Data assimilation on convective scale based on first physical principles

Dr. habil. Tijana Janjić

Ludwig Maximilian University of Munich

Outline

- Convective scale data assimilation characteristics

Outline

- Convective scale data assimilation characteristics
- Our approach of addressing some of the challenges:

Outline

- Convective scale data assimilation characteristics
- Our approach of addressing some of the challenges: Ensemble data assimilation scheme with constraints

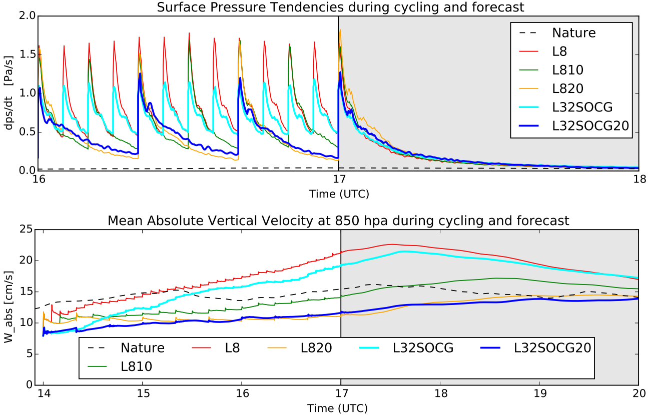
Outline

- Convective scale data assimilation characteristics
- Our approach of addressing some of the challenges: Ensemble data assimilation scheme with constraints
- Illustrate approach on simple examples

Convective scale data assimilation

- Data assimilation on convective scales needs to capture fast changing processes and many scales of motion that are resolved in high resolution models.
- Variables estimated need to be positive or in certain range.
- Rapid updates are essential (for example radar reflectivity, radial wind data assimilation 5-15 min). However, leading to problems of balance and noise.
- Background errors are non-Gaussian in nature (for example location error), model error consisting of large as well as unresolved scales and processes.
- Predictability of convective storms is couple of hours (Durran and Weyn 2016, Durran and Gingrich, 2014).

Problems with noise



Surface pressure tendency (upper). Absolute vertical velocity (bottom).

Lange, H., G. C. Craig, T. Janjić, 2017: Characterizing Noise and Spurious Convection in Convective Data Assimilation, Q. J. R. Meteorol. Soc.

Convective scale DA at DWD

- COSMO model (Baldauf et al. 2011) in the domain over Germany, 2.8km horizontal resolution, 50 hybrid levels. Deep convection explicit, shallow convection parametrized.

Convective scale DA at DWD

- COSMO model (Baldauf et al. 2011) in the domain over Germany, 2.8km horizontal resolution, 50 hybrid levels. Deep convection explicit, shallow convection parametrized.
- Kilometer-Scale Ensemble Data Assimilation (KENDA, Schraff et al. 2016) based on LETKF (Hunt et al. 2007)
- State consists of the prognostic variables of velocity, temperature, pressure perturbation, specific humidity, cloud water and ice.
- The prognostic variables of turbulent kinetic energy, rain, snow, and graupel are excluded from the analysis update.
- 1h updates

Convective scale DA at DWD

- COSMO model (Baldauf et al. 2011) in the domain over Germany, 2.8km horizontal resolution, 50 hybrid levels. Deep convection explicit, shallow convection parametrized.
- Kilometer-Scale Ensemble Data Assimilation (KENDA, Schraff et al. 2016) based on LETKF (Hunt et al. 2007)
- State consists of the prognostic variables of velocity, temperature, pressure perturbation, specific humidity, cloud water and ice.
- The prognostic variables of turbulent kinetic energy, rain, snow, and graupel are excluded from the analysis update.
- 1h updates
- **For radar data, LHN in all ensemble members**

Convective scale DA at DWD

1. 40-member ensemble
2. adaptive localization in horisontal, in vertical 0.075–0.5 in ln p
3. adaptive inflation
4. RTPP scheme with 0.75 (Zhang et al 2004)
5. Additive noise with samples from ICON's B matrix.

