Sampling errors on convective scales: What can we learn from a 1000-member ensemble?

T. Necker[1], M. Weissmann[1], Y. Ruckstuhl[2], J. Ruiz[3], T. Miyoshi[4] and J. Anderson[5]

(1) Universität Wien, Vienna, Austria (2) LMU, Munich, Germany (3) CIMA/CONICET-UBA, Buenos Aires, Argentina
(4) RIKEN Center for Computational Science, Kobe, Japan (5) NCAR/Data Assimilation Research Section, Boulder, Colorado

Abstract

Current regional forecasting systems particularly aim at the forecast of convective events and related hazards. Most weather centers apply high-resolution ensemble forecasts that resolve convection explicitly but can only afford a limited ensemble size of less than 100 members. Given that the degrees of freedom of atmospheric models are several magnitudes higher implies sampling errors. Sampling errors and fast error growth on convective scales, lead to a low predictability. Consequently, improving initial conditions and subsequent forecasts requires a better understanding of error correlations in both space and time.

For this purpose, we conducted the first convective-scale 1000-member ensemble simulation over central Europe. Several 1000-member ensemble forecasts are investigated during a high impact weather period in summer 2016 using ensemble sensitivity analysis. The spatial and spatiotemporal correlations are used to quantify sampling errors on convective scales. Correlations of the 1000-member ensemble forecast serve as truth to assess the performance of different localization approaches. Those approaches include a standard distance-based localization technique and a statistical sampling error correction method as proposed by Anderson (2012). Our study highlights advantages and disadvantages of existing methods and emphasises the need of different localization approaches for different scales and variables. Several results are published in Necker et al (2020a) and (2020b).

Outline

1) 1000-member ensemble

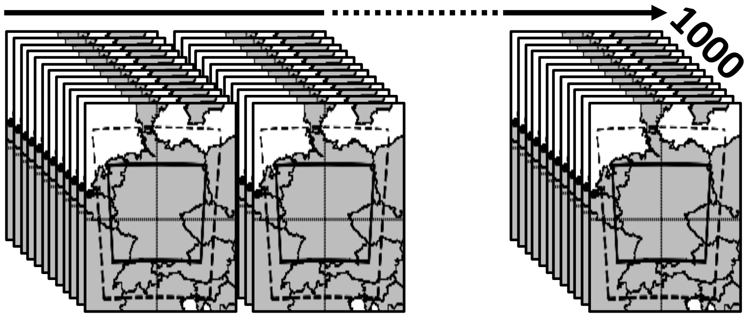
Why large ensemble?

Comparison to COSMO-DE

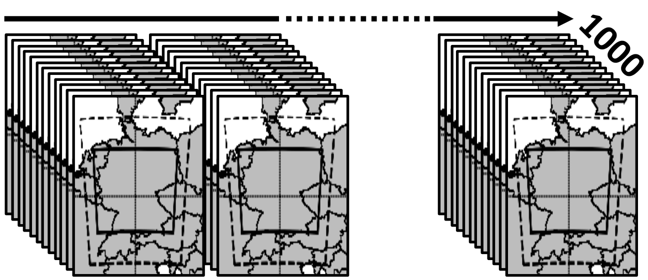
2) Sampling error correction (SEC)

Quantification of sampling errors

Localization in data assimilation

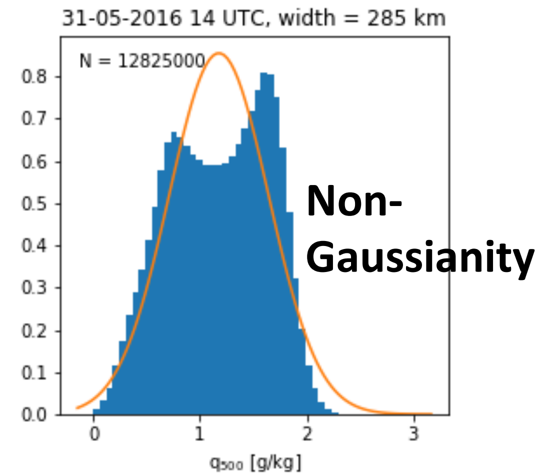


Motivation – Why 1000-member ensemble?

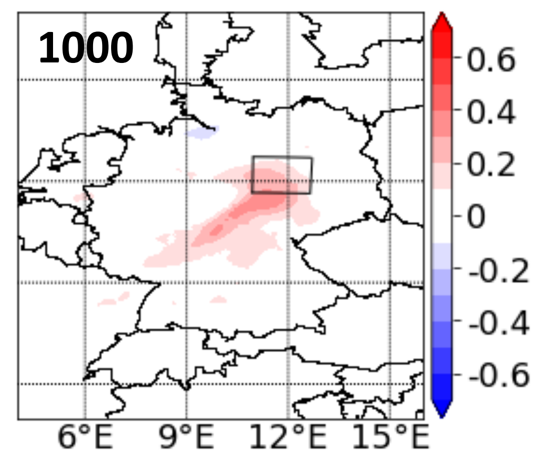


Realistic distributions

Realistic correlation



$$\hat{r} = \frac{cov(\mathbf{J},\mathbf{x})}{\sqrt{var(\mathbf{J})var(\mathbf{x})}}$$

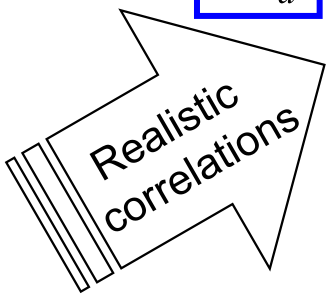


Plot: Sample correlation of 3-h precipitation forecast inside the box to initial 850 hPa humidity

