An intercomparison of statistical and dynamical downscaling models for precipitation in the European Alps

Jürg Schmidli

Atmospheric and Climate Science ETH

Acknowledgments:

M. Haylock, Y. Hundecha, J. Ribalaygua, T. Schmith, STARDEX team

Workshop on statistical downscaling, Oslo 3–4 October 2005

Intercomparison — Setup

The Downscaling Models

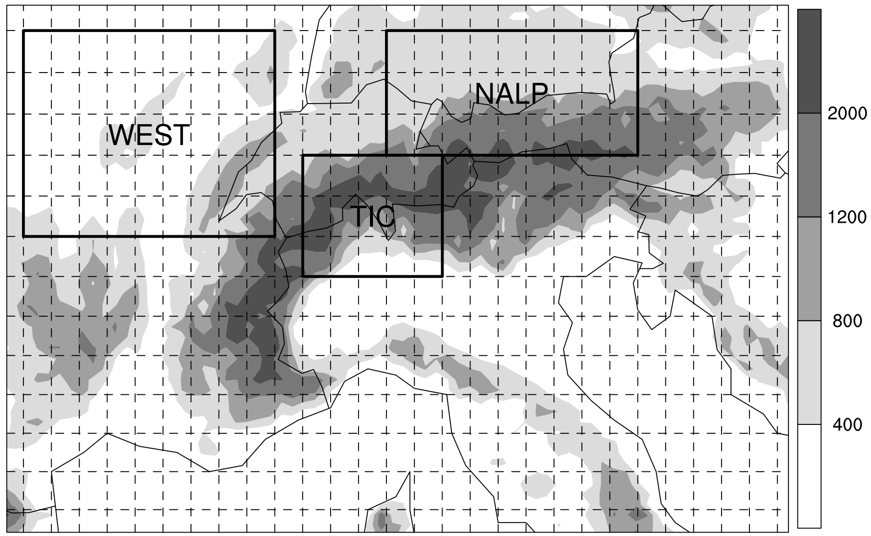
- LOCI benchmark
- 5 SDMs
  (CCA, MLR, MAR, CWG, ANA)
- 3 RCMs
  (CHRM, HadRM, HIRHAM)

Evaluation

- STARDEX indices
  (MEA, FRE, INT, Q90, XCDD, X1D, X5D, NL90)
- Independant period
- Reanalysis and HadAM3

The Predictand

- Precipitation on 0.5° grid
  (or STARDEX indices)
- 10–50 stations per grid-box



Regions

- Northern Alps, Ticino

Local Intensity Scaling (LOCI)

Idea

- Use GCM simulated precipitation as a predictor
  (integrates all relevant large-scale predictors)
- Temporal variability is well simulated (even if spatial variability is not)
- Less prone to stationarity problems

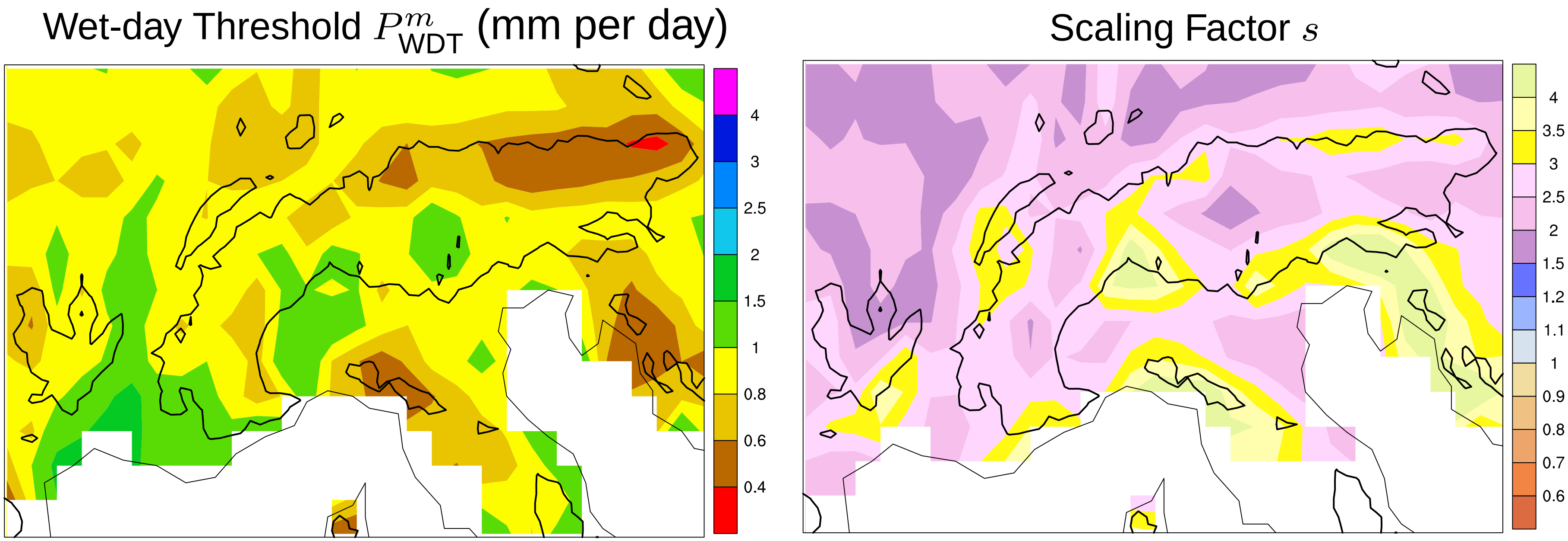
Implementation

- Correct GCM bias in wet-day frequency and precipitation intensity
- Local corrections to the precipitation frequency distribution

