WAVES TO WEATHER

Convergence of forecast probabilities with increasing ensemble size

The A6 Team: George Craig, Kirsten Tempest, Matjaž Puh, Christian Keil and Jonas Brehmer

LMU JG|U KIT TUM DLR UH

The problem of ensemble size

Motivation

- Ensembles add value in comparison to deterministic forecasts
- Current forecast ensembles are probably much too small

Questions

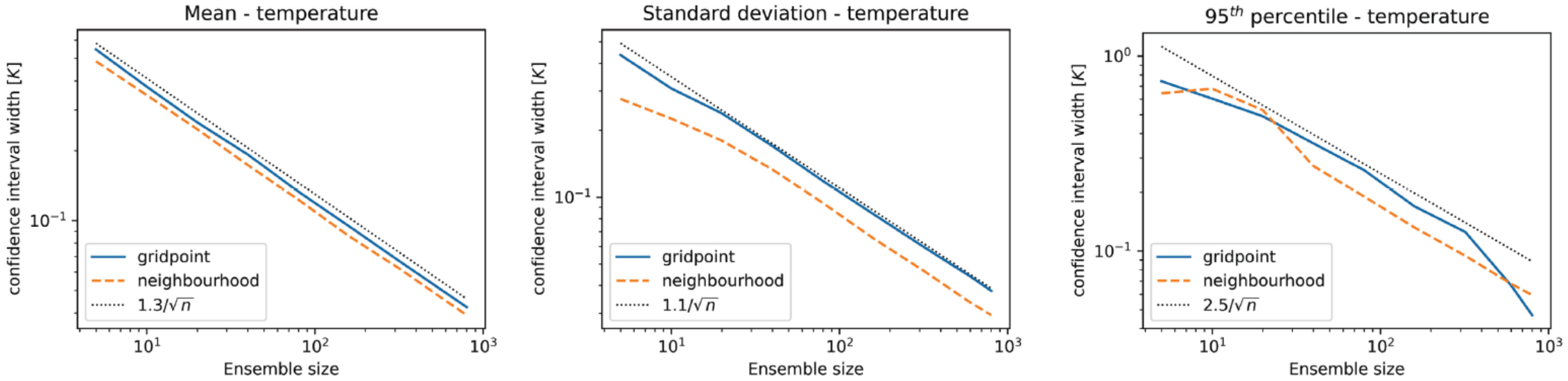
- How large are the errors caused by small ensemble size, and what do they depend on?
- How effective are strategies to compensate for small ensemble size?

First look – 1000 member convective ensemble

- SCALE-RM regional NWP system
- Resolution 2 km
- Forecast of T (z = 5km)

Width of 95% confidence interval for sample estimates of:

- mean
- standard deviation
- 95th percentile



$n^{-1/2}$ convergence
(for large enough n)

Larger n required for
tail of distribution
(rare events)

Theory – asymptotic for large *n*

σ_{n_p} – standard deviation of sample estimate of p^{th} quantile

n – ensemble size

$f(q_p)$ – probability density at p^{th} quantile q_p

$$\sigma_{n_p} = \frac{1}{\sqrt{n}} \sqrt{\frac{p(1-p)}{f^2(q_p)}}$$

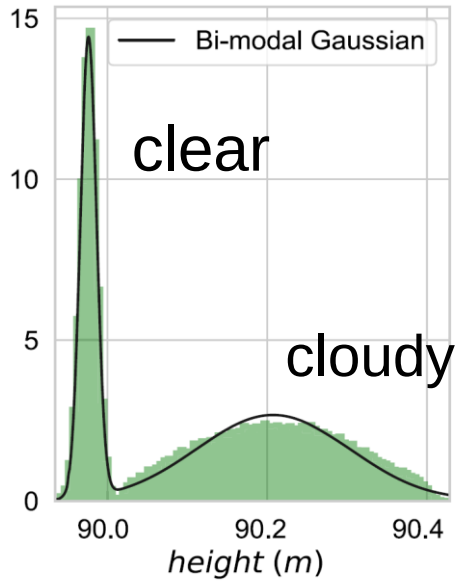
$n^{-1/2}$ convergence (for large enough n)