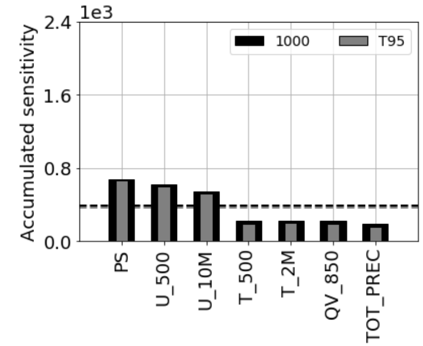
Motivation – 1000-member ensemble

Sensitivity of forecast metric J to analysis x_a

$$\boxed{\frac{\partial J}{\partial \mathbf{x}_a}}$$

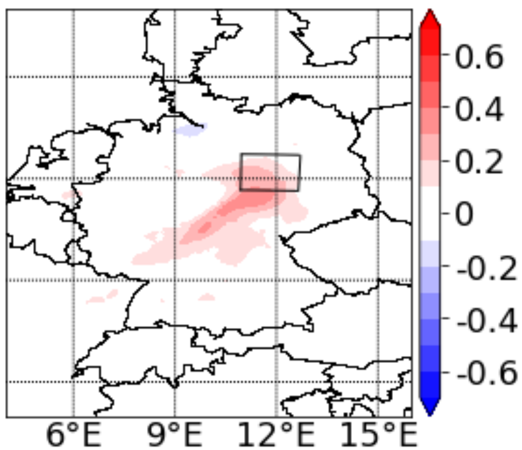


Observing strategies

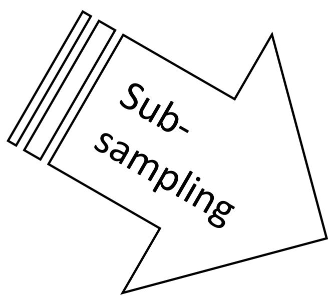


***Goal**: Potential impact of observations*

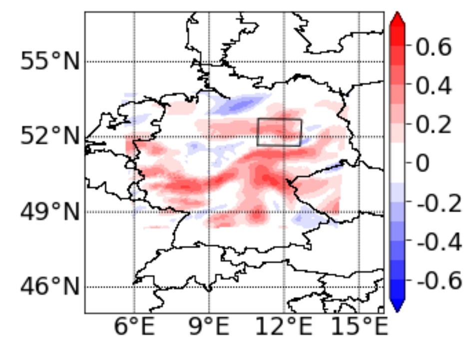
1000 members



Correlation of 3-h precipitation forecast inside the box to initial 850 hPa humidity



40 members



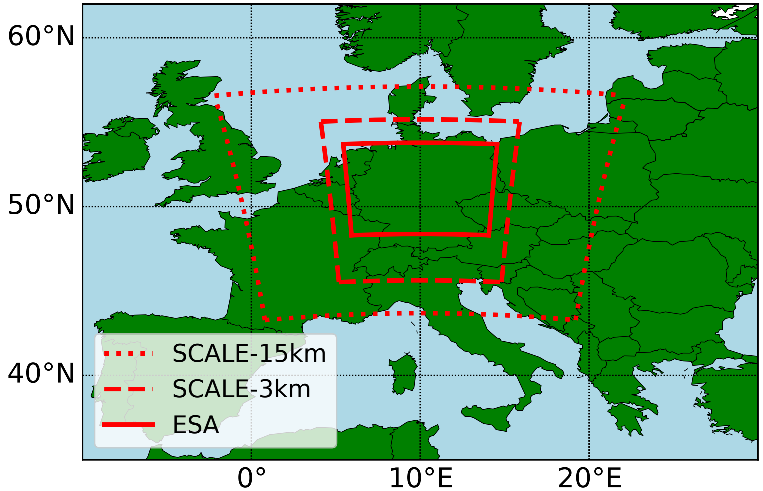
***Goal**: Quantifiaction of sampling errors*

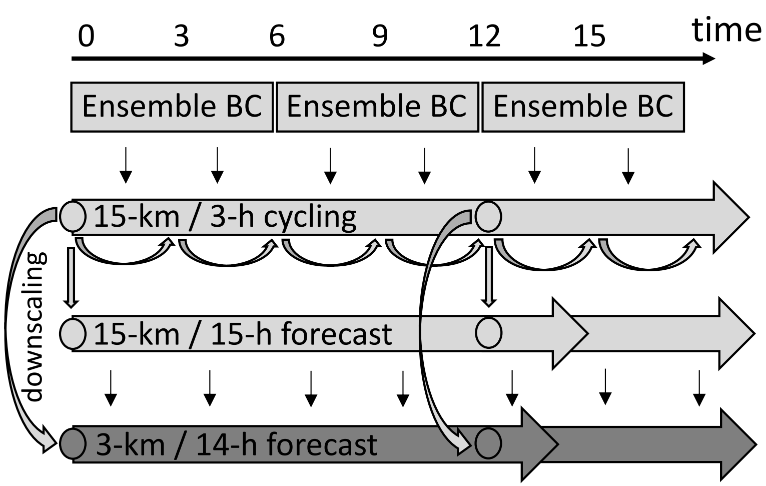
Convective-scale 1000-member ensemble

Japanese "SCALE-RM" model

- **Spin up**: 1 week
- **Period** of 5 days/10 FCs in Mai/June 2016
- Global **GFS ensemble BC** using NCEP 20-member analysis ensemble combined with 1000 random perturbations