$\Rightarrow$ Benchmark for other downscaling methods

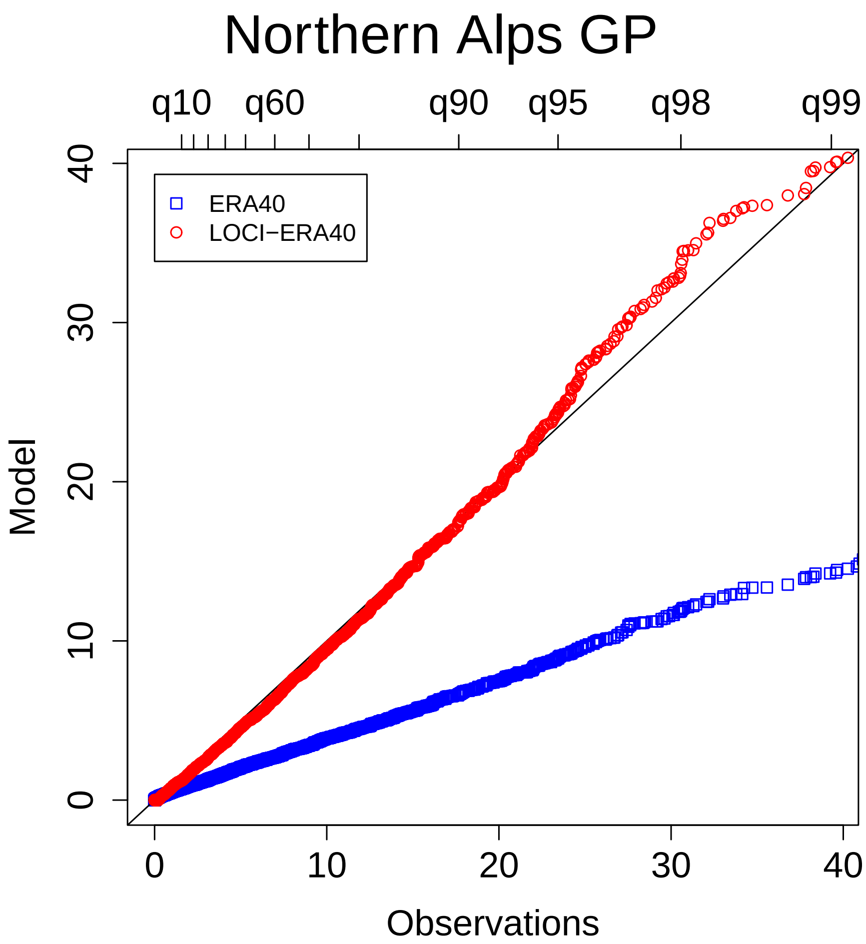
References: Schmidli et al. 2005, Widmann et al. 2003

