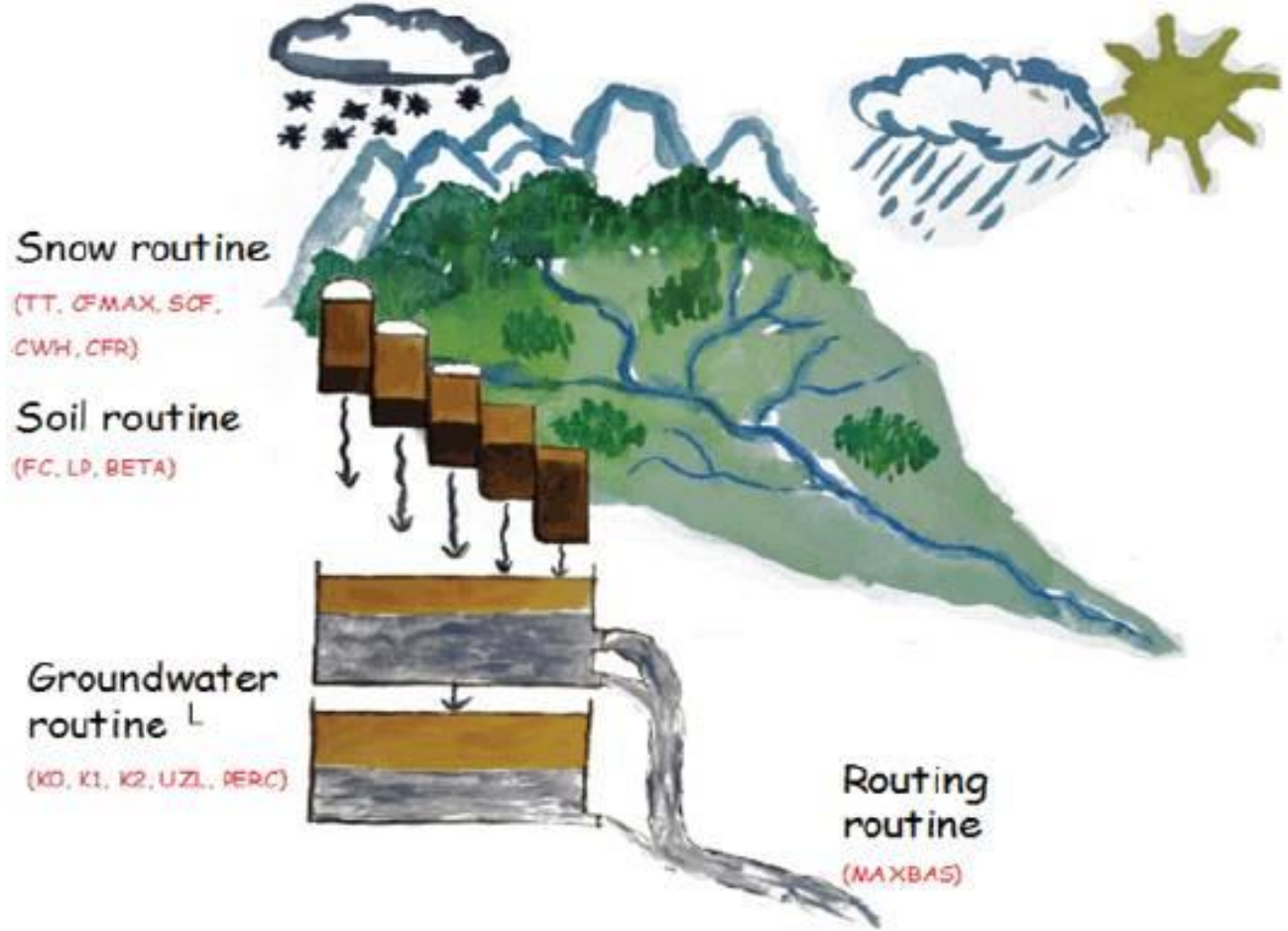
But also other changes



What calibration period to use?