- **LETKF** (15km; conventional observations)
- **Downscaling** to from 15km to 3km to initialize convective-scale forecasts
- **Forecast domain:** 3 km grid spacing, 350x250 grid points with 30 levels





***Necker et al 2020a:** A convective-scale 1000-member ensemble and potential applications. Q. J. R. Meteorol. Soc.*

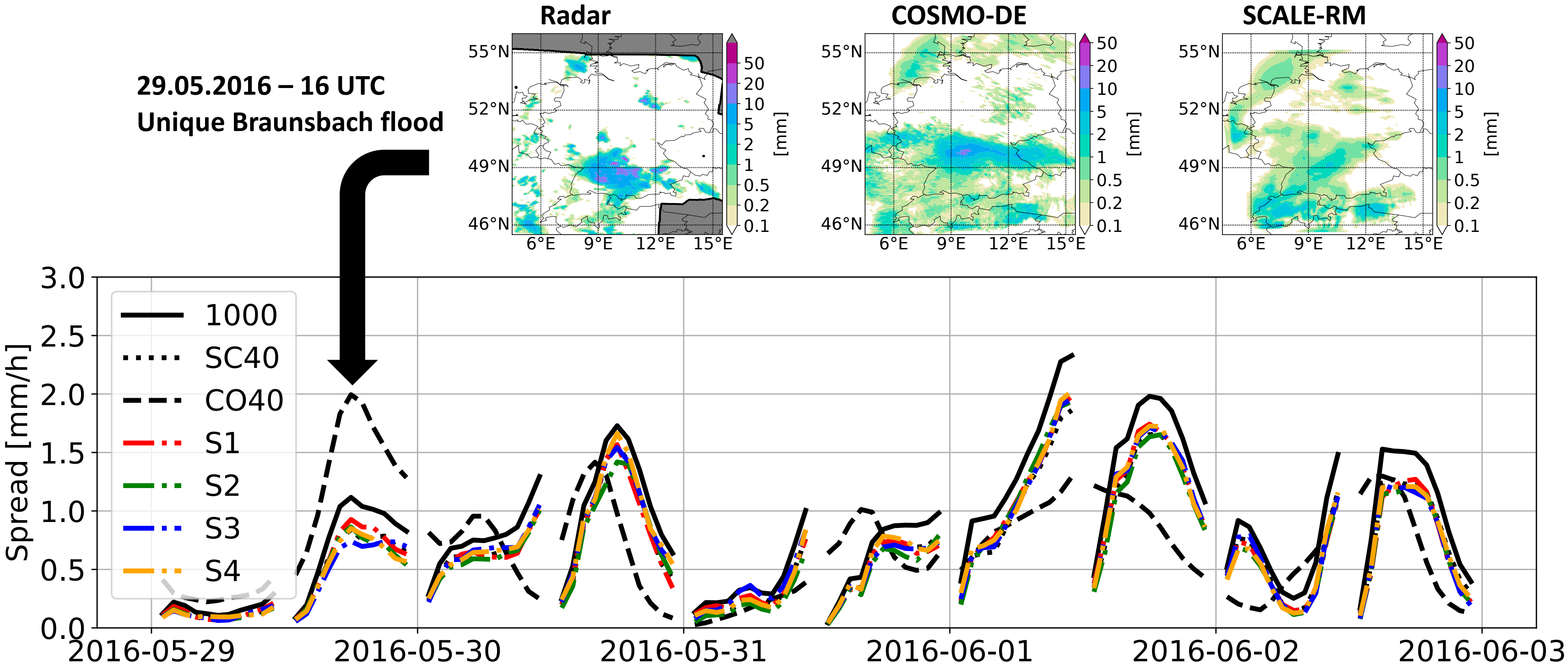
universität wien

Comparison of 1000-member ensemble to COSMO-DE –

Ensemble spread of precipitation

- Realistic evolution and amplitude
- Similar results for other variables

1000 – SCALE 1000-member ensemble
CO40 – COSMO 40-member ensemble
SC40, S1, S2, S3, S4 - different 40-member subsamples of the 1000-member ensemble.