LOCI: Calibration — ERA40

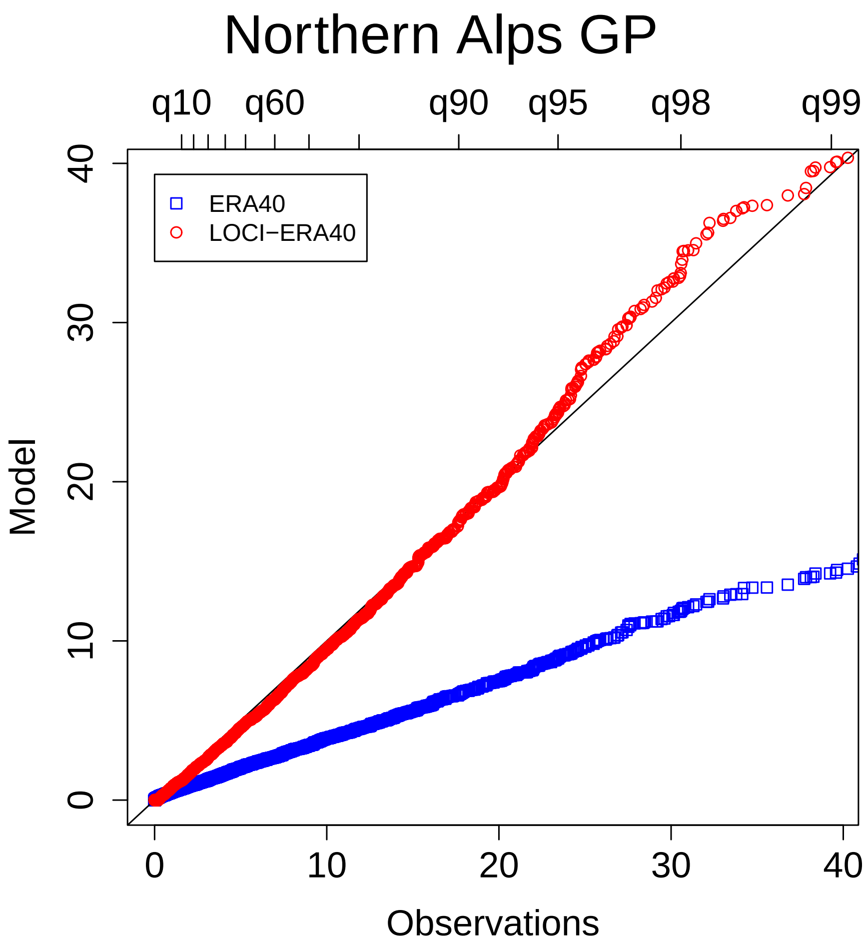


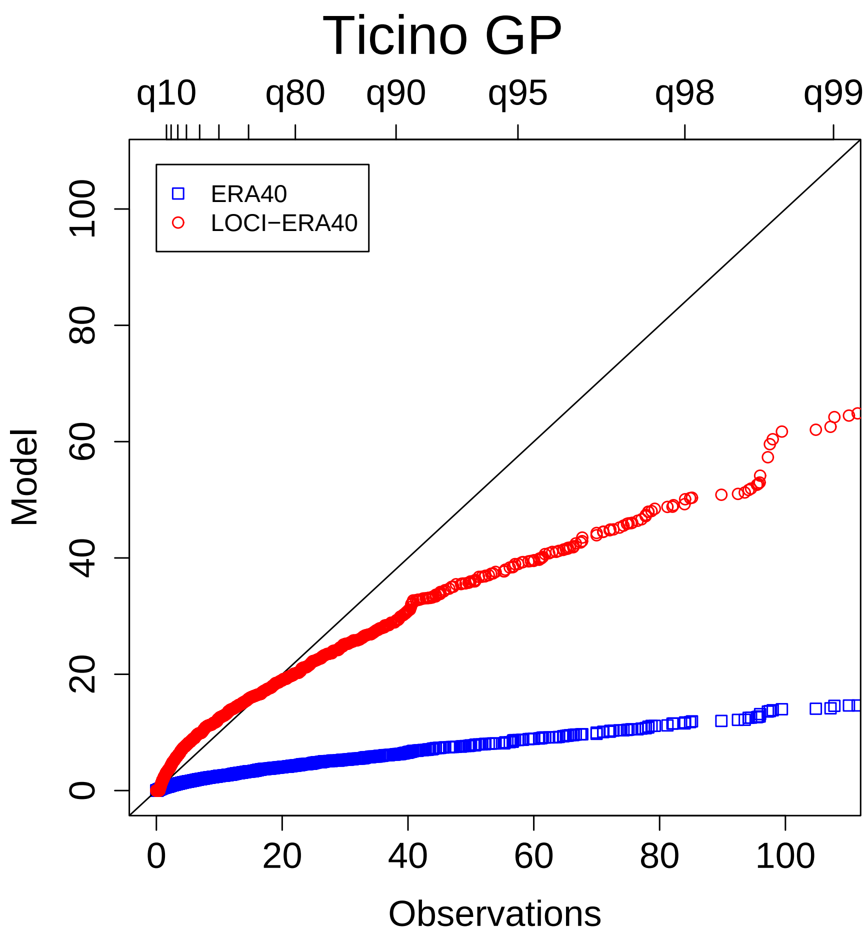
- frequency, slightly too low (mountains)
- intensity, factor 2–4 too low over mountains