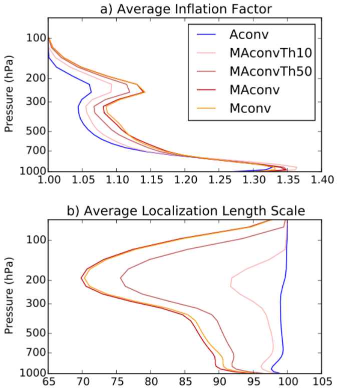
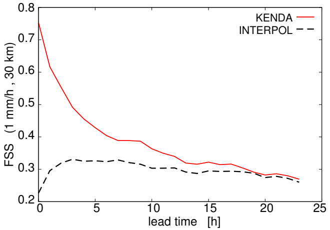


Figure from
Lange and Janjic (2016), MWR

Downscaling versus convective-scale DA

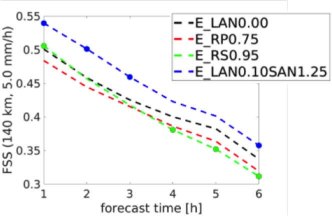


The fraction skill score corresponding to areas of 30 km $\times$ 30 km for 1.0 mm h^{-1} one hour precipitation as a function of forecast lead time for a **convective two-week period from 26 May to 9 June 2016** .

Gustafsson et al. 2018, Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres, QJRMS

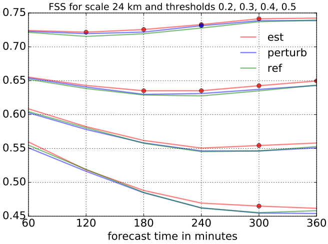
Model error

1. Additive noise Insufficient model resolution is one source of model error. Samples based on differences between 1.4km and 2.8km COSMO-DE runs. Weakly forced case, June 2016.



Zeng et al. 2018, JAMES; Zeng et al. 2019, JAMES

2. Boundary layer uncertainty:
   1. Stochastic boundary layer scheme (Kober and Craig 2017)
   2. warm-bubble
   3. Parameter perturbations or estimation e.g. roughness length. Verification against VIS/NIR data.



Ruckstuhl and Janjic 2019

— Physical properties/Conservation laws —

Outline

- Numerical discretization schemes have a long history of incorporating the most important conservation properties of the continuous system in order to improve the prediction of the nonlinear flow.

Outline

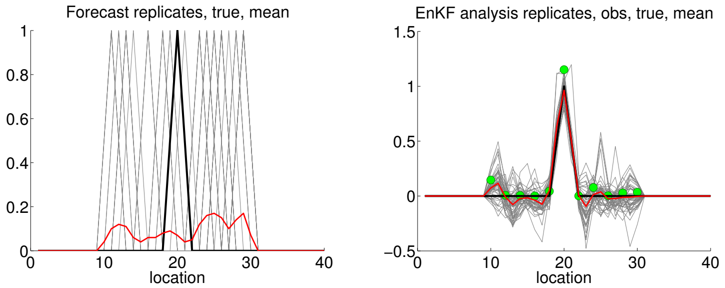
- Numerical discretization schemes have a long history of incorporating the most important conservation properties of the continuous system in order to improve the prediction of the nonlinear flow.
- The question arises, whether data assimilation algorithms should follow a similar approach?

Outline

- Numerical discretization schemes have a long history of incorporating the most important conservation properties of the continuous system in order to improve the prediction of the nonlinear flow.
- The question arises, whether data assimilation algorithms should follow a similar approach?

1. Explore which conservation properties are well recovered when using an ensemble Kalman filter
2. Include as constraints those that are not in data assimilation
3. Show implication on the prediction

Physical properties lost in the analysis step



The mean (red line) with background ensemble (left) and analysis ensemble obtained with EnKF algorithm (right). Observations (green) are the true state plus log normal noise.

Ensemble Kalman filter conserves total mass

- If no posterior adjustment are made (for example setting negative values to 0) and if

Ensemble Kalman filter conserves total mass

- If no posterior adjustment are made (for example setting negative values to 0) and if
- no adjustments are made to sample covariance, EnKF conserves mass.