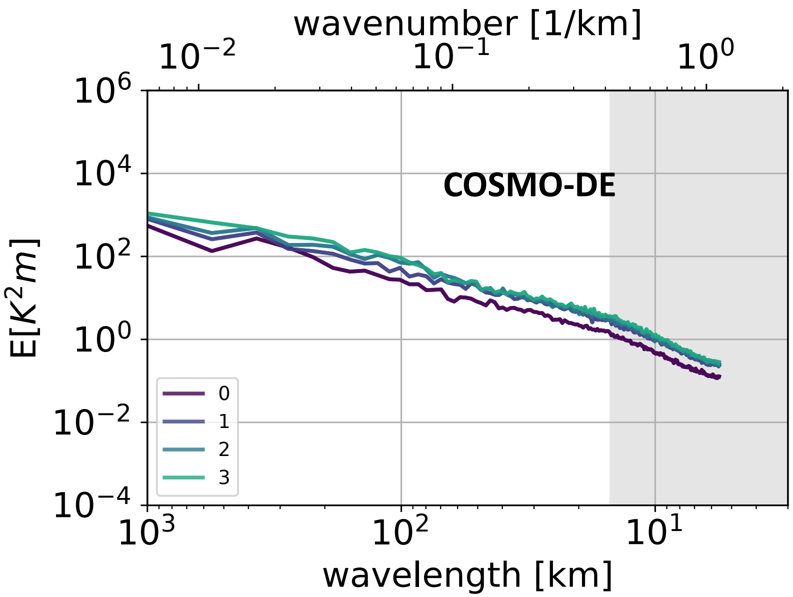


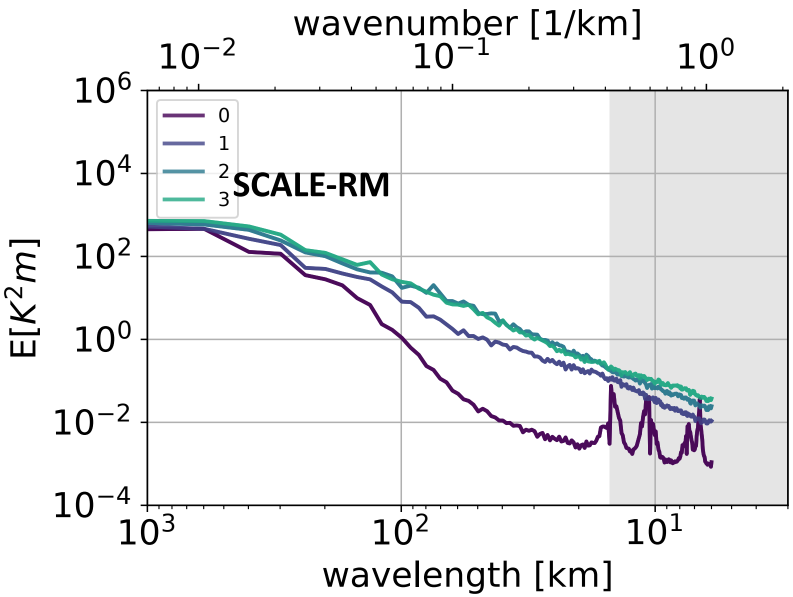
Comparison to COSMO-DE – Spectral analysis

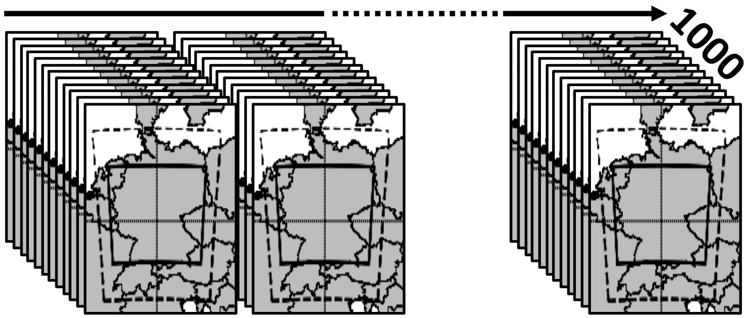
Variance spectra of temperature at 700 hPa

- Realistic spread properties after spin-up
- 1-2h spin-up that originates from the downscaled IC used by the SCALE-RM 1000-member ensemble

-> SCALE simulation is found to be realistic and comparable to COSMO







2) Sampling Error Correction (SEC)

Quantification of sampling errors

Localization in data assimilation

Sampling error correction (SEC)

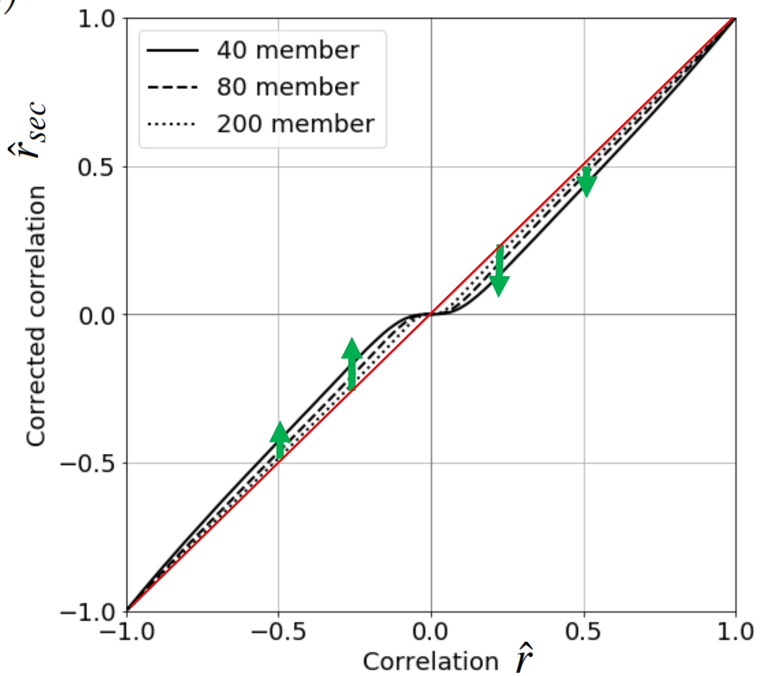
Ensemble sensitivity analysis / sample correlation:

$$S = \frac{\partial \mathbf{J}}{\partial \mathbf{x}} = \hat{r}\frac{\sigma_J}{\sigma_x} \qquad \boxed{\hat{r}} = \frac{cov(\mathbf{J},\mathbf{x})}{\sqrt{var(\mathbf{J})var(\mathbf{x})}}$$

S: sensitvity
J : response function
x : state variable
r : sample correlation (-1,1)
r_{sec}: sampling error corrected correlation
σ : sample standard deviation