LOCI: QQ-Plot — ERA40

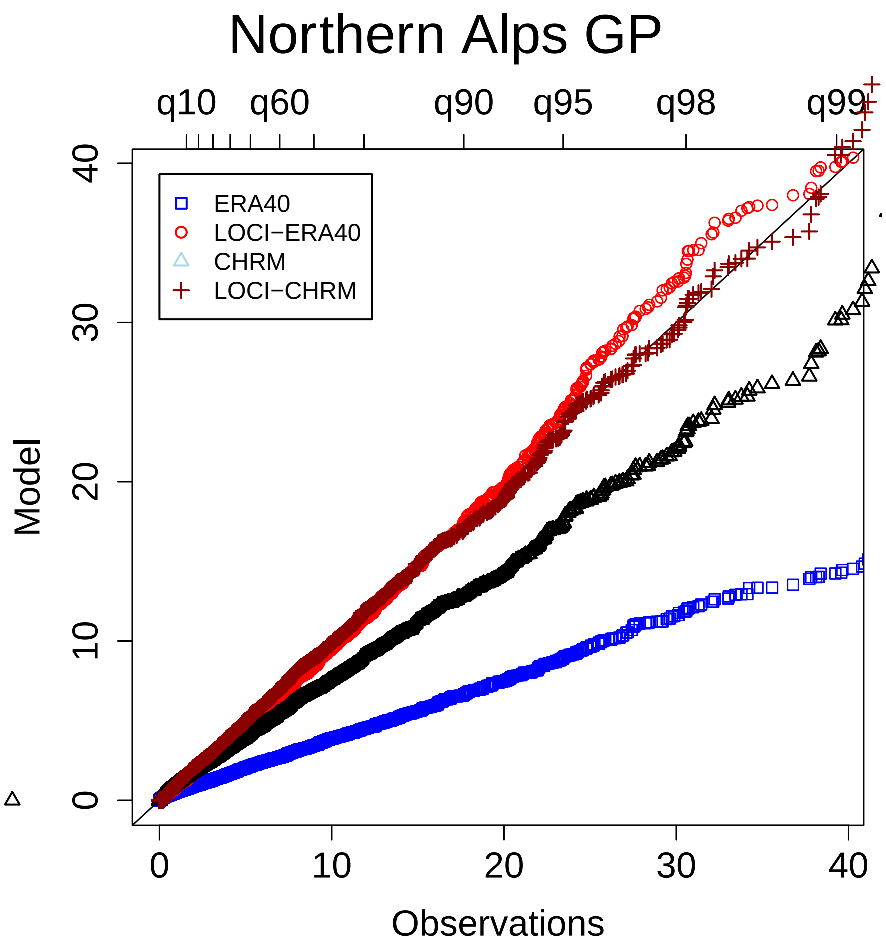


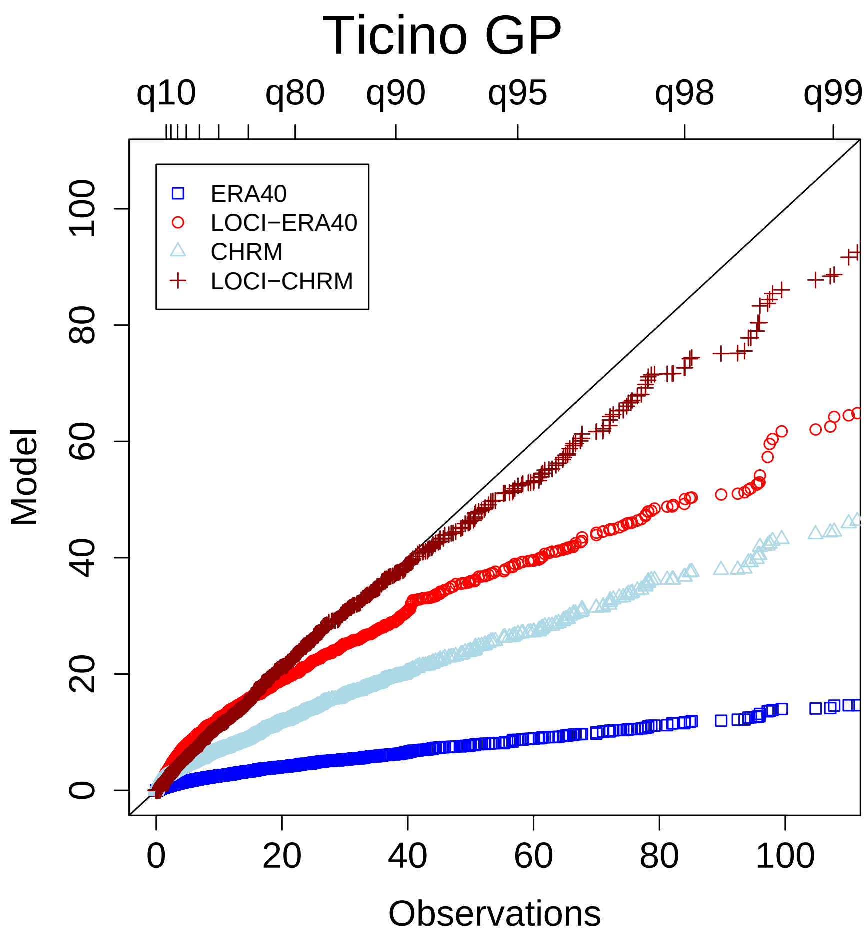
LOCI: QQ-Plot — ERA40



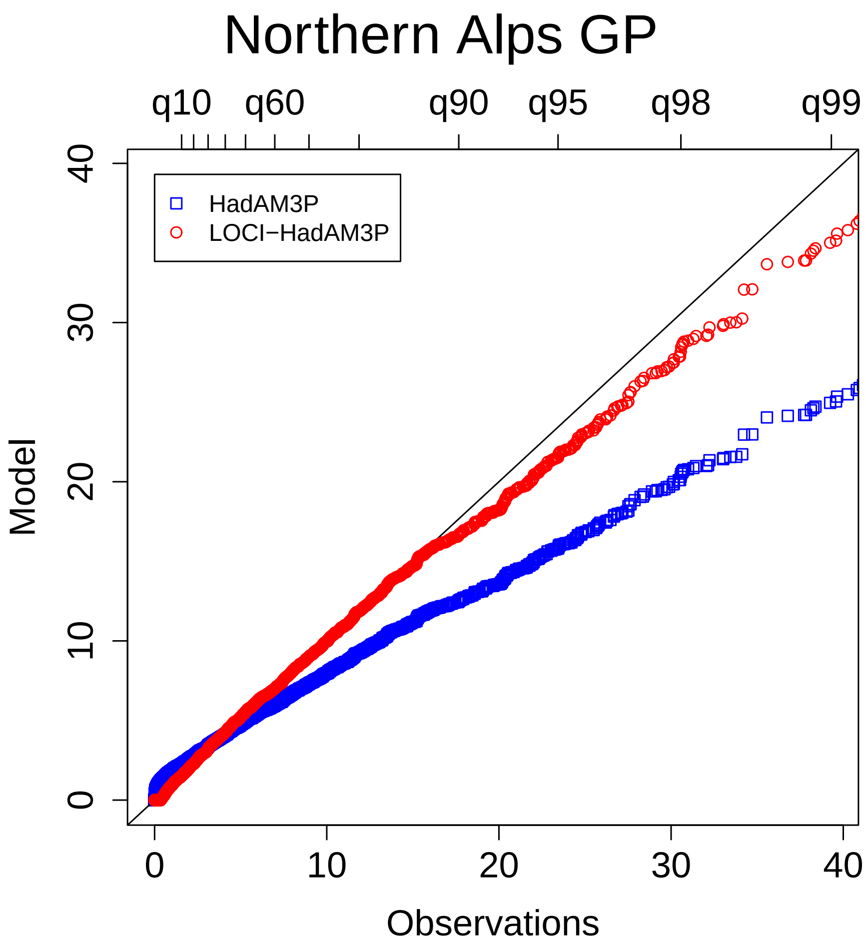


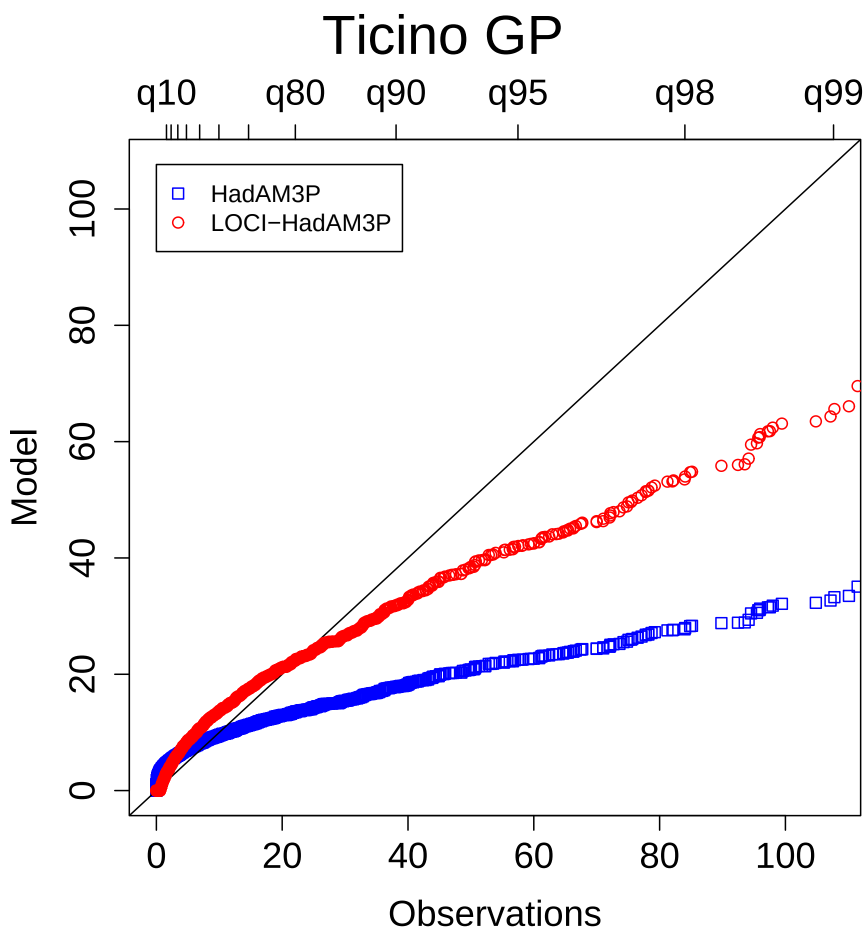
LOCI: QQ-Plot — ERA40





LOCI: QQ-Plot — HadAM3P





Evaluation: Diagnostics

STARDEX indices

- MEA: Mean precipitation
- FRE: Wet-day frequency
- INT: Precipitation intensity
- Q90: 90% quantile of wet-day precipitation amounts
- XCDD: Maximum number of consecutive dry days
- ...

