PDAF Parallel Data Assimilation Framework

ALFRED-WEGENER-INSTITUT
HELMHOLTZ-ZENTRUM FÜR POLAR-
UND MEERESFORSCHUNG

Hybrid nonlinear-Kalman ensemble transform filtering for data assimilation in systems with different degrees of nonlinearity

Lars Nerger

Alfred Wegener Institute
Helmholtz Center for Polar and Marine Research
Bremerhaven, Germany

7th International Workshop on Nonhydrostatic Models (NHM-WS 2025), Morioka, Japan, Nov. 17, 2025

Motivation

Data Assimilation (DA) into non-linear models (e.g. nonhydrostatic dynamics)

- Nonlinearity in DA → non-Gaussianity of error distributions
  - → Non-linear model dynamics lead to different degree of non-Gaussianity
- For Gaussian distributions: (Ensemble) Kalman filter is optimal
- For non-Gaussian distributions:
  - → Kalman filter suboptimal or failing
  - → Non-linear DA methods (e.g. Particle filters): possibly better estimates, but higher sampling errors than KFs

→ Aim for DA method that

- adapts to non-Gaussianity
- Allows to utilize optimality of Kalman filter for Gaussian distributions
  - → **Utilize hybrid combination of Kalman and nonlinear filters**

Linear and Nonlinear Ensemble Filters

- Represent state and its error by ensemble $\mathbf{X}$ of N states
- **Forecast**:
  - Integrate ensemble with numerical model
- **Analysis step:**
  - update ensemble mean $\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{X}'^f \tilde{\mathbf{w}}$
  - update ensemble perturbations $\mathbf{X}'^a = \mathbf{X}'^f \mathbf{W}$

  (both can be combined in a single step)
- Ensemble Kalman & nonlinear filters: Different definitions of
  - weight vector $\tilde{\mathbf{w}}$ (dimension N)
  - Transform matrix $\mathbf{W}$ (dimension $N \times N$)

ETKF (Bishop et al., 2001)

- **Ensemble Transform Kalman filter**
  - Assume Gaussian distributions
  - Transform matrix

$$\mathbf{A}^{-1} = (N-1)\mathbf{I} + (\mathbf{H}\mathbf{X}'^f)^T\mathbf{R}^{-1}\mathbf{H}\mathbf{X}'^f$$

  - Mean update weight vector

$$\tilde{\mathbf{w}} = \mathbf{A}(\mathbf{H}\mathbf{X}'^f)^T\mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}^f}\right)$$

(depends linearly on $\mathbf{y}$)

  - Transformation of ensemble perturbations

$$\mathbf{W} = \sqrt{N-1}\;\mathbf{A}^{1/2}\mathbf{\Lambda}$$

$\mathbf{\Lambda}$: mean-preserving random matrix or identity

($\mathbf{W}$ depends only on $\mathbf{R}$, not $\mathbf{y}$)

Note: $\mathbf{W}$ depends only on $\mathbf{R}$, not on $\mathbf{y}$

NETF (Tödter & Ahrens, 2015)

- **Nonlinear Ensemble Transform Filter**
  - Mean update from Particle Filter weights:
    for Gaussian observation errors for all particles *i*

$$\tilde{w}^i \sim \exp\left(-0.5(\mathbf{y}-\mathbf{H}\mathbf{x}_i^f)^T\mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}\mathbf{x}_i^f)\right)$$

(nonlinear function of observations $\mathbf{y}$)

  - Ensemble update
    - Transform ensemble to fulfill analysis covariance (like ETKF, but not assuming Gaussianity)
    - Derivation gives

$$\mathbf{W} = \sqrt{N}\left[\mathrm{diag}(\tilde{\mathbf{w}}) - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^T\right]^{1/2}\Lambda$$

($\mathbf{\Lambda}$: mean-preserving random matrix; useful for stability)

Note: $\mathbf{W}$ depends on $\mathbf{y}$ and $\mathbf{R}$ (higher sampling errors than ETKF)

NETF: second-order exact particle filter using transformation rather than resampling

Tödter, J. and Ahrens, B. (2015) *Mon. Wea. Rev.* **143**,1347–1367

ETKF-NETF – Hybrid Filter Variants

Factorize the likelihood: $p(\mathbf{y}|\mathbf{x}) = p(\mathbf{y}|\mathbf{x})^{\gamma} p(\mathbf{y}|\mathbf{x})^{(1-\gamma)}$

1-step update (*HSync*)

$$\mathbf{X}^a_{HSync} = \overline{\mathbf{X}}^f + (1-\gamma)\Delta\mathbf{X}_{NETF} + \gamma\Delta\mathbf{X}_{ETKF}$$