Sampling Error Correction *(SEC)*:

- Designed to replace/reduce need of localization
- Offline Monte-Carlo technique -> look-up table
- r_{sec} table depends on ensemble size, sample correlation and assumed prior (normal) correlation distribution



-> Samling error corrected sensitivity:

$$S_{sec} = \boxed{\hat{r}_{sec}}\frac{\sigma_J}{\sigma_x}$$

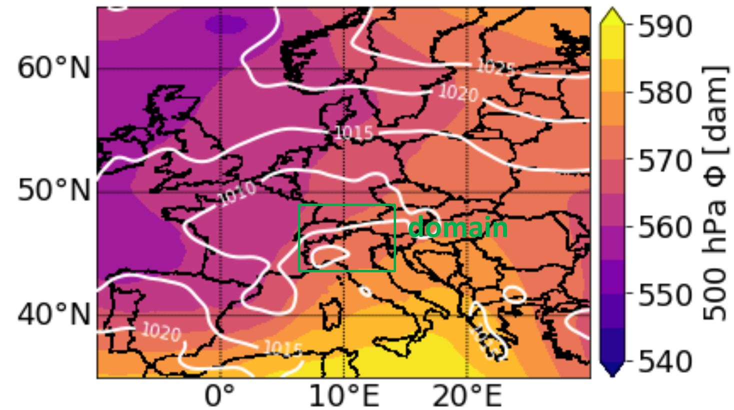
Anderson (2012): *Localization and Sampling Error Correction in EnKF DA*

Example of spatiotemporal correlations

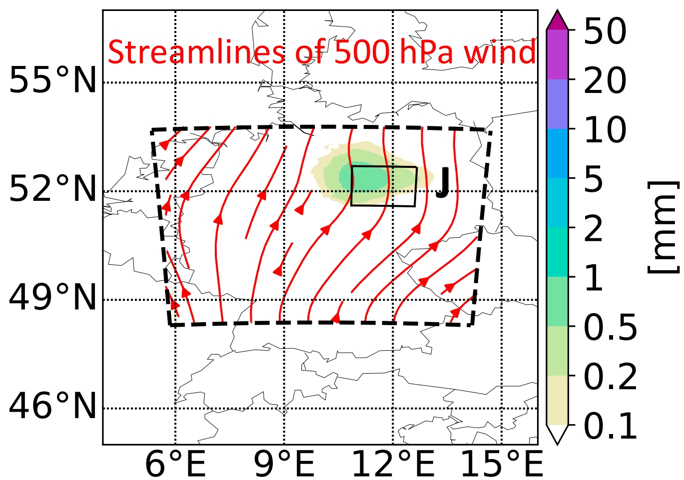
ESA setup: (3-h lead time forecast)

- Response function (black square): Precipitation coarse grained over 40x40 grid points

