A Hybrid Nudging-Ensemble Kalman Filter Approach to Data Assimilation in WRF/DART

Lili Lei[1,2], David R. Stauffer[1] and Aijun Deng[1]

[1] The Pennsylvania State University

[2] NCAR/ASP

Outline

- Introduction
- Methodology
- Experiment design
- Results
- Conclusions

- Lei, L., Stauffer, D. R., Haupt, S. E. and Young, G. S., 2012. A hybrid nudging-ensemble Kalman filter approach to data assimilation. Part I: application in the Lorenz system. *Tellus A*, **64**, 18484.
- Lei, L., Stauffer, D. R. and Deng, A., 2012. A hybrid nudging-ensemble Kalman filter approach to data assimilation. Part II: application in a shallow-water model. *Tellus A*, **64**, 18485.
- Lei, L., Stauffer, D. R. and Deng, A., 2012. A hybrid nudging-ensemble Kalman filter approach to data assimilation in WRF/DART. *Q. J. R. Meteorol. Soc.* doi: 10.1002/qj.1939

Introduction

- The EnKF uses the statistical properties of an ensemble model forecast to estimate the flow-dependent background error covariance, and then determines how an observation modifies the forecast background fields (Evensen 1994; Houtekamer and Mitchell 1998).
- As an intermittent data assimilation approach, the EnKF uses a cycle of a model integration period, analysis step, and then another model integration period, often causing discontinuities/error spikes around the observation times (e.g., Hunt et al. 2004).
- A time-continuous, seamless meteorological field is preferred in many applications, especially for use in driving air-quality and atmospheric-chemistry models (e.g., Stauffer et al. 2000; Tanrikulu et al. 2000; Otte 2008a, b).

Introduction

- Thus we propose a hybrid method that uses nudging-type terms to apply the EnKF gradually in time and uses the EnKF to provide flow-dependent and time-dependent nudging coefficients. This hybrid nudging-EnKF approach can effectively combine the strengths of the nudging and EnKF while avoiding their individual weaknesses.
- It is hypothesized that if the EnKF could be applied gradually in time, a continuous and seamless analysis can be produced and the analysis will be improved by reducing the intermittent discontinuities or bursts. These improved and seamless meteorological background fields could also improve the transport and dispersion simulation.

Methodology for the HNEnKF

Model equation with
additional nudging terms:

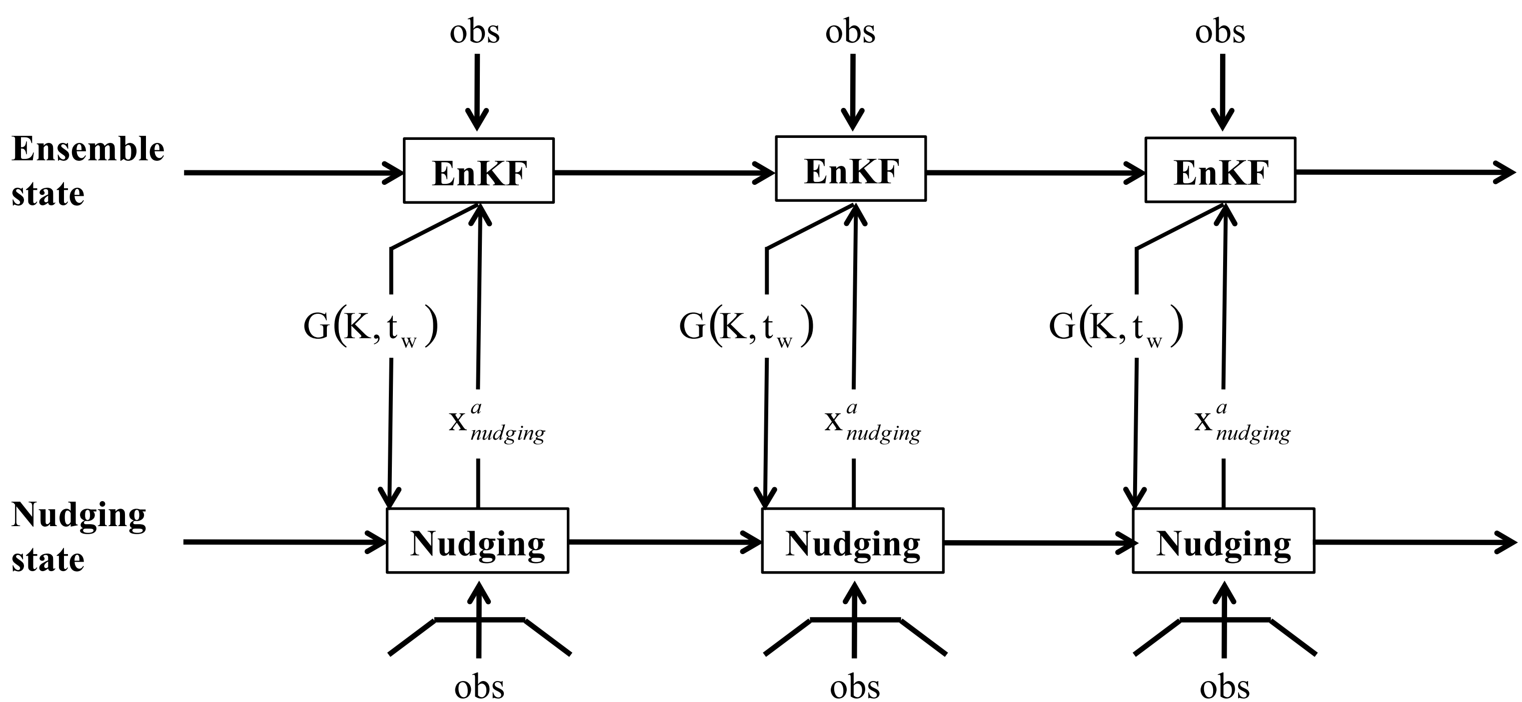
$$\frac{d\mathbf{x}}{dt} = f\left(\mathbf{x}\right) + \mathbf{G} \cdot \mathrm{w}_s \cdot \mathrm{w}_t \cdot \left(\mathbf{y}^o - \mathbf{H}\mathbf{x}\right)$$