Larger n required for tail of distribution (rare events)

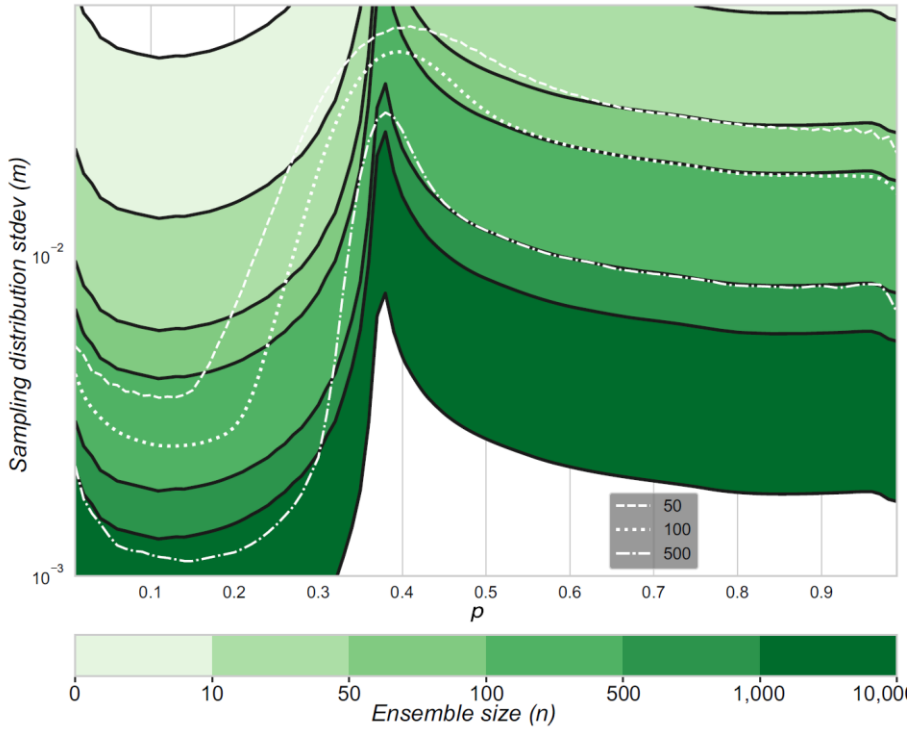
Importance of underlying distribution

- Toy model: shallow water equations with modification for latent heat release
- Ensemble size n = 100 000
- Forecast for h – proxy for total water (vapor plus liquid)

Forecast PDF:



Ensemble size required to estimate quantile p with a given accuracy – larger for rare events