- $\Delta\mathbf{X}$: assimilation increment of a filter
- γ: hybrid weight (between 0 and 1; 1 for fully ETKF)

2-step updates

Variant 1 (*HNK*): NETF followed by ETKF

$$\tilde{\mathbf{X}}^a_{HNK} = \mathbf{X}^a_{NETF}[\mathbf{X}^f, (1-\gamma)\mathbf{R}^{-1}]$$

$$\mathbf{X}^a_{HNK} = \mathbf{X}^a_{ETKF}[\tilde{\mathbf{X}}^a_{HNK}, \gamma\mathbf{R}^{-1}]$$

- Both steps computed with increased $\mathbf{R}$ according to γ

Variant 2 (*HKN*): ETKF followed by NETF

Choosing hybrid weight γ

- Hybrid weight shifts filter behavior

Some possibilities:

- Fixed value
- Adaptive - According to which condition?
  - Frei & Kuensch (2013) suggested using effective sample size $N_{eff} = \sum \frac{1}{(w^i)^2}$ (Usual choice for 'tempering')
    - γ_α : Choose γ so that N_{eff} is as small as possible but above minimum limit α (done iteratively)

Issue: Using N_{eff}

- only ensures non-collapsing ensemble
- does not ensure good analysis result

Using skewness and kurtosis to define hybrid weight γ

- Sampling errors are larger in NETF than ETKF
  - → Always use ETKF for Gaussian (linear) cases
- Skewness and kurtosis describe deviation from Gaussianity
- mean absolute skewness (mas) and kurtosis (mak) of observed ensemble (with localization: use locally assimilated observations)
- Use normalized means:

$$nmas = \frac{1}{\sqrt{\kappa}} mas \qquad nmak = \frac{1}{\kappa} mak$$

standard value:
$\kappa = N_e$

Now define

$$\gamma_{sk,\alpha} = \max\left[\min(1 - nmak,\ 1 - nmas),\ \gamma_\alpha\right]$$

stronger influence of $nmas$ and $nmak$ limited by N_{eff}

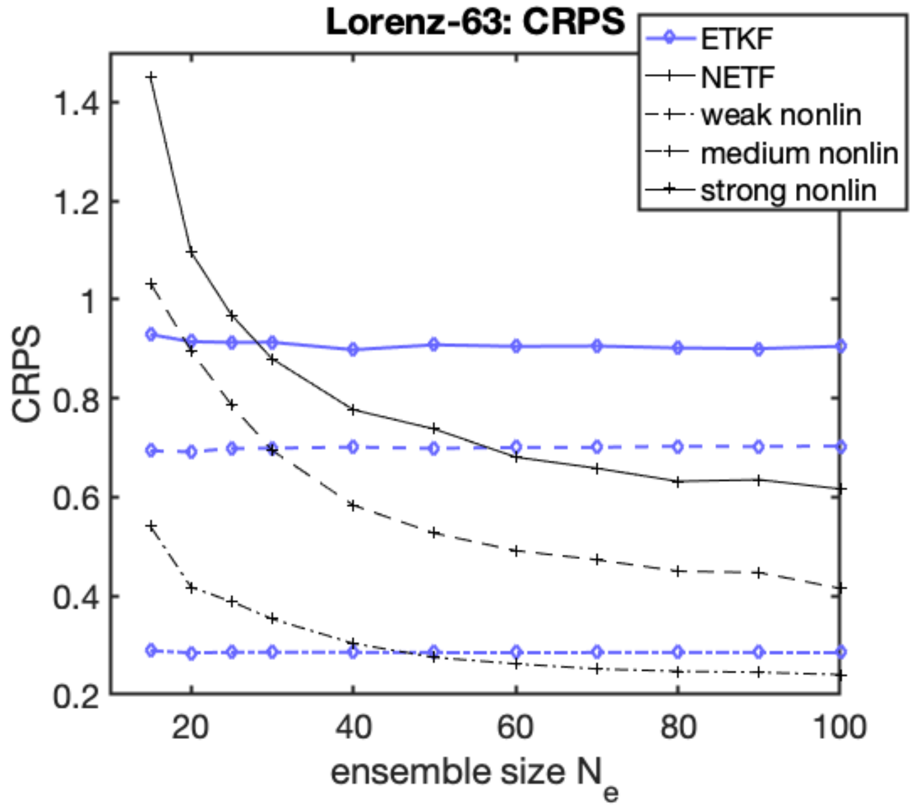
Note: There are sampling errors, e.g. for skewness $\sigma_{skew} \sim \sqrt{6/N_e}$