HBV Model

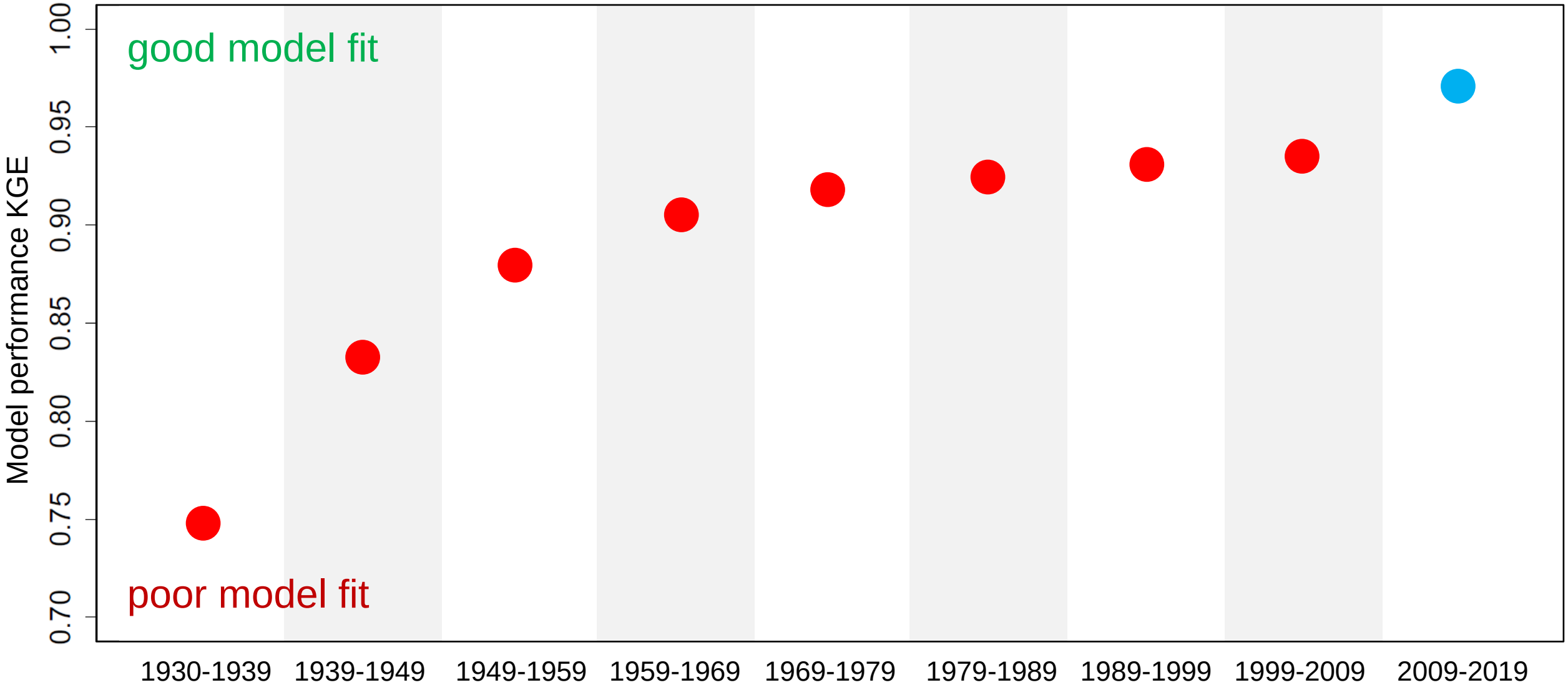
- Lumped box type model
- 13-14 parameters
- Calibration based on optimizing the Kling-Gupta Efficiency (KGE)