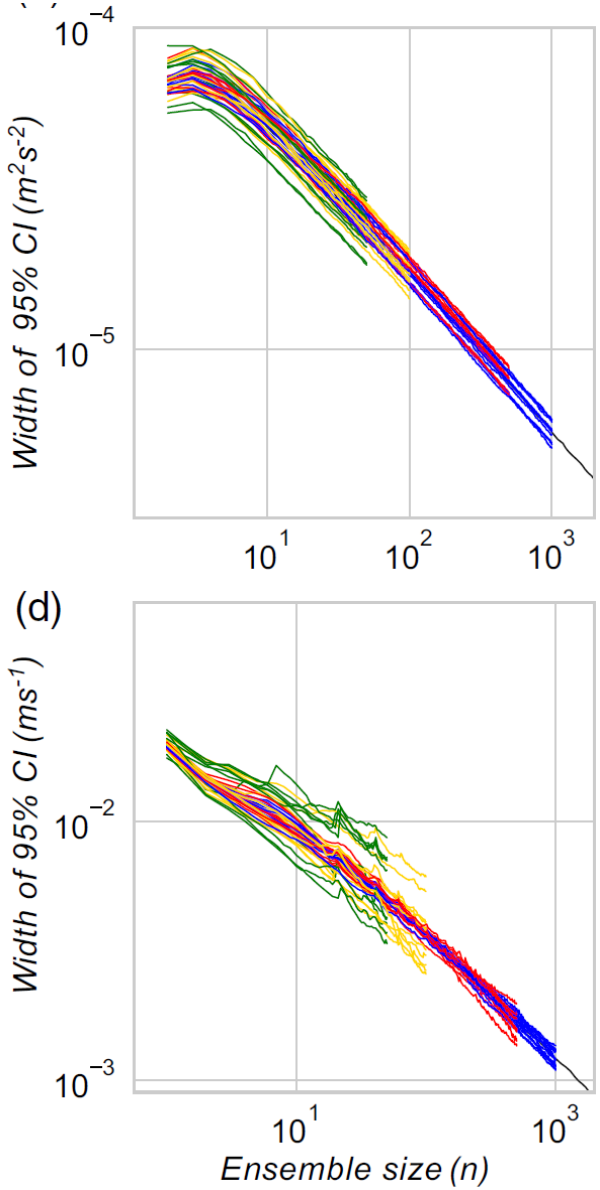


Small ensembles

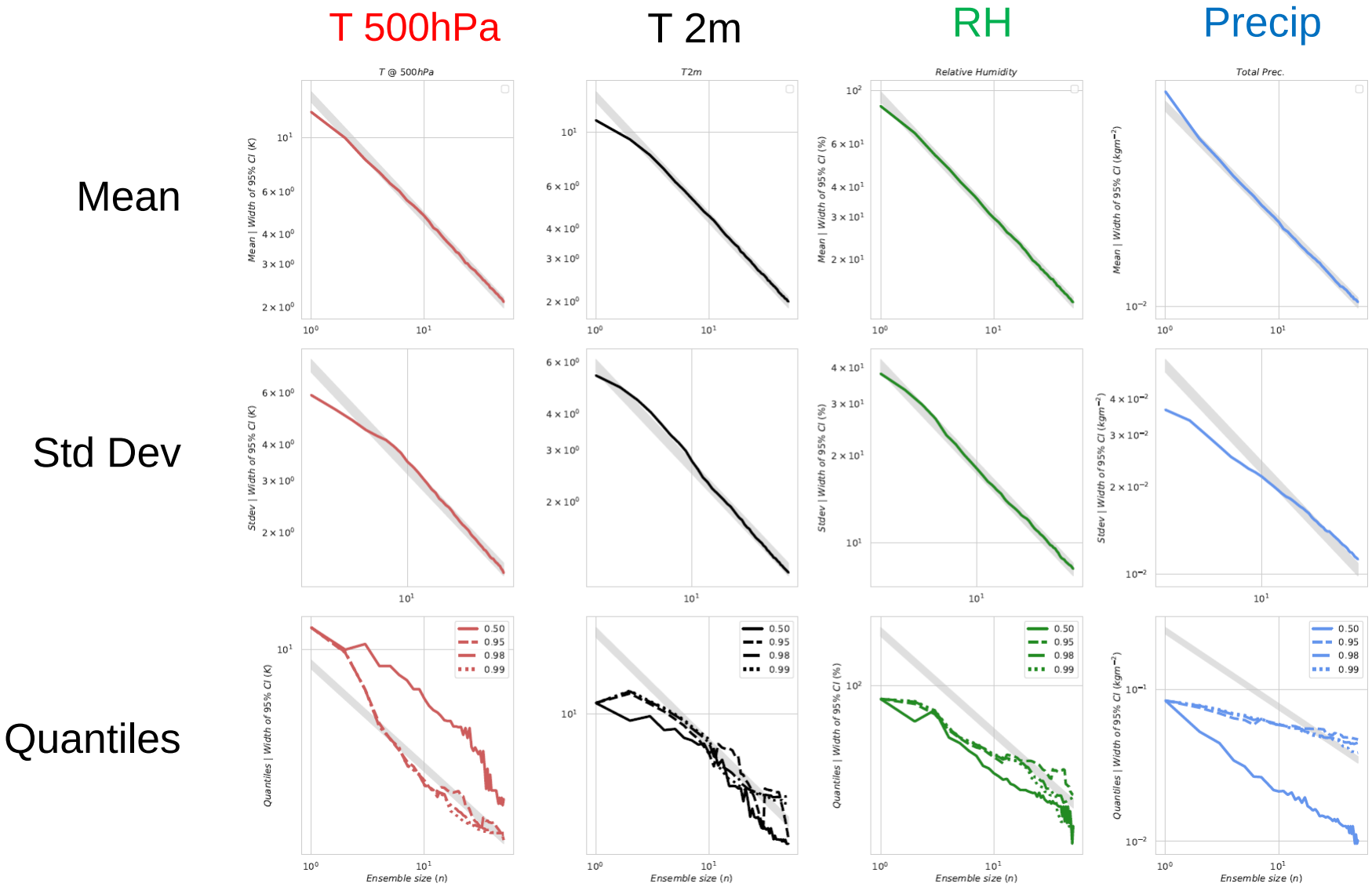
- See $n^{-1/2}$ scaling if in convergence regime, but with some offset