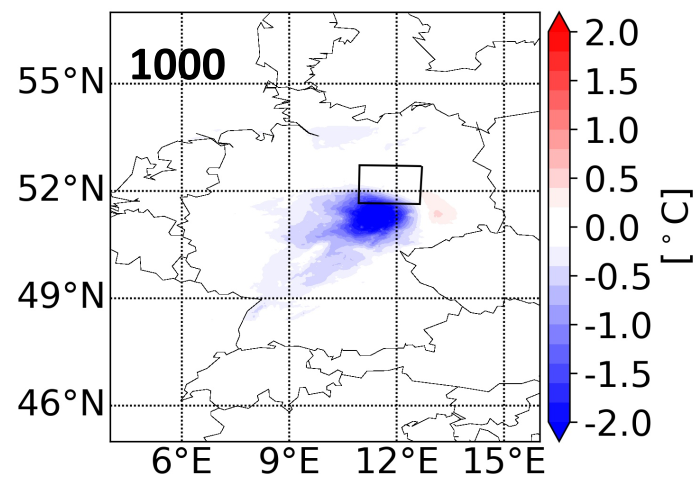
ECMWF analysis, 29.05.2016 – 0 UTC



1000-member ensemble mean precipitation



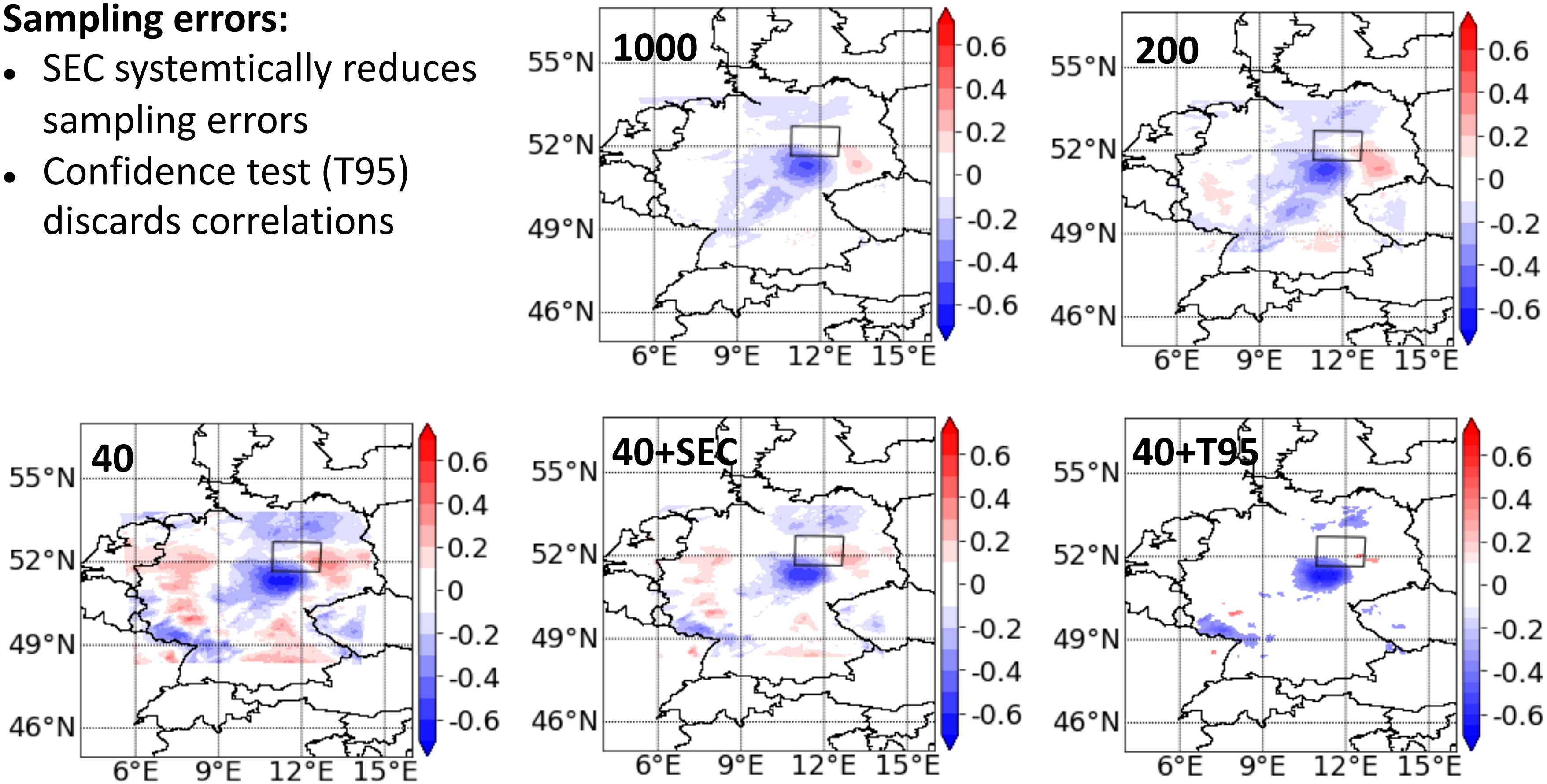
Correlation of precipitation forecast to 2-m temperature field



Sampling error – Qualitative analysis of spatiotemporal correlations

Sampling errors:

- SEC systemtically reduces sampling errors
- Confidence test (T95) discards correlations



Correlation of 3-h precipitation forecast to initial 2-m temperature, 1 forecast

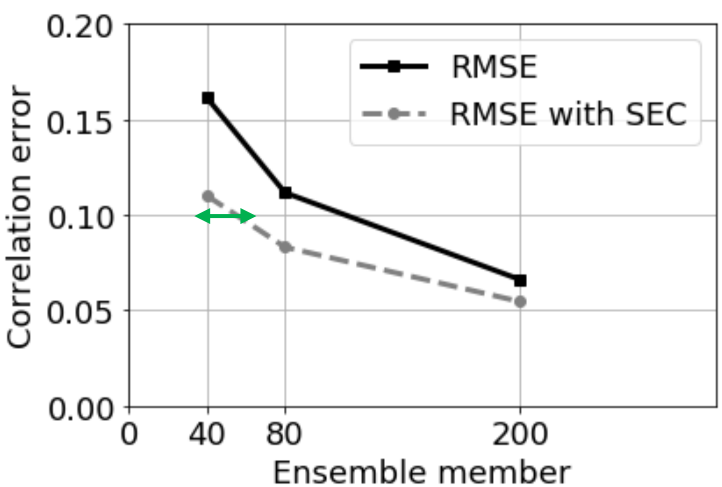
Sampling error as function of ensemble size evaluating spatiotemporal correlations

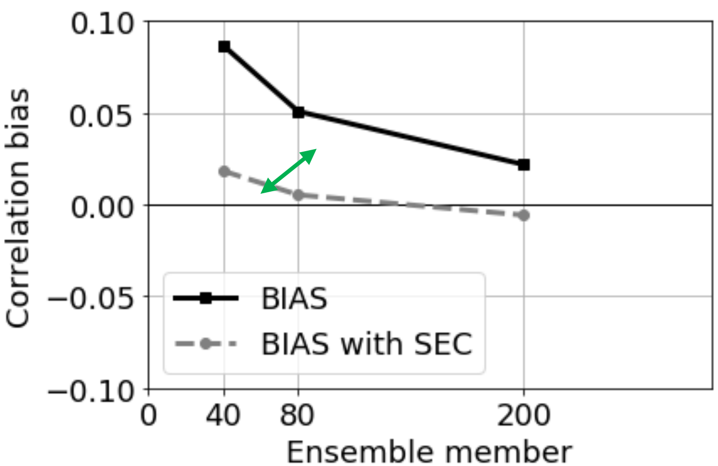
Sampling errors:

- Doubling the ensemble size from 40 to 80 member decreases sampling error by 30%
- SEC significantly reduces sampling errors for all investigated ensemble sizes
- 40 member + SEC performs better then 80 member

$$RMSE_{40} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(r_{40,n} - r_{1000,n}\right)^2}$$

$$BIAS_{40} = \frac{1}{N}\left(\sum_{n=1}^{N}|r_{40,n}| - \sum_{n=1}^{N}|r_{1000,n}|\right)$$