Only two basic groups:

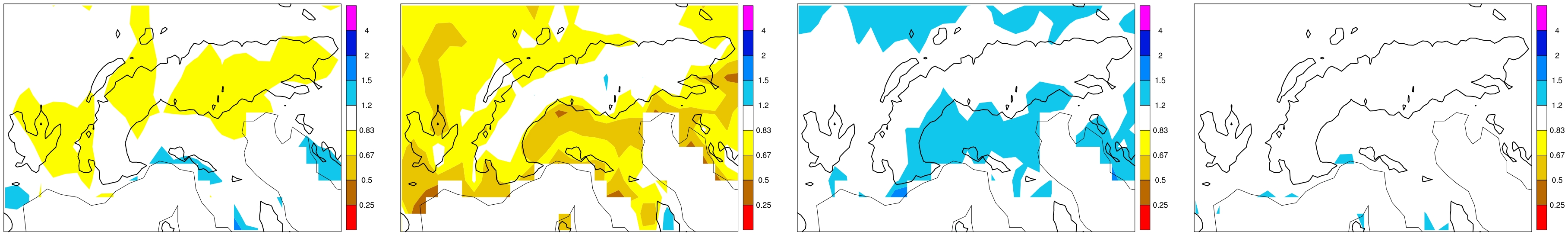
- Occurence process: FRE (MEA, XCDD, NL90)
- Intensity process: INT (Q90, X1D, X5D)

$\Rightarrow$ Other indices are highly correlated with these two basic indices

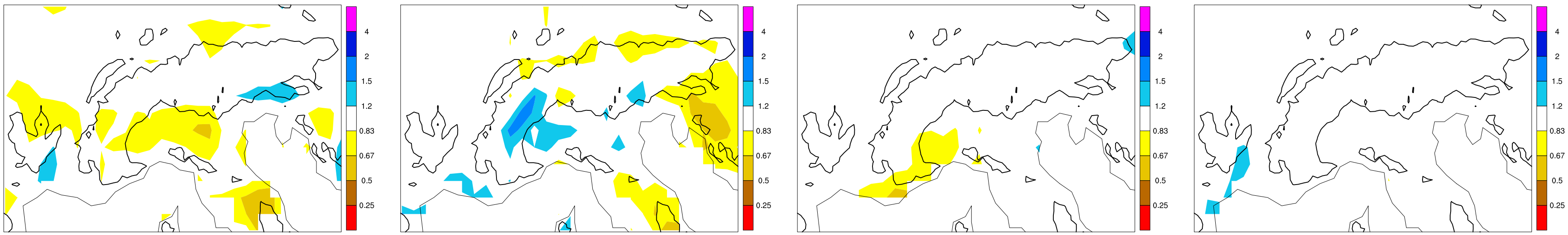
Evaluation: Relative Bias (JJA)

HadAM3 1961–1990 / OBS 1966 –1990

Wet-day Frequency (FRE)



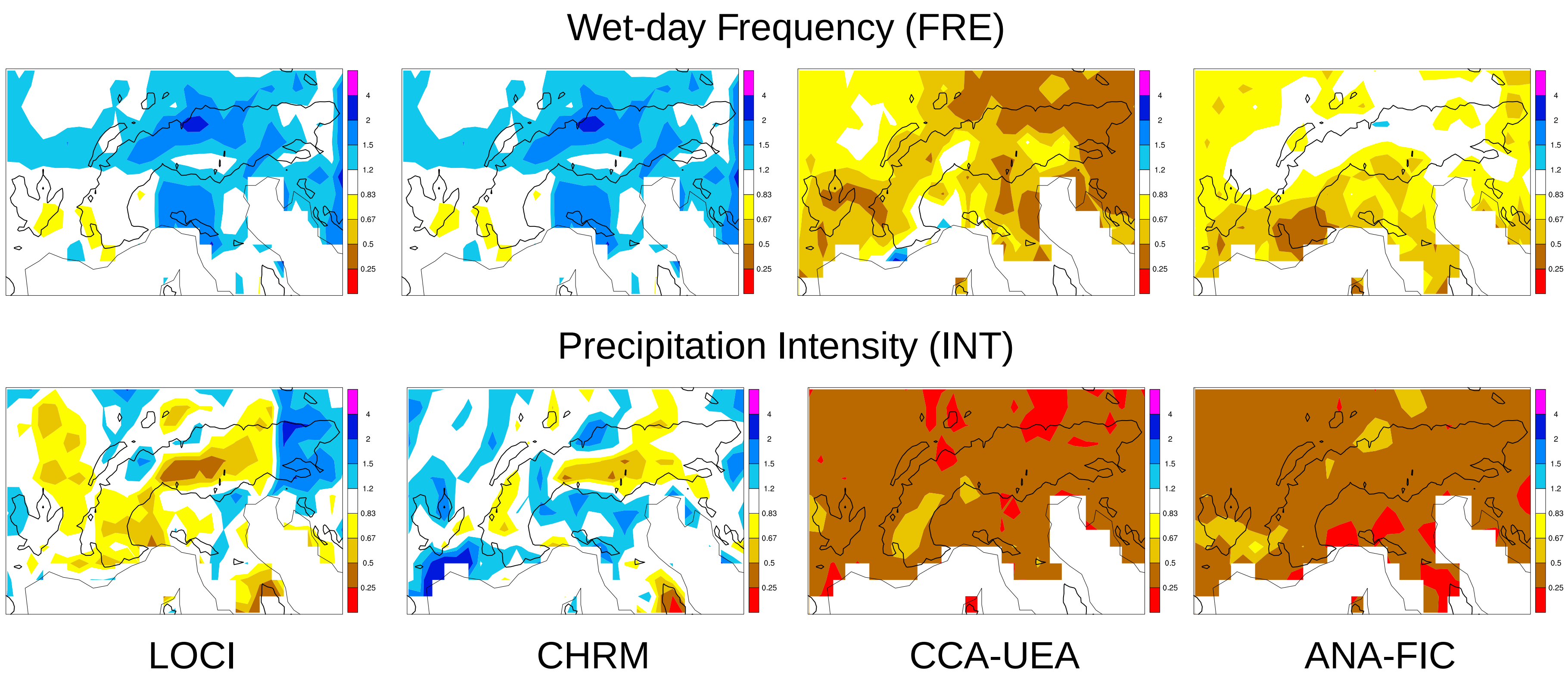
Precipitation Intensity (INT)