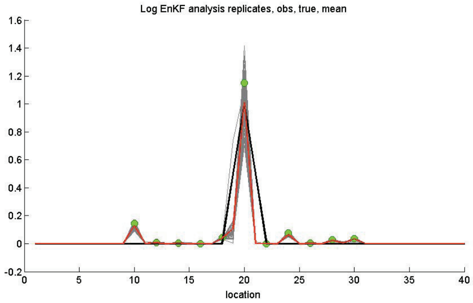
Ensemble Kalman filter conserves total mass

- If no posterior adjustment are made (for example setting negative values to 0) and if
- no adjustments are made to sample covariance, EnKF conserves mass.
  If $\mathbf{e}^T \mathbf{w}_k^{b,i} = M$ then the ensemble mean has also mass M, hence
  $\mathbf{e}^T(\mathbf{w}_k^{b,i} - \mathbf{w}_k^b) = 0$ for all i and $\mathbf{e}^T \mathbf{P}_k^b = 0$, that is $\mathbf{e}^T \mathbf{K}_k = 0$
  $\mathbf{e}^T \mathbf{w}_k^a = \mathbf{e}^T \mathbf{w}_k^b$ and

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^b (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T$$

  $\mathbf{e}^T \mathbf{P}_k^a = 0$
  $\mathbf{e}^T \mathbf{w}_k^{a,i} = \mathbf{e}^T \mathbf{w}_k^a$.

Physical properties lost in analysis step



The analysis mean (red line) and analysis ensemble obtained with log transformed EnKF algorithm. Analysis ensemble and the analysis are positive. Mass $1.42 < 2$ is not conserved.

Preserving physical properties

- Study conservation of mass, energy and enstrophy
- including dependence of the results on the observational type and localization radius

- Non-linear dynamics with 2D nonlinear shallow water model

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} - v\frac{\partial u}{\partial y} + fv - g\frac{\partial h}{\partial x}$$
$$\frac{\partial v}{\partial t} = -u\frac{\partial v}{\partial x} - v\frac{\partial v}{\partial y} - fu - g\frac{\partial h}{\partial y}$$
$$\frac{\partial h}{\partial t} = -\frac{\partial}{\partial x}(hu) - \frac{\partial}{\partial y}(hv)$$

h is the free surface height,
u is the zonal wind, v is the meridional wind,
g is the gravity acceleration and f is the Coriolis parameter.

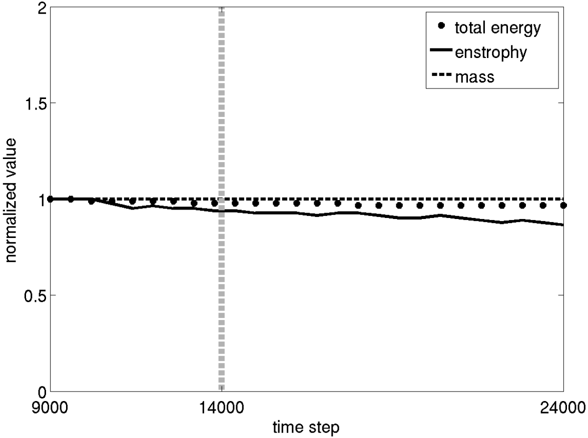
Preserving physical properties

- Numerical discretization of the dynamics is such that mass, energy and momentum are conserved and enstrophy for non divergent flow.

  We will consider changes due to data assimilation in

- Mass $M = \int\int h(x,y)dxdy$,
- Total energy
  $$E = \frac{1}{2}\int\int h(x,y)\left[u(x,y)^2 + v(x,y)^2\right] + g\left[h(x,y) - h_0\right]^2 dxdy$$
- Enstrophy
  $$\mathcal{E} = \int\int h(x,y)\left[\frac{\partial v}{\partial x}(x,y) - \frac{\partial u}{\partial y}(x,y)\right]^2 dxdy.$$

Nonlinear shallow water model



Time evolution of mass, total energy and enstrophy, normalized with respective initial values, in a nature run.