→ For N_e=25: ~10% error in γ

Numerical experiments

- Lorenz-63 and Lorenz-96 models
- Configurations that ensure high nonlinearity
  - long forecast times
  - Lorenz-96 also incomplete observations
- Assimilation implemented using PDAF (https://pdaf.awi.de)

- NETF yields smaller errors than ETKF if ensemble size large enough
  - → Size decreases for larger nonlinearity
  - → Improvement by NETF stronger for higher nonlinearity



ETKF and NETF for 3 different nonlinearities

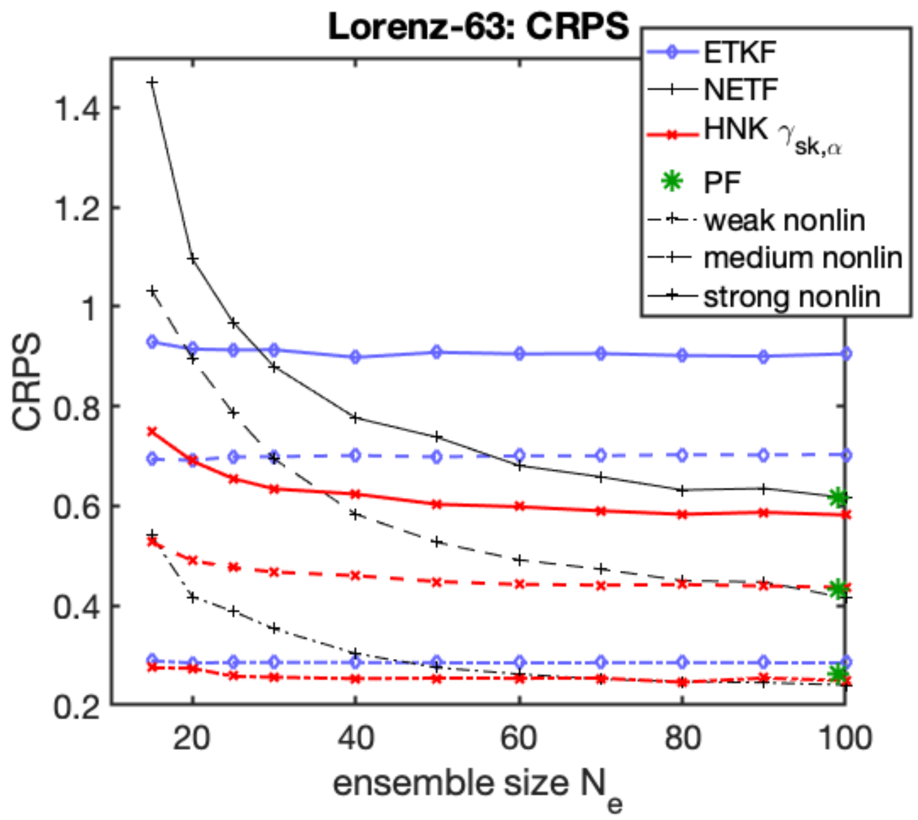
(weak Δt=0.1, medium Δt=0.4, strong Δt=0.7)

PDAF Parallel Data Assimilation Framework

AWI

Assimilation with Lorenz-63 model

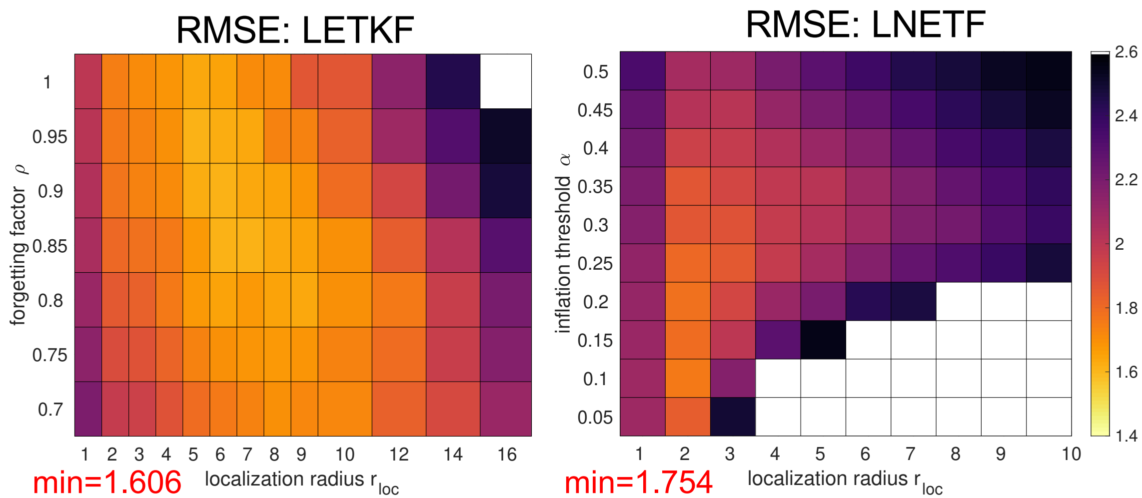
- **Hybrid filter HNK**
  - particular strong effect for small N_e
  - CRPS from NETF and HNK converge for large N_e
  - errors reduced up to 28%
- **Particle Filter** $*$
  - comparable CRPS for large N_e
    - PF expected to be superior if N_e sufficiently large (the full nonlinear filter)