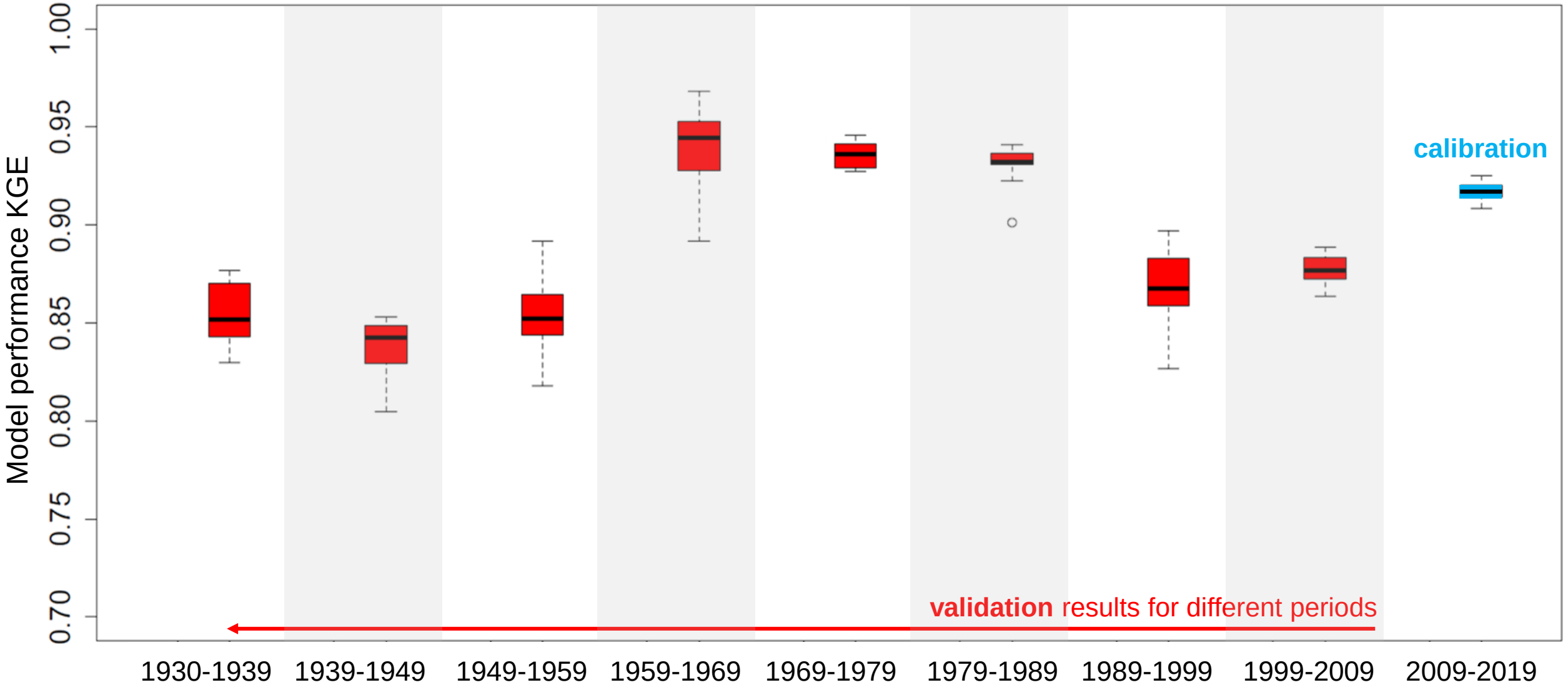
Seibert and Vis, 2012, HESS

Drawing by Petra Seibert

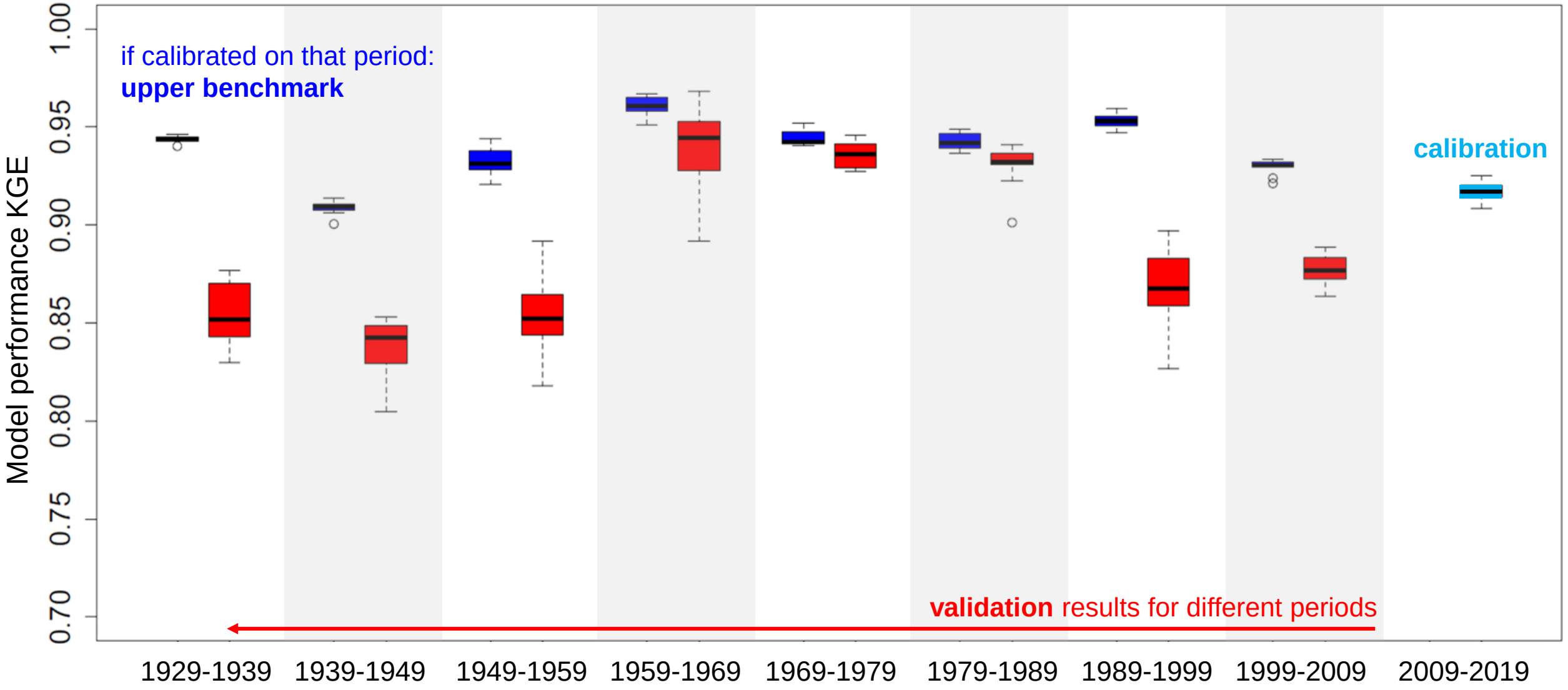