LETKF experiments

- different localization and the observational coverage
- 32 members + 1 deterministic run, constant inflation = 1.05
- 50 assimilation cycles
- Observations, u, v and h, or u and v, or h only from nature run
- Linear observation operator
- Gaussian observation error with standard deviations of 1.5m/s and 50 m.
- 1h updates

Diagnostics for analysis (ensemble mean)

1. RMSE
2. Normalized divergence
3. Noise (e.g. Janjic et al. 2011)

$$\mathcal{N} = \frac{\sum_{i,j=1}^{N_x,N_y} [\nabla^2 u(i,j)]^2 + [\nabla^2 v(i,j)]^2}{\sum_{i,j=1}^{N_x,N_y} [u(i,j)^2 + v(i,j)^2]}$$

Diagnostics for analysis (ensemble mean)

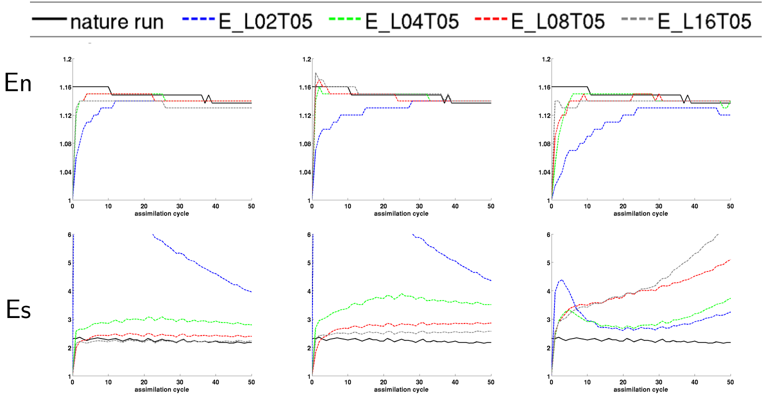
1. RMSE
2. Normalized divergence
3. Noise (e.g. Janjic et al. 2011)

$$\mathcal{N} = \frac{\sum_{i,j=1}^{N_x,N_y} [\nabla^2 u(i,j)]^2 + [\nabla^2 v(i,j)]^2}{\sum_{i,j=1}^{N_x,N_y} [u(i,j)^2 + v(i,j)^2]}$$

Relative to:

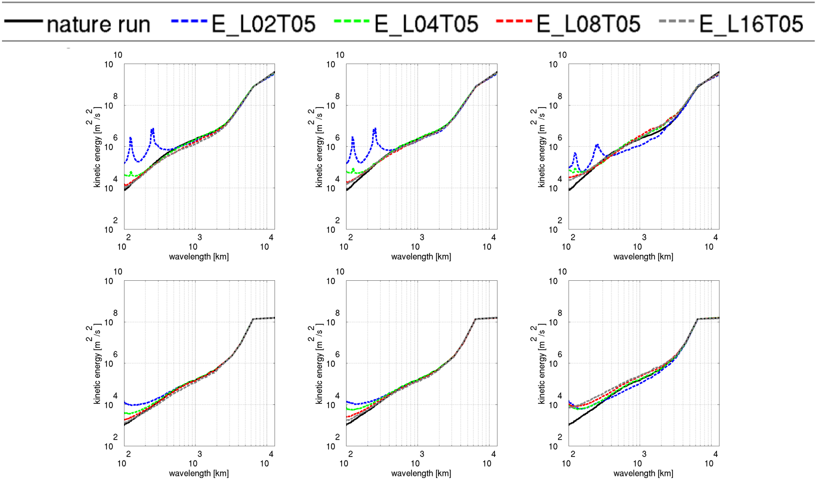
- nature run

Energy and Enstrophy



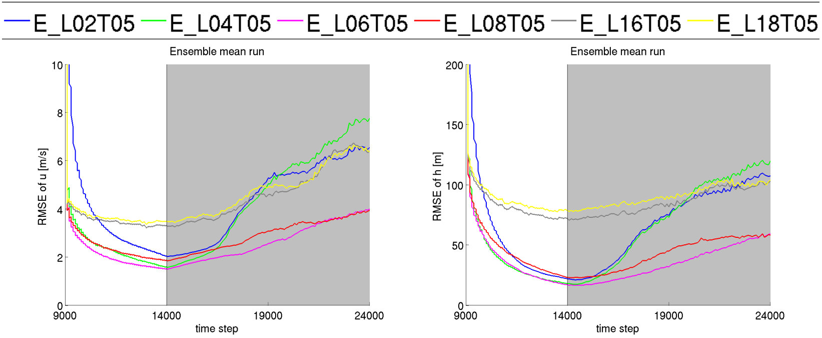
Obs u,v and h　　Obs u and v　　Obs h

Kinetic energy spectra



Averaged over the first (upper) and last five assimilation cycles (lower).