LOCI CHRM CCA-UEA ANA-FIC

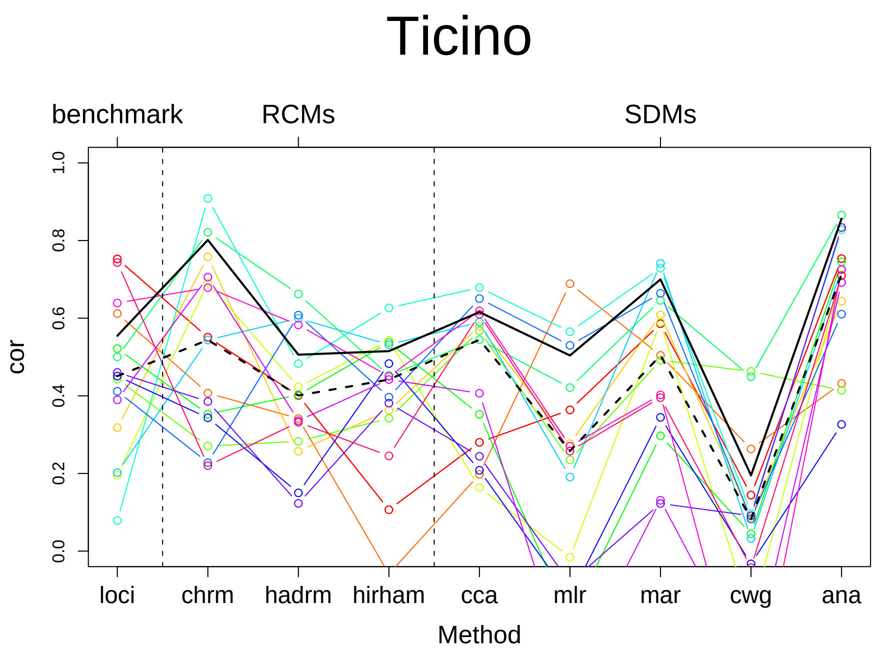
Evaluation: Standard Deviation Ratio (JJA)

HadAM3 1961–1990 / OBS 1966 –1990



Interannual Variability of INT — Summer (JJA)

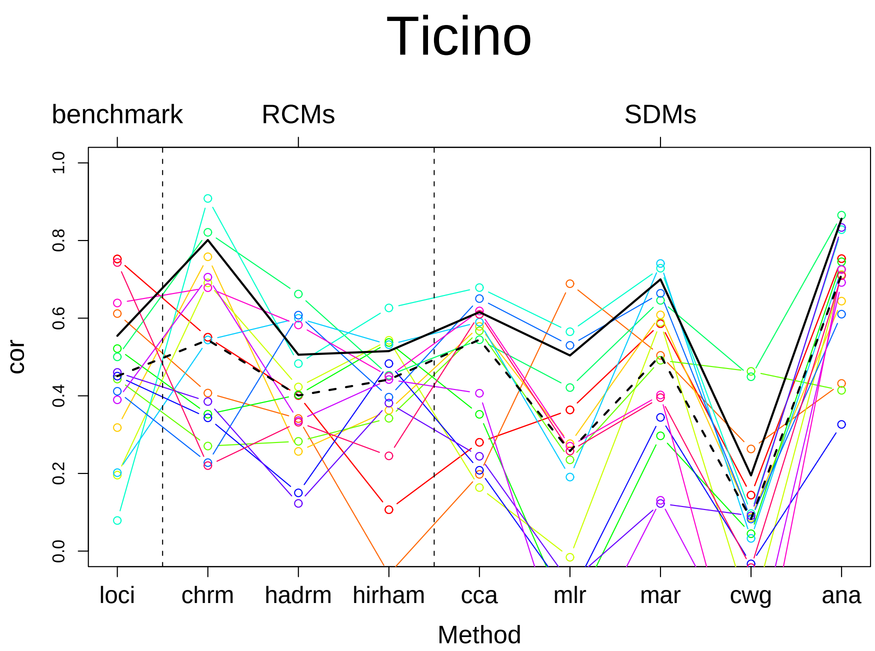
Correlation skill for precipitation intensity (INT), for individual GPs