Hybrid nudging coefficients:

$$\mathbf{G} \cdot \mathrm{w}_s = \frac{1}{\left(\sum_{t=t^o-\tau_{\mathrm{N}}}^{t^o} \mathrm{w}_t \cdot \Delta t\right)} \cdot \mathbf{K}$$

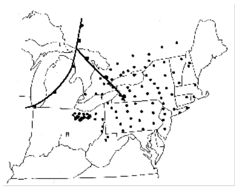
- The hybrid nudging coefficients come directly from the EnKF gain matrix **K** that contains information from the flow-dependent background error covariances computed from an ensemble forecast.
- Thus there is no need to specify the nudging strength or the spatial nudging weighting coefficient.
- The hybrid nudging terms include not only the standard diagonal terms (i.e., u correction in u-equation, v correction in v-equation, etc.) in the nudging magnitude matrix **G**, but also the off-diagonal terms (i.e., inter-variable influence).
- This statistical inter-variable influence is included in the model's relaxation terms to gradually and continuously force the model towards the observations.
- Thus the error spikes and dynamic imbalances often produced by intermittent EnKF are reduced.

Methodology for the HNEnKF

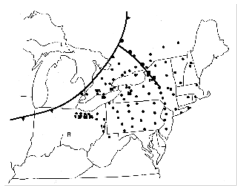


Overview of the 18-20 September 1983 CAPTEX-83 case

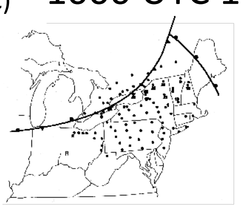
(a) 2200 UTC 18 September



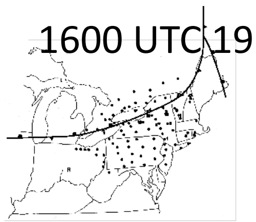
(b) 0400 UTC 19 September



(c) 1000 UTC 19 September



(d) 1600 UTC 19 September



(After Deng et al. 2004)

WRF/DART model description

- WRF-ARW version 3.1.1 (Skamarock et al. 2008)
- Horizontal grid-spacing of 12 km on a 208×190 horizontal grid with 33 vertical levels and a model top at 100 hPa
- Model physics: Longwave Rapid Radiative Transfer Model (RRTM: Mlawer et al. 1997), the shortwave Dudhia radiation scheme (Dudhia 1989), and the thermal diffusion land-surface scheme
- 48-h simulation period: 1200 UTC 18 – 1200 UTC 20 September 1983
- Initial conditions (ICs) and lateral boundary conditions (LBCs): generated by the NCEP/ NCAR Global Reanalyses, and then enhanced with observations by OBSGRID
- DART (Data Assimilation Research Testbed, Anderson et al. 2009): EAKF (a deterministic ensemble square root filter)

The WRF 12-km domain (outer domain) and the 12-km SCIPUFF AT&D domain (inner domain).