Correlation of 3-h precipitation forecast to initial 2-m temperature, 10 forecasts

Spatial correlations for data assimilation

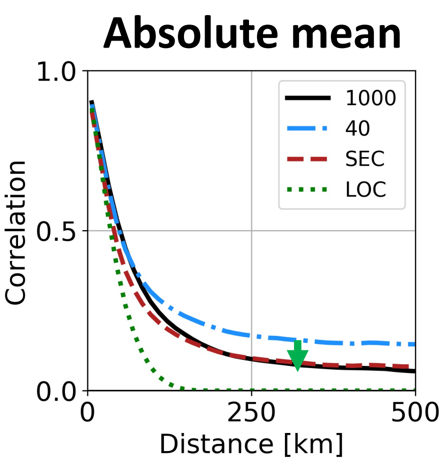
Comparison SEC vs LOC:

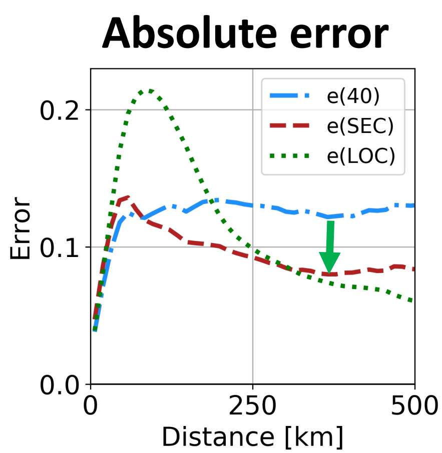
- SEC outperforms LOC on short distances
- LOC best for long-range correlations >250km
- Combination of both approaches seems most beneficial

LOC – Gaspari Cohn localization

Correlation

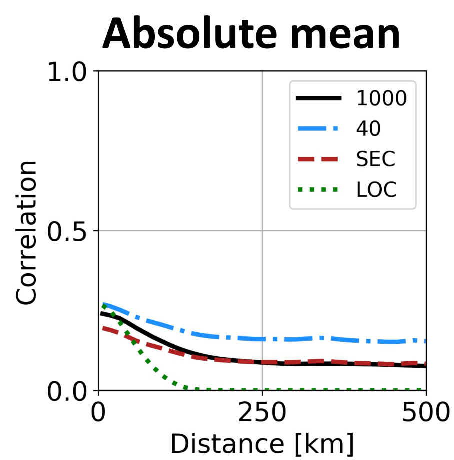
T2m to T2m

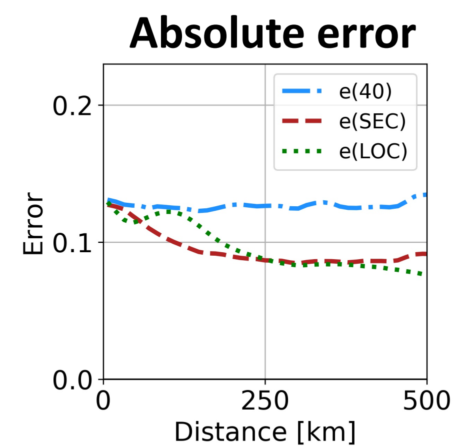




Cross - Correlation

T2m to U10M





universität wien

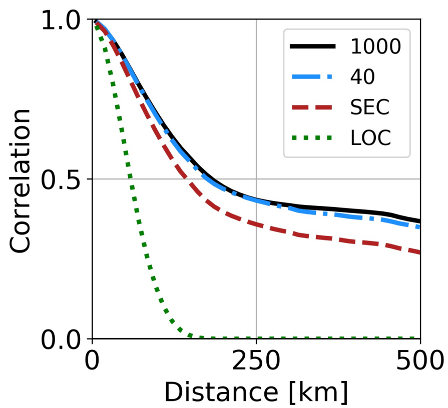
Spatial correlations for data assimilation

Spatial correlations as function of horizontal distance:

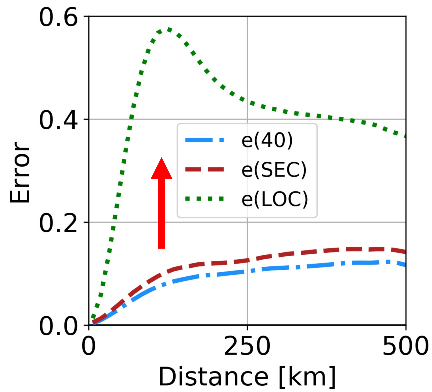
- Localization can degrade the performace in particular cases
- For highly positivly correlated variables SEC and LOC are not suitable
  → no improvements due to insufficient uniform prior or to narrow localization radius

Correlation T 500hPa to T 500hPa

Absolute mean



Absolute error



Possible solutions:

→ Use different localization radii for different variables

→ Use different or adaptive prior computing a SEC (see next slide)