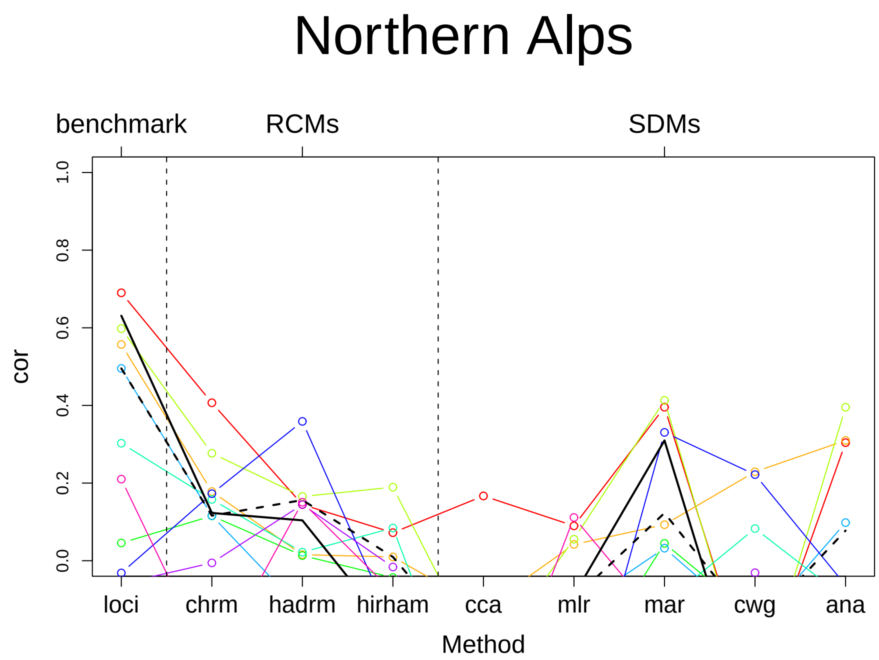


⇒ Very large variability within region!

Interannual Variability of INT — Summer (JJA)

Correlation skill for precipitation intensity (INT), for individual GPs

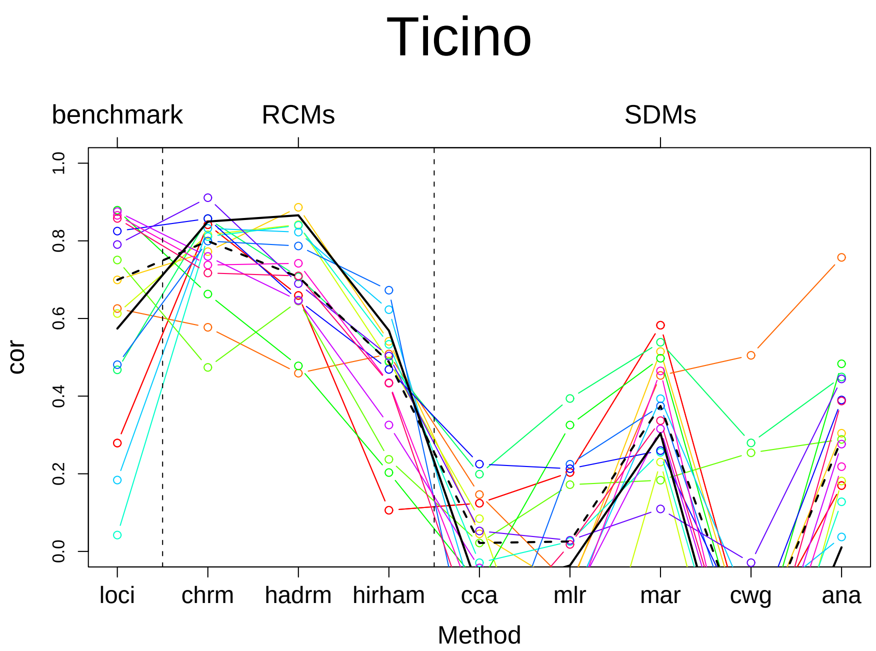


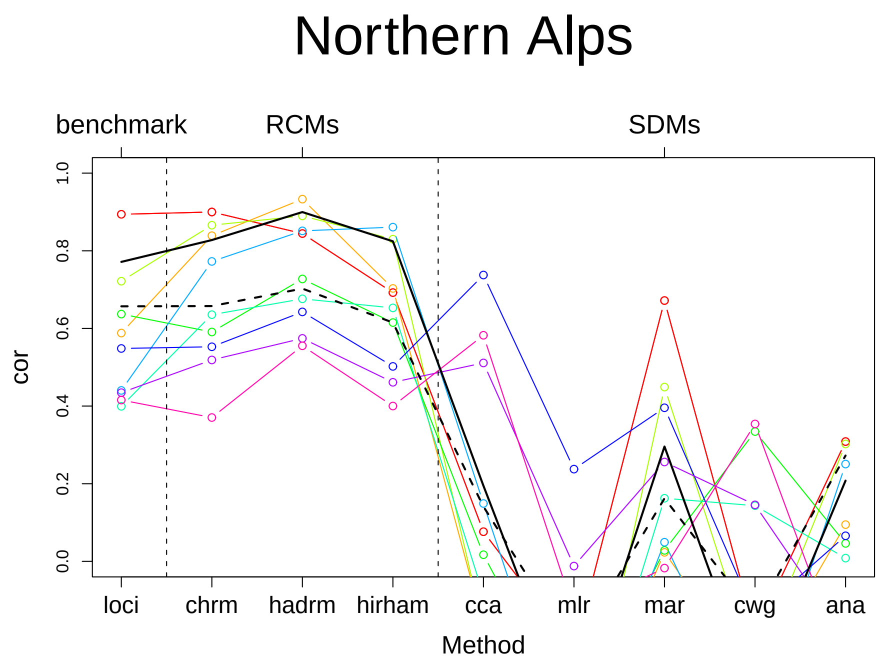


$\Rightarrow$ Very large variability within region! Between regions!

Interannual Variability of INT — Winter (DJF)

Correlation skill for precipitation intensity (INT), for individual GPs





$\Rightarrow$ And between seasons! Intercomparsion?

Conclusions

- LOCI: Current GCMs may have good skill, despite large biases
- Large variations between regions, seasons, and within regions!
- Alps: RCMs, ”daily” SDMs
- FRE good proxy for occurence-related diagnostics (XCDD, MEA, . . .)
- INT good proxy for intensity-related/extreme diagnostics (Q90, X1D, . . .)

Recommendations for Intercomparison

→ Use LOCI as benchmark

→ Evaluate pdf (e.g. mean, variance) at *local* scale

→ BUT: Evaluate temporal correspondance at *regional* scale

→ Evaluate occurence and intensity process (FRE and INT)

→ Spatial aggregation is essential for intercomparison
(of methods, regions, and seasons)

→ Evaluate predictors! Perfect? Stationarity?