What we expected due to change in bare area/forest cover



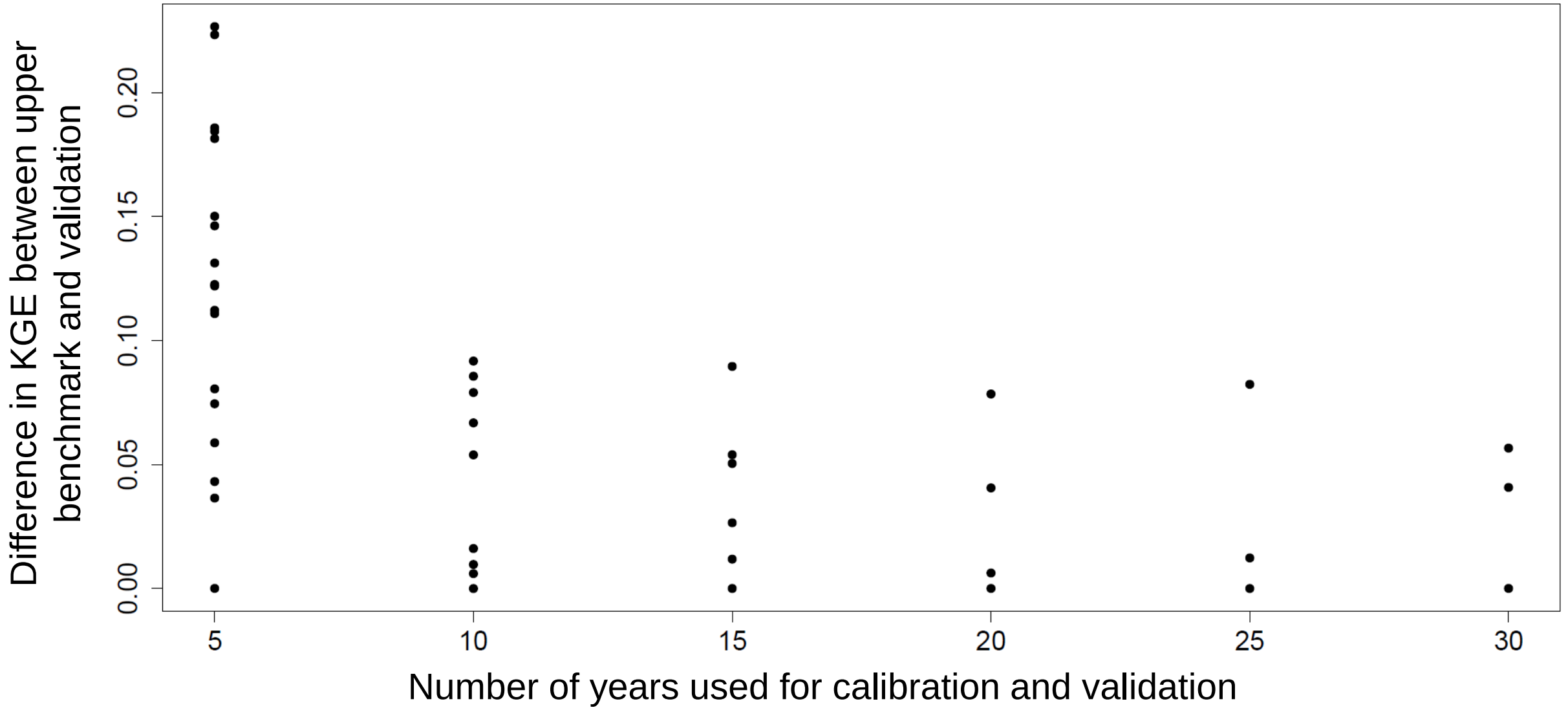
Actual results if calibrated on last 10 years of data