Note on Bootstrapping

- Estimate confidence interval assuming ensemble accurately represents true distribution
  - *n* large – all ok
  - *n* small – inaccurate estimate of confidence interval
- Should be possible to detect whether forecast is in asymptotic convergence regime

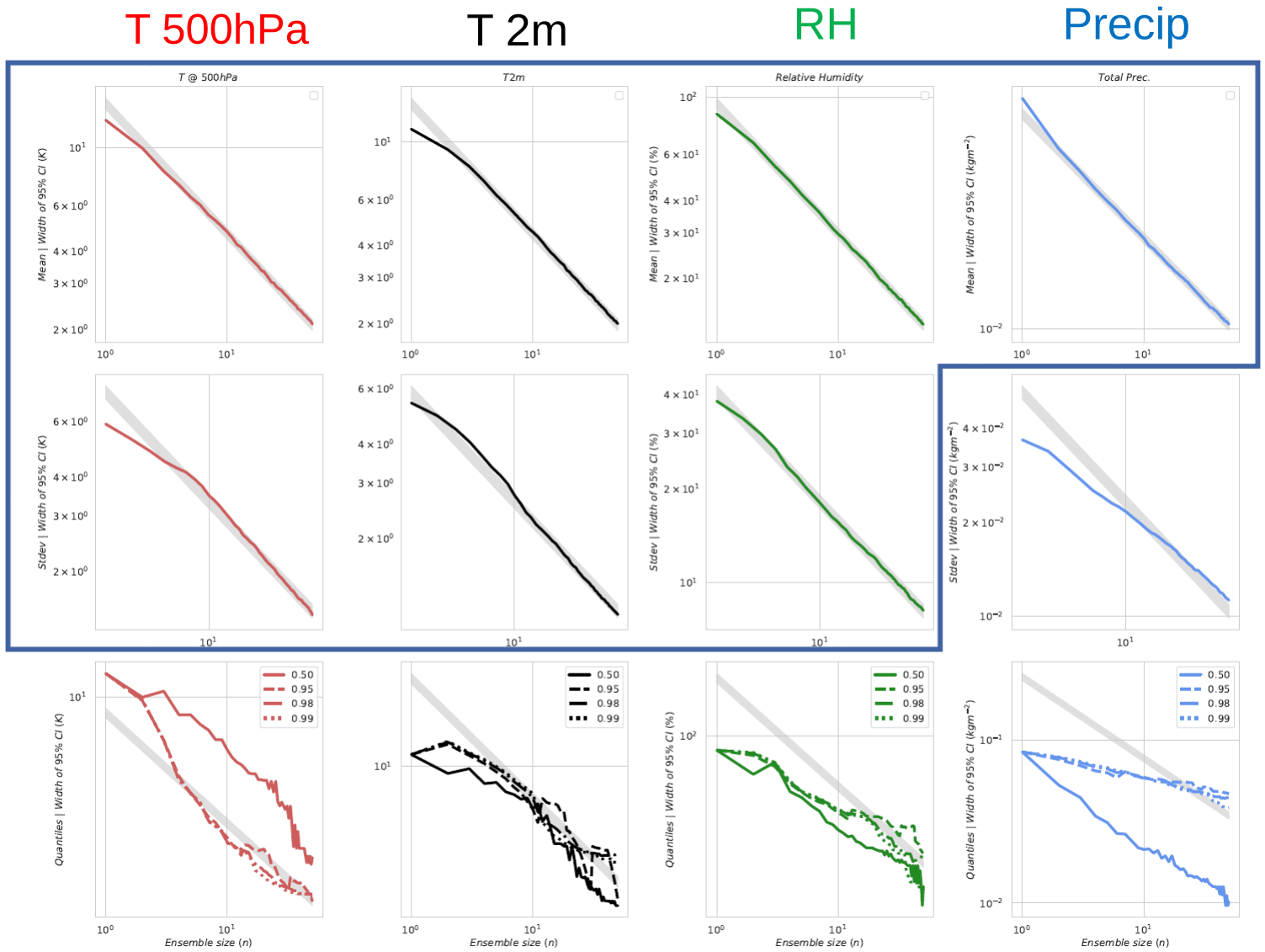


Second look – ECMWF forecasts



- Global NWP model
- Selected variables at a single gridpoint
- Ensemble size $n = 51$
- Ensemble designed for reasonable spread-skill relation