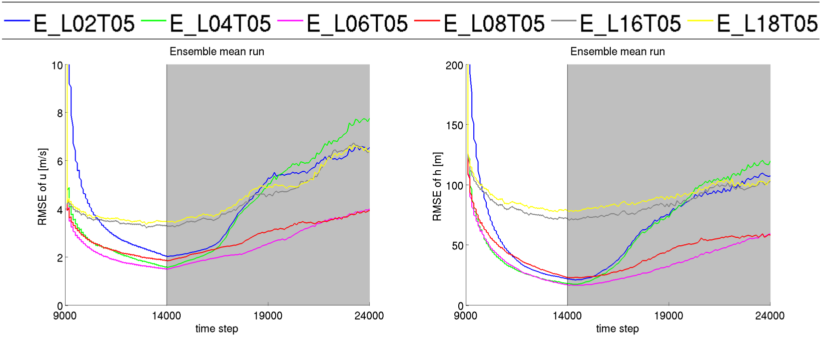
Prediction



RMSE for u

RMSE for h

Prediction



RMSE for u

RMSE for h

Y. Zeng and T. Janjic, 2016: Study of Conservation Laws with the Local Ensemble Transform Kalman Filter, Q. J. R. Meteorol. Soc.,142:699, 2359–2372.

— EnKF with constraints —

Janjic, T., D. McLaughlin, S. E. Cohn, M. Verlaan, 2014: Conservation of mass and preservation of positivity with ensemble-type Kalman filter algorithms, Mon. Wea. Rev., 142, No. 2, 755-773.

Zeng, Y., T. Janjić, Y. Ruckstuhl and M. Verlaan, 2017: Ensemble-type Kalman filter algorithm conserving mass, total energy and enstrophy, Q. J. R. Meteorol. Soc., 143:708, 2902–2914, doi:10.1002/qj.3142.

QPEns

Propagation step. Propagate the mean and the covariance with the dynamics between observations. Prior to new observation we have $\mathbf{w}_k^b$ and its covariance $\mathbf{P}_k^b$.

$$\mathbf{w}_k^{b,i} = \mathcal{M}\mathbf{w}_{k-1}^{a,i} + \mathbf{q}_k^i \quad i = 1, \dots N$$

$$\mathbf{P}_k^b = \frac{1}{N-1}\sum_{i=1}^{N}[\mathbf{w}_k^{b,i} - \mathbf{w}_k^b][\mathbf{w}_k^{b,i} - \mathbf{w}_k^b]^T.$$

Kalman analysis.

$$\mathbf{w}_k^{a,i} = \mathbf{w}_k^{b,i} + \mathbf{K}_k(\mathbf{w}_k^o + r^i - \mathbf{H}_k\mathbf{w}_k^{b,i}),$$
$$\mathbf{K}_k = \mathbf{P}_k^b\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_k^b\mathbf{H}_k^T + \mathbf{R}_k)^{-1}$$
$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T\mathbf{P}_k^b$$

Derived using $q^i \sim \mathcal{N}(0, \mathbf{Q})$, $r^i \sim \mathcal{N}(0, \mathbf{R})$, $\mathbf{w}_0^b \sim \mathcal{N}(0, \mathbf{P_0^b})$ and all uncorrelated.

QPEns algorithm

Inverse of ensemble derived background error covariance can be used to minimize the cost function to obtain the analysis

$$\mathbf{w}_k^{a,i} = \mathbf{w}_k^{b,i} + \arg\min_{\delta w^i} \frac{1}{2}[\delta\mathbf{w}^{i\,T}(\mathbf{P}_k^b)^{-1}\delta\mathbf{w}^i + \mathbf{f}^{i\,T}\mathbf{R}_k^{-1}\mathbf{f}^i]$$

subject to

$$\delta\mathbf{w}^i \geq -\mathbf{w}_k^{b,i}.$$

where

$$\delta\mathbf{w}^i = \mathbf{w}_k^{a,i} - \mathbf{w}_k^{b,i}, \mathbf{f}^i = \mathbf{w}_k^{o,i} - \mathbf{H}_k\mathbf{w}_k^{b,i} - \mathbf{H}_k\delta\mathbf{w}^i - \bar{\mathbf{r}}_k^o.$$

Janjic, T., D. McLaughlin, S. E. Cohn, M. Verlaan, 2014: Conservation of mass and preservation of positivity with ensemble-type Kalman filter algorithms, Mon. Wea. Rev., 142, No. 2, 755-773.