Experiment design – observations and verification techniques

- Three-hourly WMO surface observations and twelve-hourly rawinsonde observations are assimilated in the 48-h simulation period.
- Verification:
  - Three-hourly forecasts (priors)
  - Independent tracer data verification

    Use hourly WRF analyses as input in the SCIPUFF AT&D model (Sykes et al. 2004) to predict surface tracer concentrations, and then verify them against the observed surface tracer concentration data.

    The observed surface tracer concentration data is available at four times (2200 UTC 18 September 1983, 0400 UTC 19 September 1983, 1000 UTC 19 September 1983, and 1600 UTC 19 September 1983).

Experiment design – ensemble configuration

- IC ensemble: contains perturbations of the ICs and LBCs.
  - Adding perturbations, which are drawn from a multivariate normal distribution by use of the WRF-3DVAR, to the ICs and LBCs.
  - Ensemble size is 24
- ICPH ensemble: contains multi-physics members in addition to the perturbed ICs and LBCs members
  - Eight physics configurations are used
  - Ensemble size is 24

Physics configuration	Microphysics	Convective	PBL
1	WSM-3	Kain-Fritsch	MYJ
2	Lin et al.	Kain-Fritsch	MYJ
3	WSM-3	Betts-Miller-Janjic	MYJ
4	WSM-3	Kain-Fritsch	YSU
5	Lin et al.	Betts-Miller-Janjic	MYJ
6	Lin et al.	Kain-Fritsch	YSU
7	WSM-3	Betts-Miller-Janjic	YSU
8	Lin et al.	Betts-Miller-Janjic	YSU