Second look – ECMWF forecasts

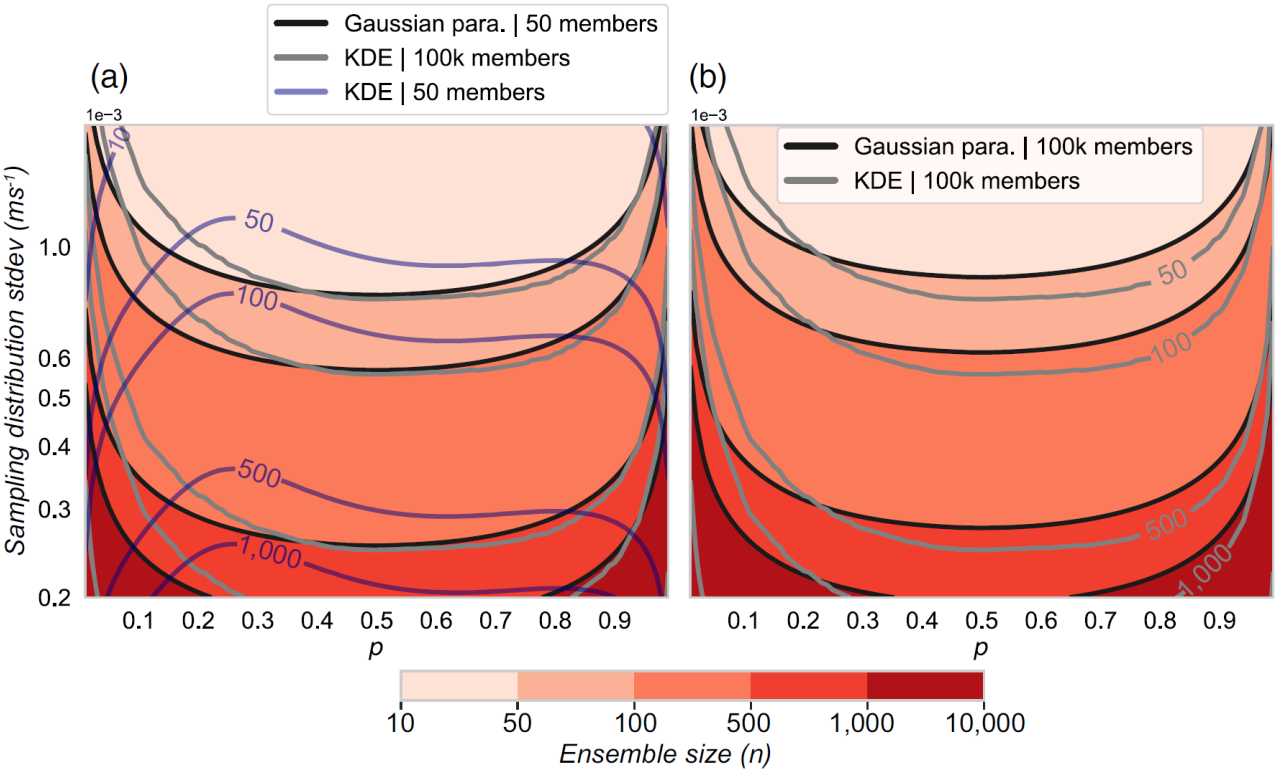


$n^{-1/2}$ convergence (for large enough n)

Recall: Ensemble designed for reasonable spread-skill relation

no evidence of convergence

Using small ensembles I – parametric fits



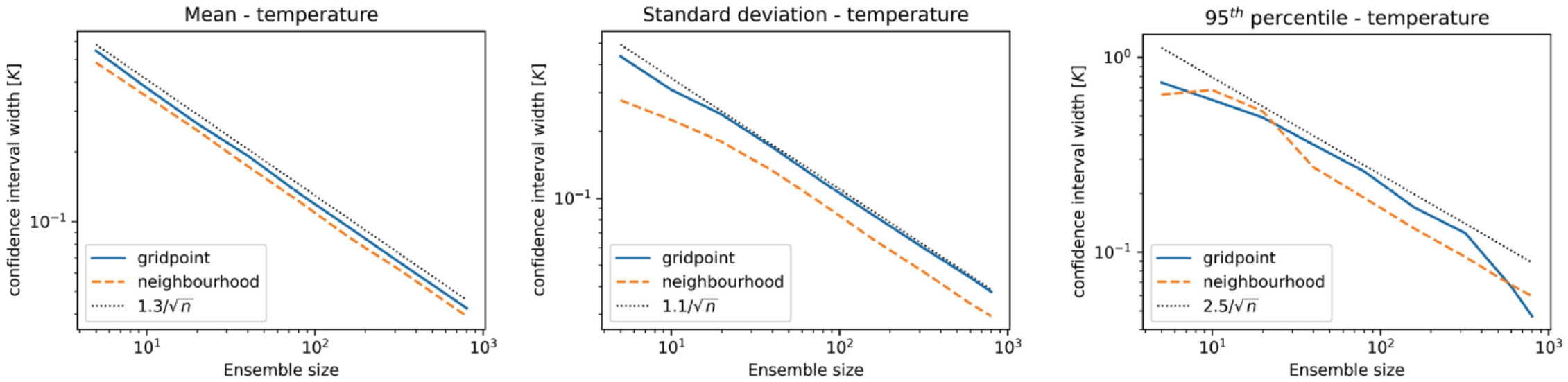
- Fit Gaussian to 50 member ensemble of wind
- Good estimate, comparable to Gaussian fitted to 100k members
- Much better than from raw 50-member ensemble
- This method may be implicitly included in statistical post-processing

Using small ensembles II – neighborhoods

- SCALE-RM regional NWP system
- Resolution 2 km
- Forecast of *T* (*z* = 5km)