SQPEns algorithm

Inverse of ensemble derived analysis error covariance can be used to minimize the cost function to obtain the analysis

$$\mathbf{w}_k^{a,i} = \mathbf{w}_k^{b,i} + \arg\min_{\delta w^i} \frac{1}{2}[\delta \mathbf{w}^{i\,T} (\mathbf{P}_k^b)^{-1} \delta \mathbf{w}^i + \mathbf{f}^{i\,T} \mathbf{R}_k^{-1} \mathbf{f}^i]$$

subject to

$$c_j(\delta \mathbf{w}_i) \leq 0, \;\; j \in \{1, 2, ..., m_1\}$$
$$g_k(\delta \mathbf{w}_i) = 0, \;\; k \in \{1, 2, ..., m_2\}$$

where

$$\delta \mathbf{w}^i = \mathbf{w}_k^{a,i} - \mathbf{w}_k^{b,i}, \mathbf{f}^i = \mathbf{w}_k^{o,i} - \mathbf{H}_k \mathbf{w}_k^{b,i} - \mathbf{H}_k \delta \mathbf{w}^i - \bar{\mathbf{r}}_k^o.$$

Zeng, Y., T. Janjić, Y. Ruckstuhl and M. Verlaan, 2017: Ensemble-type Kalman filter algorithm conserving mass, total energy and enstrophy, Q. J. R. Meteorol. Soc., 143:708, 2902–2914, doi:10.1002/qj.3142.

QPEns algorithm in ensemble space

$\rho = Rank(\mathbf{P}^b)$, which is no larger than $N - 1$

$$\delta \mathbf{w}^i = \mathbf{L}\eta^i$$

$$\mathbf{P}^b = \mathbf{L}\mathbf{L}^T$$

QPEns algorithm in ensemble space

$\rho = Rank(\mathbf{P}^b)$, which is no larger than $N-1$

$$\delta\mathbf{w}^i = \mathbf{L}\eta^i$$

$$\mathbf{P}^b = \mathbf{L}\mathbf{L}^T$$

QPEns Algorithm in ensemble space

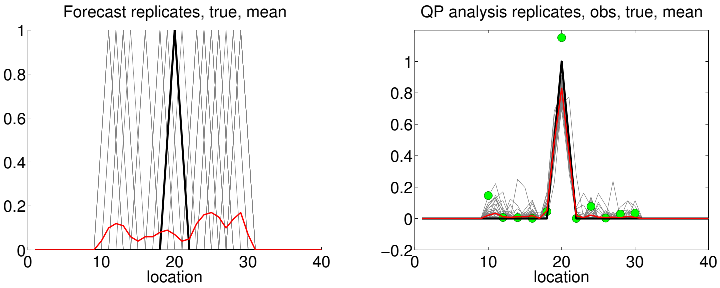
$$\eta^i = \arg\min_{\eta^i} \frac{1}{2}[{\eta^i}^T \eta^i + {\mathbf{f}^i}^T \mathbf{R}^{-1}\mathbf{f}^i]$$

subject to the following non-negativity constraint:

$$-\mathbf{L}\eta^i \leq \mathbf{w}_k^{b,i}.$$

The algorithm reduces to EnKF if there are no constraints present.

Preserving physical properties



QPEns analysis in ensemble space with positivity constraint. Both mass conservation and positivity constraint improve analysis.