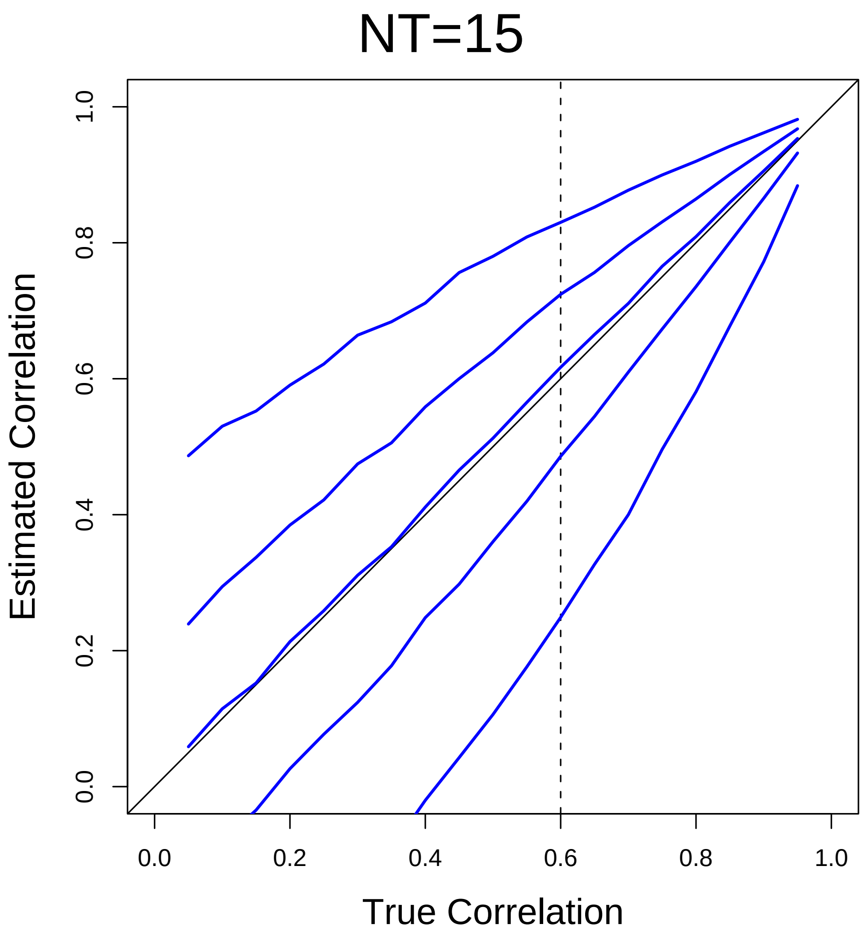
→ Data exchange infrastructure (e.g. OpenDAP/DODS)

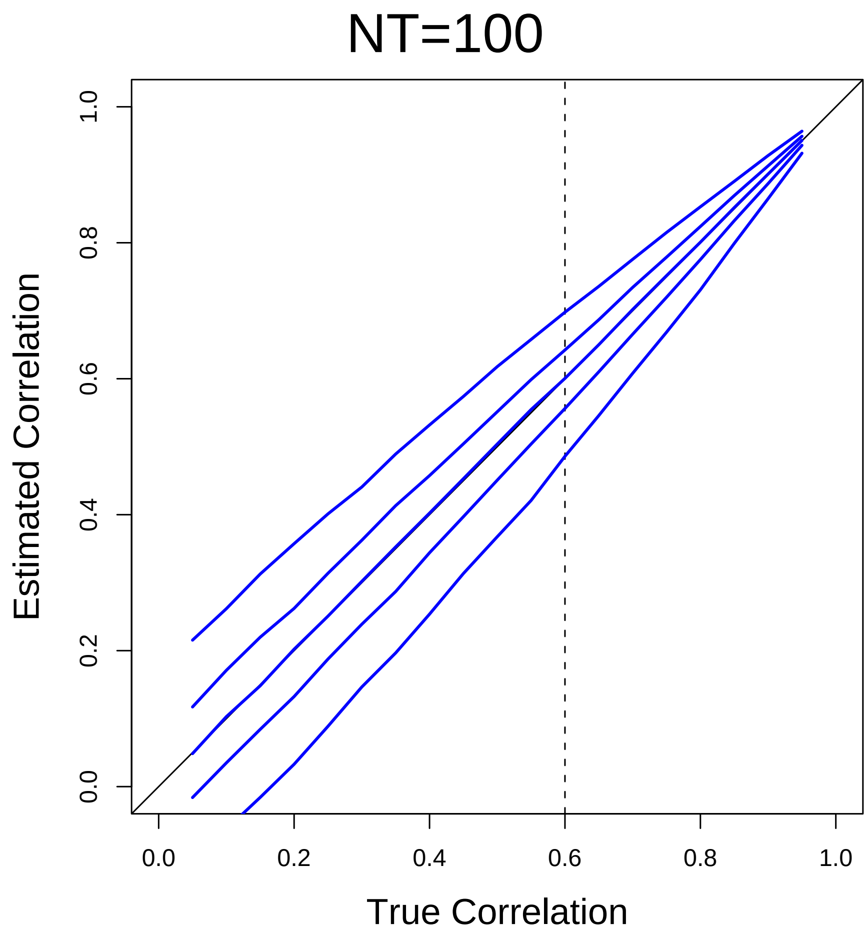
Recommendations for Intercomparison

- B: Extremes: Are some indicators better/more robust? Methods?
  $\longrightarrow$ include basic diagnostics (understanding)
  $\longrightarrow$ stronger extremes may be more predictable
- E: Methods: General recommendations?
  $\longrightarrow$ no method is always superior (but ...)
  $\longrightarrow$ always use a range of methods
- F: Uncertainty associated with statistical downscaling?
  $\longrightarrow$ spatial variability, predictability, sampling error
- H: SMIP: Basic principles/standard for comparison of models?
  $\longrightarrow$ see previous slide

Estimation of the Correlation Coefficient

Dependance of Sampling Error on Sample Size





50% and 90% confidence interval for the correlation coefficient