Width of 95% confidence interval for sample estimates of:

- mean
- standard deviation
- 95th percentile



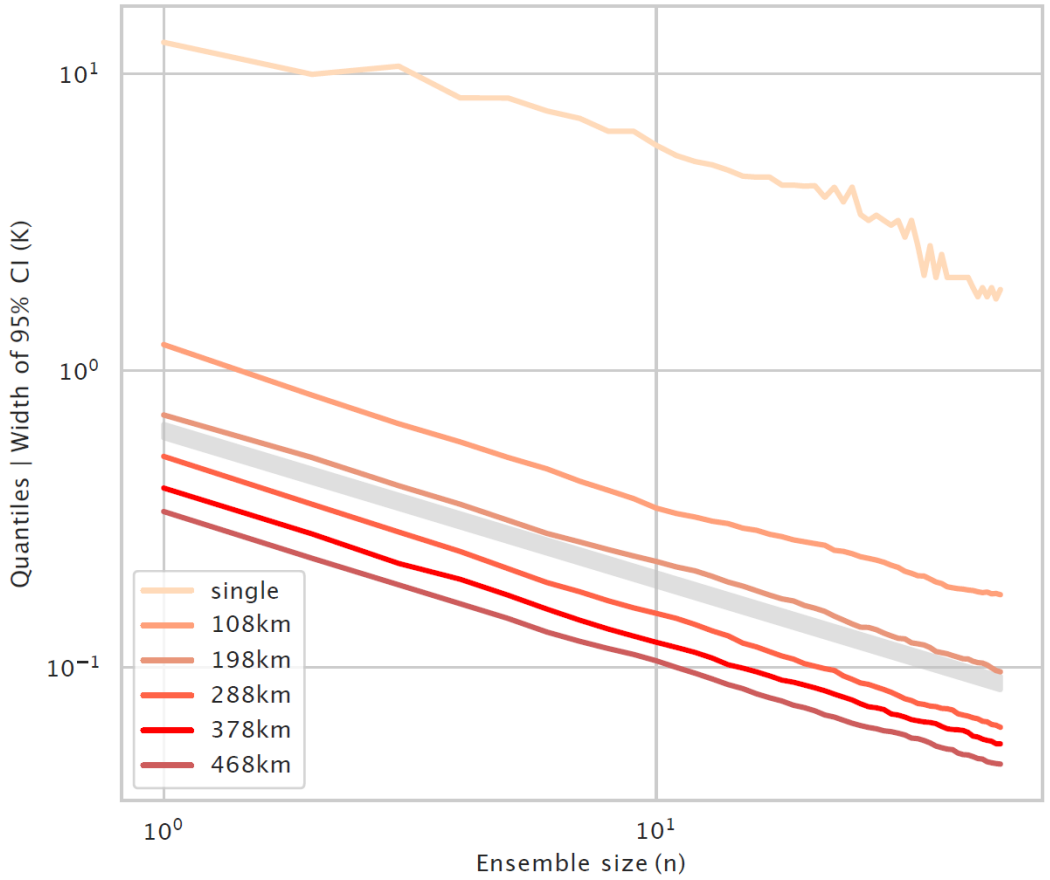
$n^{-1/2}$ convergence (for large enough n)

Neighborhood increases effective ensemble size

Using small ensembles II – neighborhoods

Treat points within a neighborhood as independent realizations of gridpoint state

- ECMWF temperature forecasts of 50th percentile
- Single gridpoint and neighborhood sizes from 108 to 468 km
- $n^{-1/2}$ convergence for neighborhoods larger than 198 km
- Effective ensemble size increased by about ¼ neighborhood size