Modified shallow water model

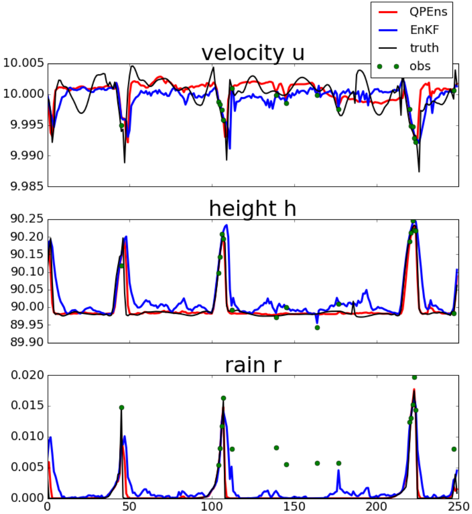
$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{\partial(\phi + \gamma^2 r)}{\partial x} = \beta_u + D_u\frac{\partial^2 u}{\partial x^2}, \phi = \begin{cases} \phi_c & \textit{if } \ h > h_c \\ gh & \textit{otherwise}, \end{cases}$$

$$\frac{\partial r}{\partial t} + u\frac{\partial r}{\partial x} = D_r\frac{\partial^2 r}{\partial x^2} - \alpha r - \begin{cases} \delta\frac{\partial u}{\partial x}, & \textit{if } h > h_r \textit{and} \frac{\partial u}{\partial x} < 0 \\ 0 & \textit{otherwise}, \end{cases}$$

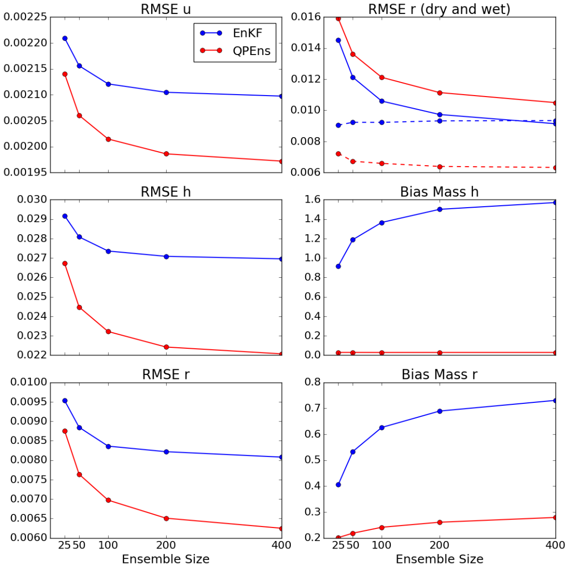
$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} = D_h\frac{\partial^2 h}{\partial x^2}.$$

Wuersch and Craig 2014: A simple dynamical model of cumulus convection for data assimilation research., Meteorol. Z., 23, 483-490.

EnKF vs. QPEns



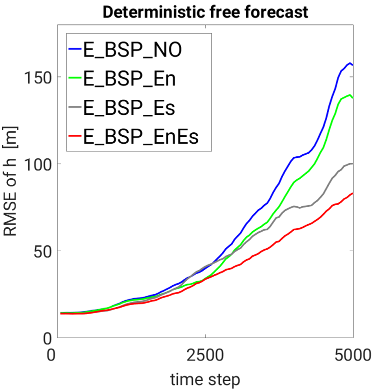
EnKF vs. QPEns analysis with positivity and mass constraint (Ruckstuhl and Janjic 2018) for modified shallow water model (Wuersch and Craig 2014).

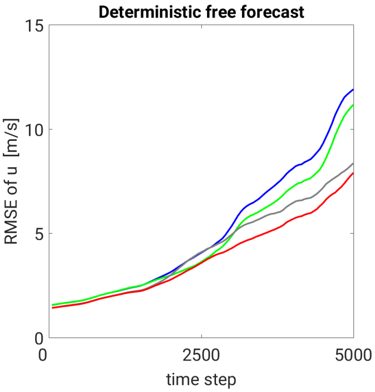


Ruckstuhl and Janjic 2018: Parameter and state estimation with ensemble Kalman filter based algorithms for convective scale applications.

Q.J.R. Meteorol. Soc.. 144:712, 826–841, doi:10.1002/qj.3257.