How good the model can be if calibrated on 10 y of data



Difference in model performance



Long-term data from Switzerland: Sperbelgraben

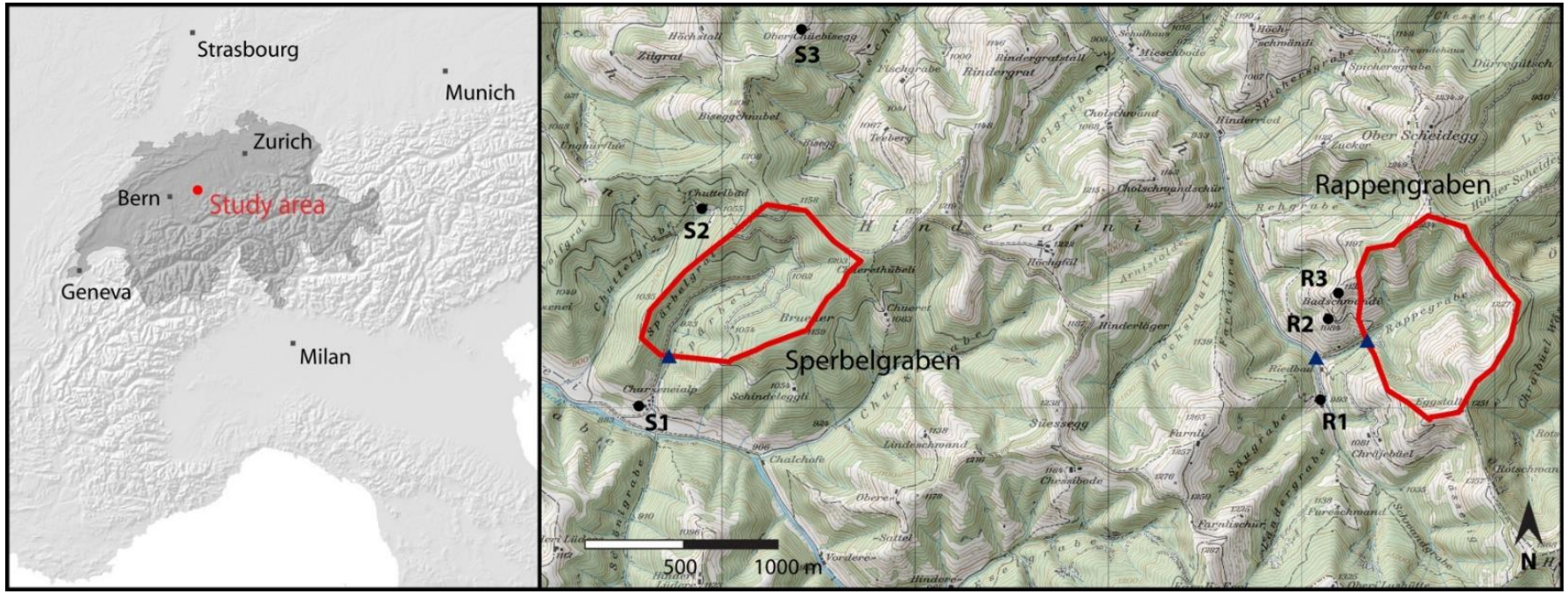
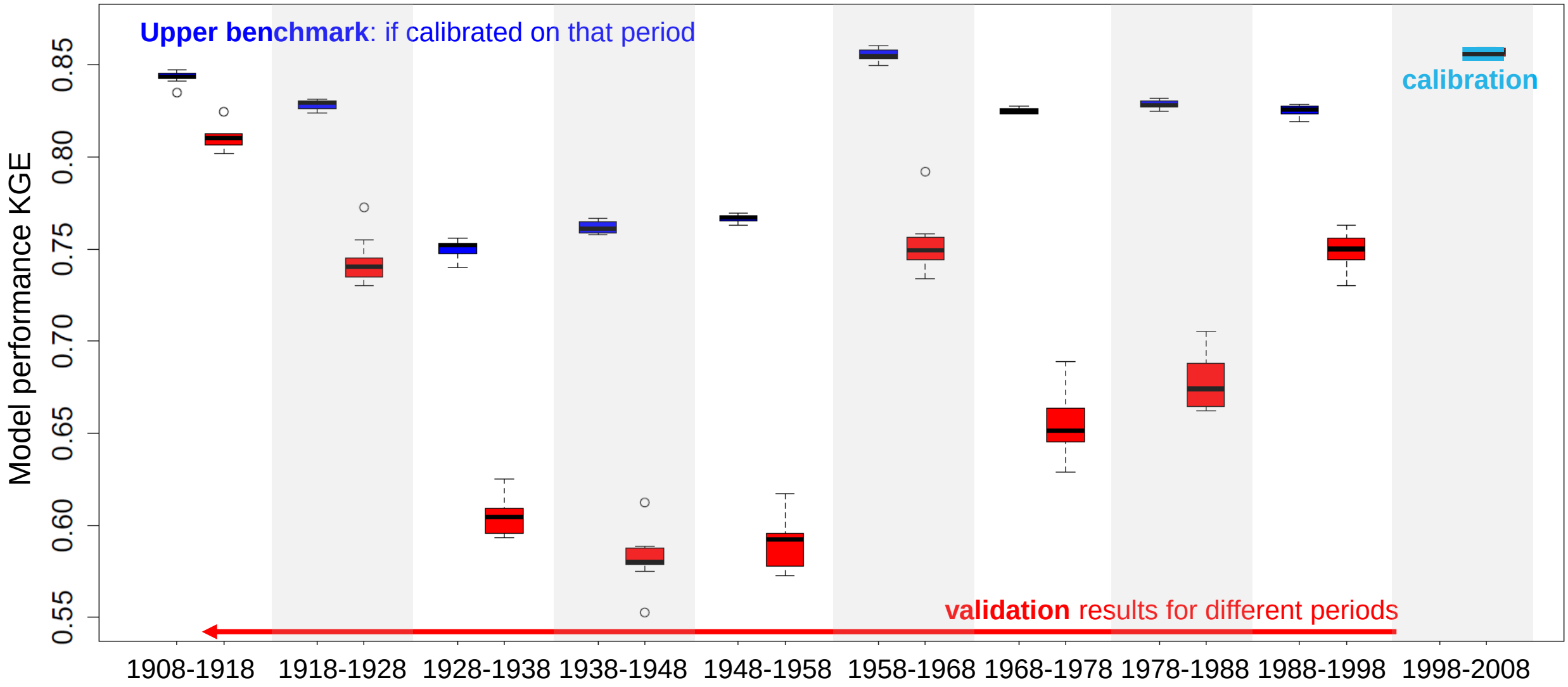