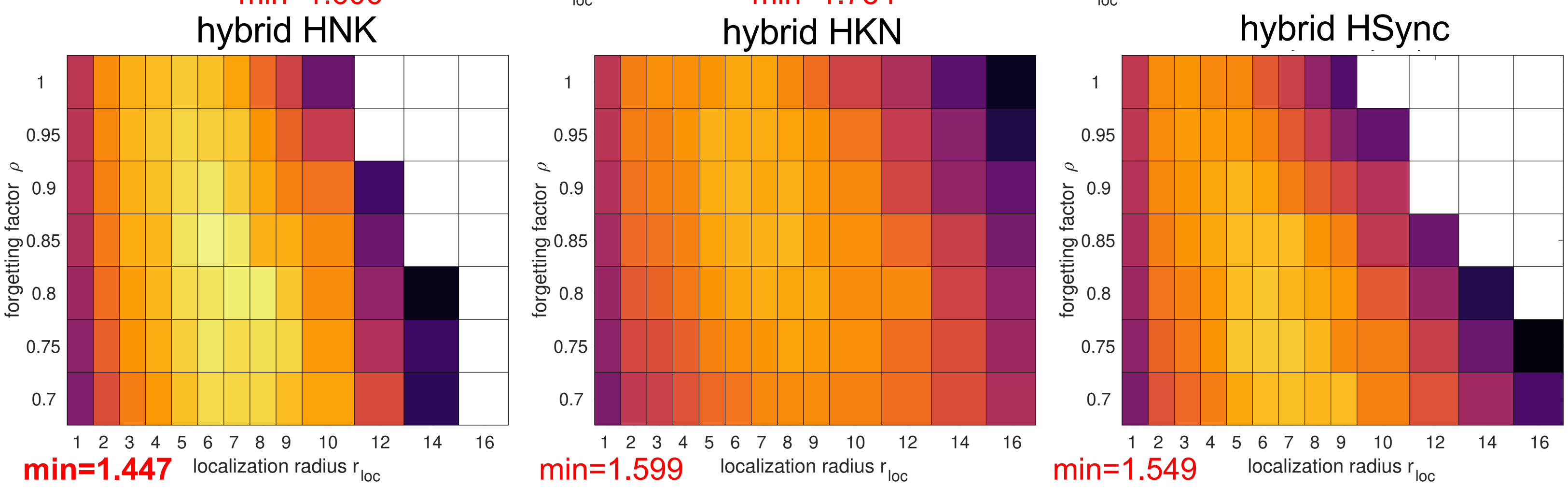


- Note: Easy to get large ensemble for Lorenz-63, difficult for higher dimensional models

Test with Lorenz-96 model

Ensemble size 15; Forecast length: 8 time steps; 20 observations





- Show RMS errors as function of inflation (forgetting factor or α) and localization radius
- LNETF worse than LETKF
- All hybrid variants improve the state estimate
- *Smallest errors:* Hybrid HNK (10% error reduction)

Summary

PDAF Parallel Data Assimilation Framework

Hybrid nonlinear-Kalman ensemble transform filter (LKNETF)

- Combine LETKF and LNETF methods
- hybrid weight γ shifts between both methods

Experiments with Lorenz models

- Hybrid filter reduces errors compared to both LETKF and LNETF
- Best results for particle filter applied before LETKF
- Compute γ based on skewness and kurtosis
  → allows to control nonlinearity of filter based on non-Gaussianity

Ongoing work

- more applications to complex models to understand performance (Tests: ocean 1/4°: little effect; marine biogeochemistry: large effect; PM2.5 components modeling: large effect)
- mathematical consistent description for hybrid weight