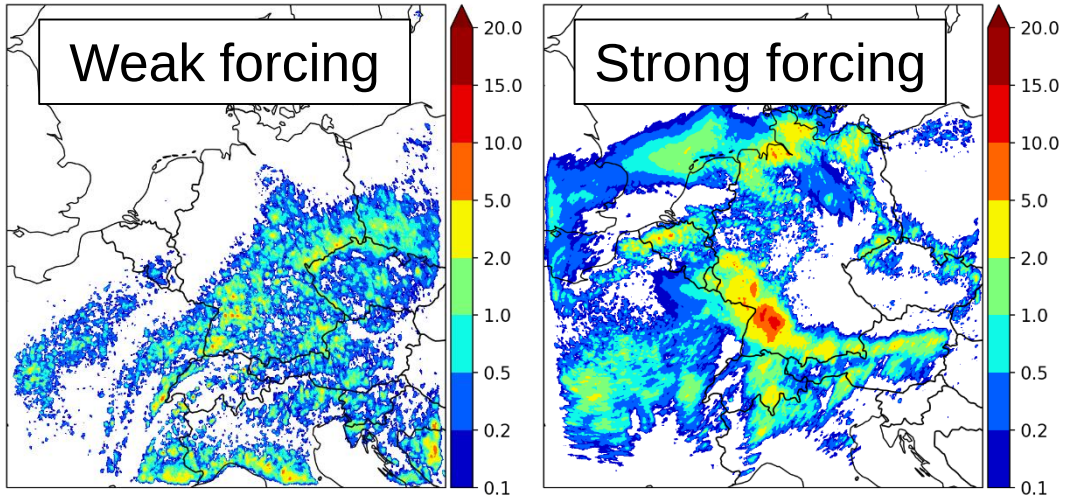


Variability with location and weather regime

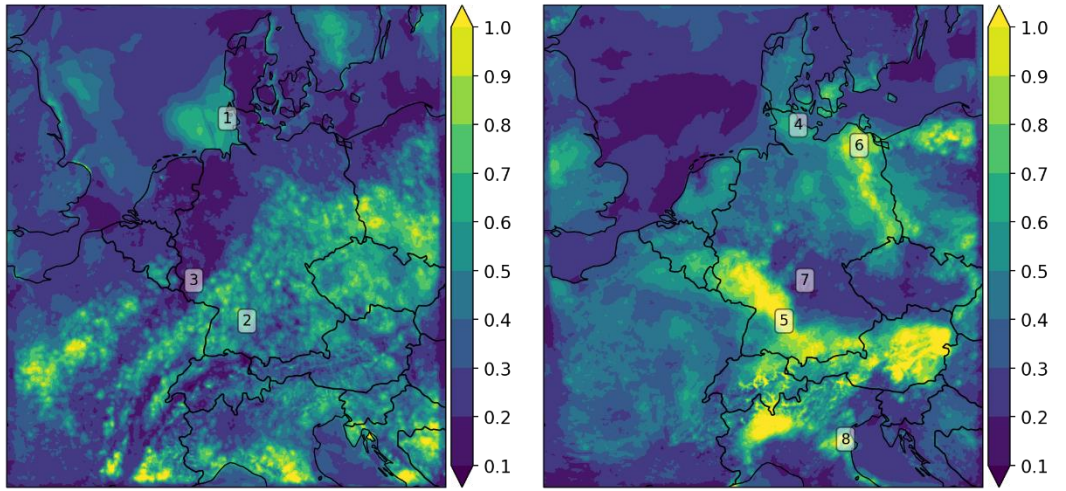
ICON-D2 NWP ensemble

- Resolution 2 km
- 120 members

Precipitation



Width of 95% confidence interval for ensemble mean 2m temperature



Increased uncertainty in regions of convection (broad distributions)

Synoptic forcing adds additional uncertainty

Conclusions

- Asymptotic theory predicts universal convergence law
- Bootstrap estimates of confidence intervals show if in convergence regime
- Current ensembles show convergence for some variables, typically ensemble mean and variance, but not rare events
- Parametric fits, neighborhoods effective in extracting useful estimates from small samples
- Current work – dependence on forecast variable, location, weather regime