Photo by Jan Seibert

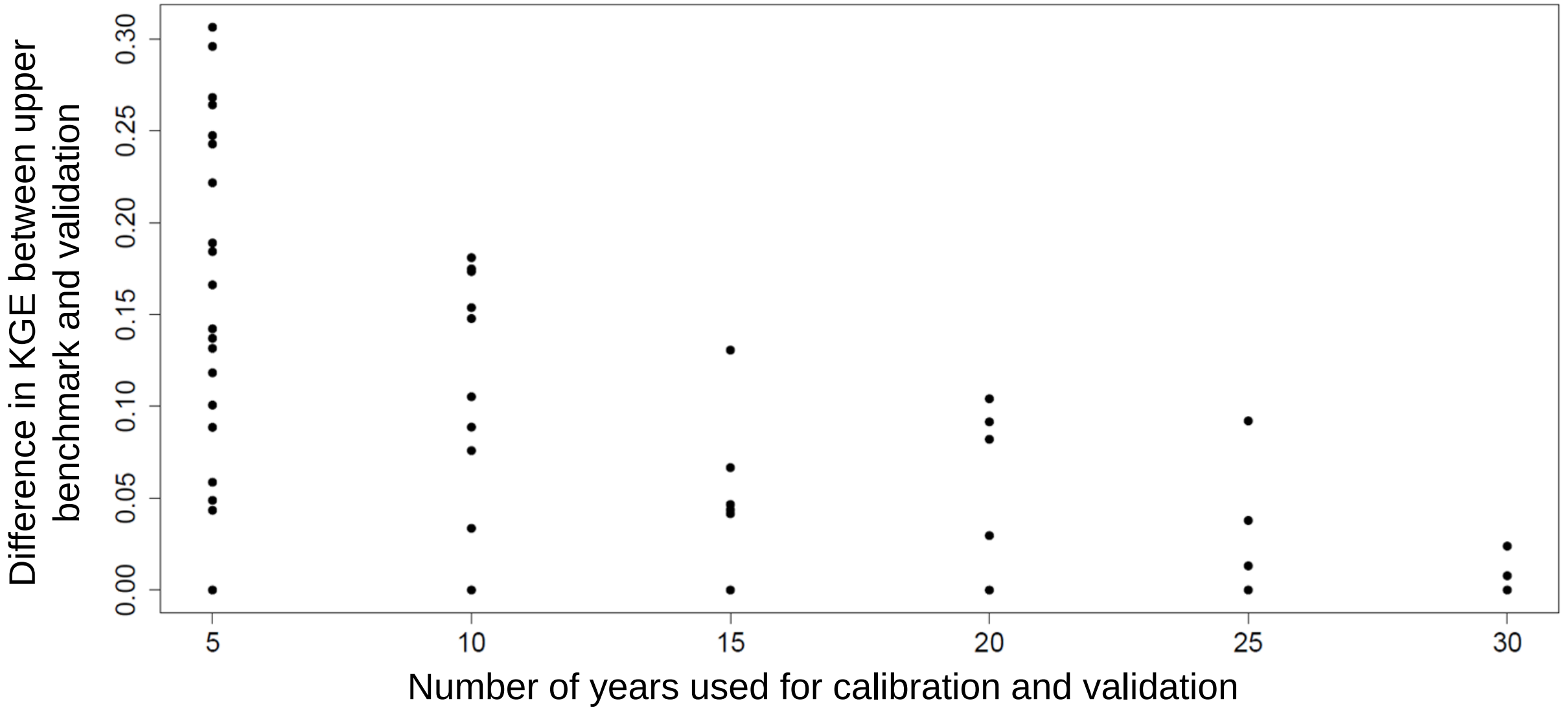
Data from 1908 onwards
0.5 km^2
1640 mm/y (snow: 11%)
Forest cover in 1915: 97%, in 2009: 100%
Large windstorm in 1999
Cambisols on conglomerates and marls

Stähli, et al., Environ. Monit. Assess. 2011, https://doi.org/10.1007/s10661-010-1757-0

Swiss example: Sperbelgraben



Swiss example: Sperbelgraben



Take home messages and open questions

Take home messages

- Difference between calibration and validation performance may depend strongly on the selected time period.
- Don’t forget that catchments change and think about the length of the calibration period – longer is not always better

Open questions

- What causes the high variability in validation results?
- Results for other catchments?

Thank you