Prediction 2D SW



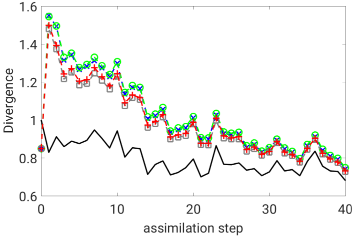


RMSE for h

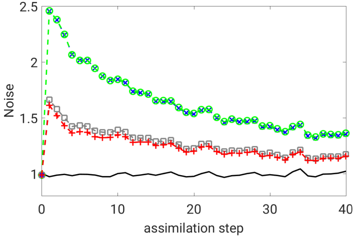
RMSE for u

Zeng, Y., T. Janjić, Y. Ruckstuhl and M. Verlaan, 2017: Ensemble-type Kalman filter algorithm conserving mass, total energy and enstrophy, Q. J. R. Meteorol. Soc., 143:708, 2902–2914, doi:10.1002/qj.3142.

Diagnostics



Divergence

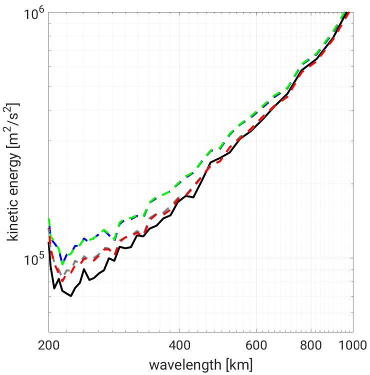


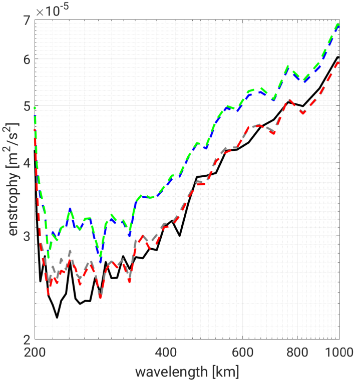
Noise

Variations of model diagnostics of divergence and noise within the data assimilation in experiments

E_BSP_NO E_BSP_En E_BSP_Es and E_BSP_EnEs.

Small scale spectra





Energy spectra

Enstrophy spectra

E_BSP_NO E_BSP_En E_BSP_Es E_BSP_EnEs.

Conclusion

- QPEns a method for addressing positivity
- Method is by construction multivariate
- Allows inclusion of other linear and nonlinear constraints
- Improves accuracy and bias in simple problems
- Adjoint not needed

Conclusion

- QPEns a method for addressing positivity
- Method is by construction multivariate
- Allows inclusion of other linear and nonlinear constraints
- Improves accuracy and bias in simple problems
- Adjoint not needed

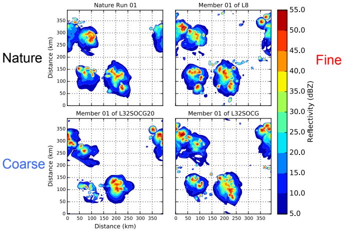
- Although total energy of the analysis ensemble mean converges towards the nature run value with time, enstrophy does not.
- Imposing the conservation of enstrophy within the data assimilation effectively avoids the spurious energy cascade of rotational part and this way succesfully suppresses the noise.
- Conserving mass and positivity reduces the noise in convective scale data assimilation applications.

Outlook

- Tests of imposing physical constraints on a hierarchy of 2D models for robustness across scales and in presence of sources and sinks, boundary conditions, etc.

Outlook

- Tests of imposing physical constraints on a hierarchy of 2D models for robustness across scales and in presence of sources and sinks, boundary conditions, etc.
- Application to high dimensional systems (either through optimization research as in T. Janjic, Y. Ruckstuhl and P. L. Toint, 2019 or through machine learning)

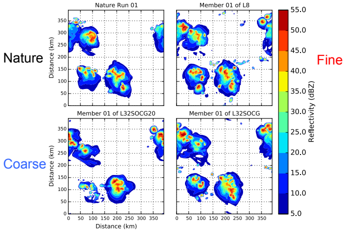


- Continued research in model and representation error for both data assimilation (e.g. exploring stochastic approaches to representation error) and ensemble forecasting.

Outlook

- Tests of imposing physical constraints on a hierarchy of 2D models for robustness across scales and in presence of sources and sinks, boundary conditions, etc.

- Application to high dimensional systems (either through optimization research as in T. Janjic, Y. Ruckstuhl and P. L. Toint, 2019 or through machine learning)



- Continued research in model and representation error for both data assimilation (e.g. exploring stochastic approaches to representation error) and ensemble forecasting.

- Predictability studies