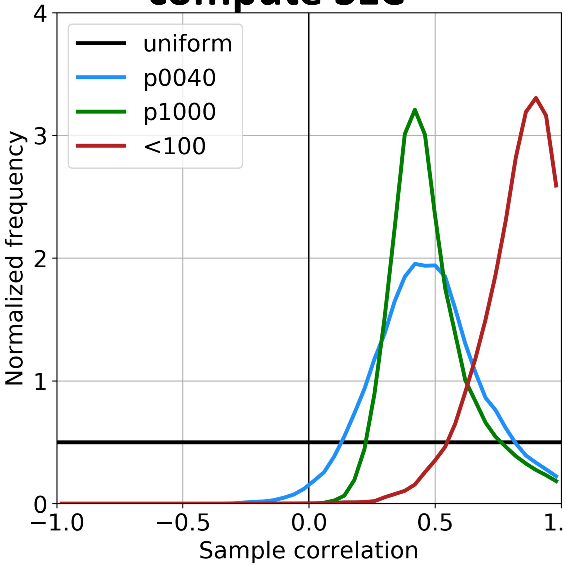
Sampling error correction – Different prior assumption

Conclusion:

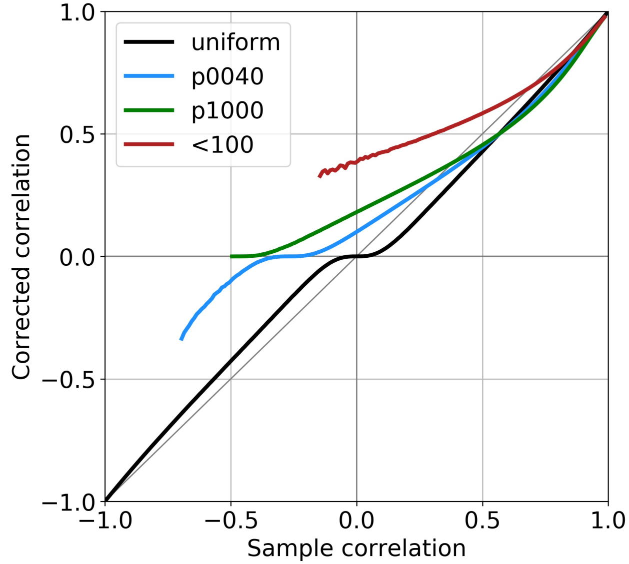
- Suitable prior assumptions can improve SEC
- Distance depended prior performs best for short distances

T 500hPa   to T 500hPa

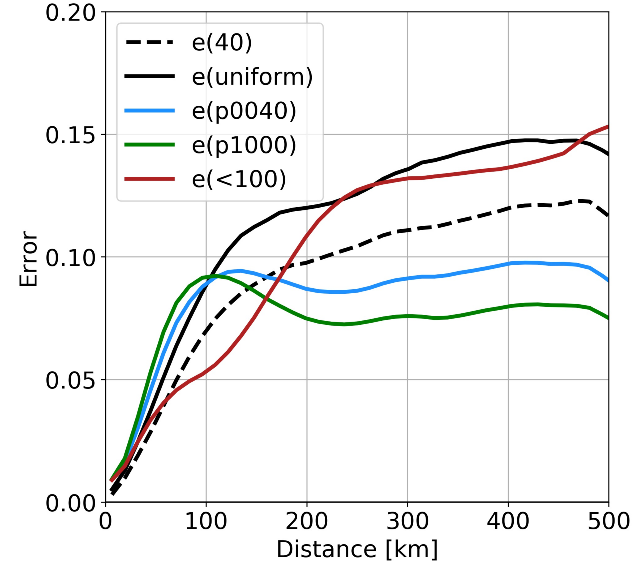
Prior distributions used to compute SEC



Resulting SEC

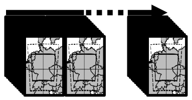


Absolute error

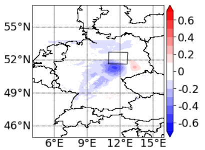


Necker et al 2020b: *Sampling error correction evaluated using a convective-scale 1000-member ensemble. Mon. Wea. Rev.*

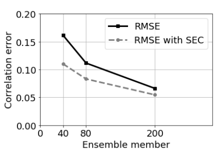
Conclusions



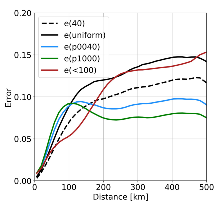
➢ 1000-member ensemble provides unique dataset for various studies on: Potential impact of observations, non-Gaussianity, sampling errors or localizaion in data assimiation



➢ **This talk presented temporal and spatial correlations** obtained for the **convective-scale 1000-member ensemble simulation** over Europe



➢ **Sampling error correction (SEC for spatiotemporal correlations):**
- Significantly reduced sampling errors
- Simple prior assumption is suitable



➢ **Sampling error correction (SEC for DA/ spatial correlations)**:
- Promissing especially for convective-scale and vertical application
- Different priors required for better performance

➢ **For more details please see our journal publications**

universität wien

References

Ancell and Hakim, 2007: Comparing Adjoint- and Ensemble-Sensitivity Analysis with Applications to Observation Targeting. *Mon. Wea. Rev.,* 135, 4117-4134

Torn, R. D., 2010: Ensemble-Based Sensitivity Analysis Applied to African Easterly Waves. *Weather and Forecasting.*

Anderson, J. L. 2012: Localization and Sampling Error Correction in Ensemble Kalman Filter Data Assimilation. *Mon. Wea. Rev*

Anderson, J. L., 2016: Reducing Correlation Sampling Error in Ensemble Kalman Filter Data. Assimilation. Mon. Wea. Rev.

Necker et al 2020a : A convective-scale 1000-member ensemble and potential applications. *Q. J. R. Meteorol. Soc.*

Necker et al 2020b: Sampling error correction evaluated using a convective-scale 1000-member ensemble. *Mon. Wea. Rev.*