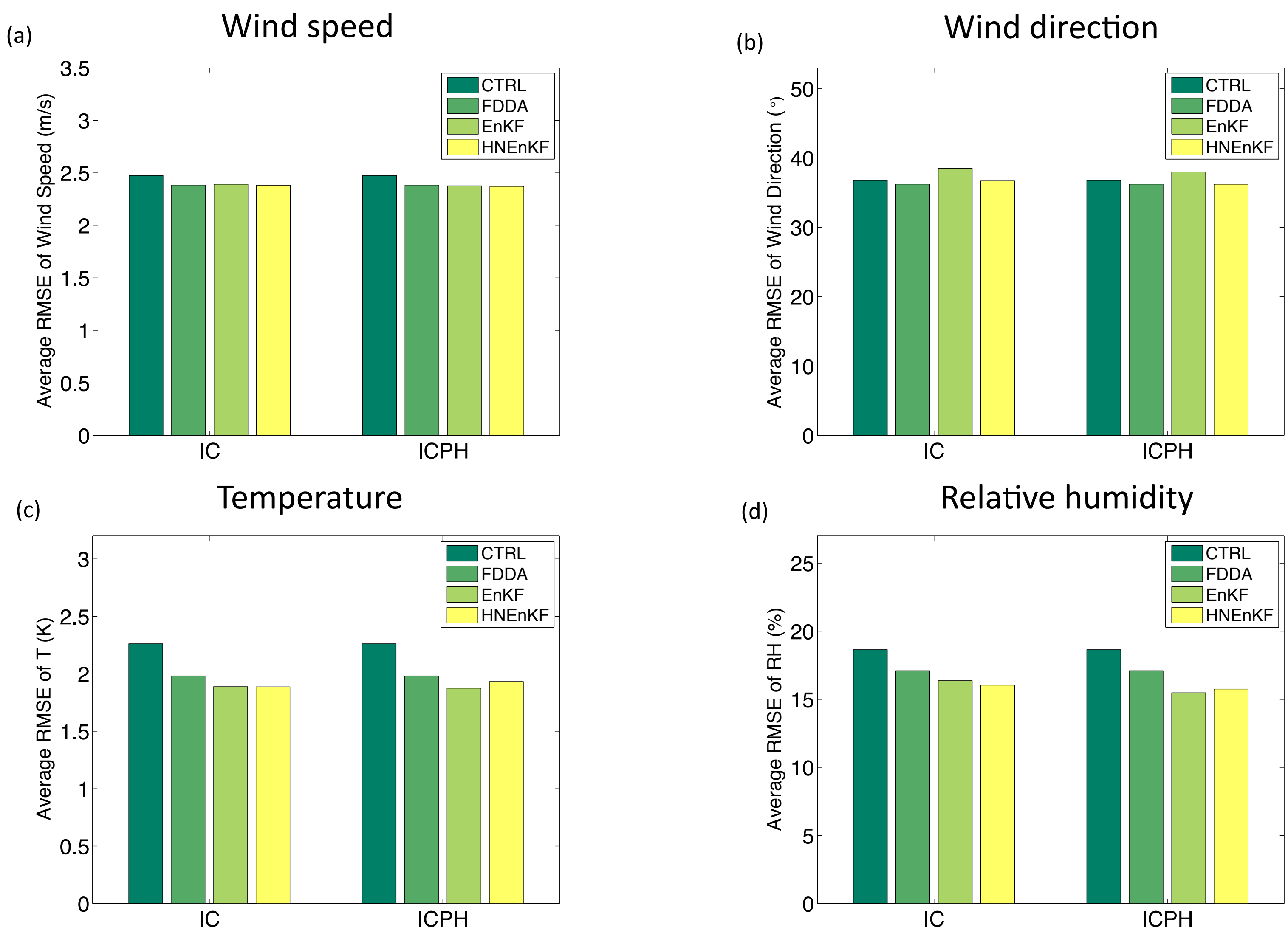
Experiment design – parameters

Experiment	Nudging strength	Horizontal radius of influence	Surface data vertical radius of influence (stable PBL)	Surface data vertical radius of influence (unstable PBL)	Half-period of nudging time window	Horizontal error covariance localization	Vertical error covariance localization	Error covariance inflation
FDDA	4×10^{-4} s^{-1}	67-200 km	100 m	PBL top plus 50m	1-2 h	---	---	---
EnKF	---	---	---	---	---	533 km	150 hPa	Adaptive inflation
HNEnKF	---	---	---	---	1-2 h	533 km	150 hPa	Adaptive inflation

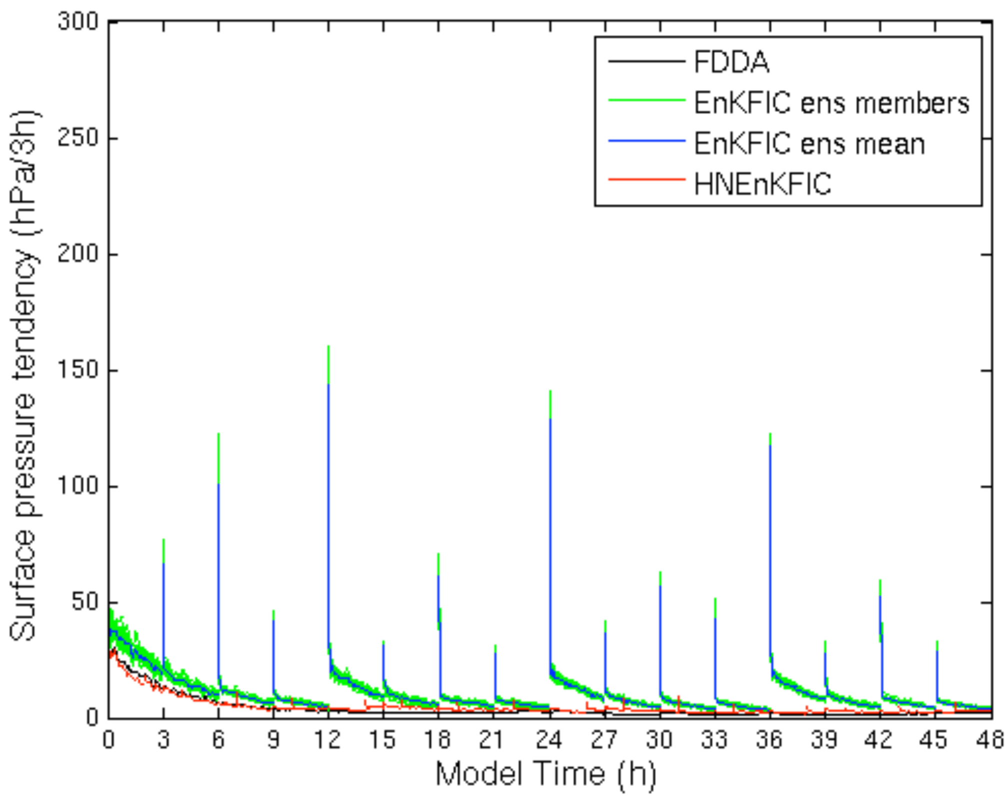
Experiment design – experiment description

Exp. Name	Exp. Description
CTRL	Assimilate no observations
FDDA	Assimilate observations by observation nudging
EnKFIC	Assimilate observations by EnKF with IC ensemble
EnKFICPH	Assimilate observations by EnKF with ICPH ensemble
HNEnKFIC	Assimilate observations by HNEnKF with IC ensemble
HNEnKFICPH	Assimilate observations by HNEnKF with ICPH ensemble

Evaluation using meteorological data (priors)

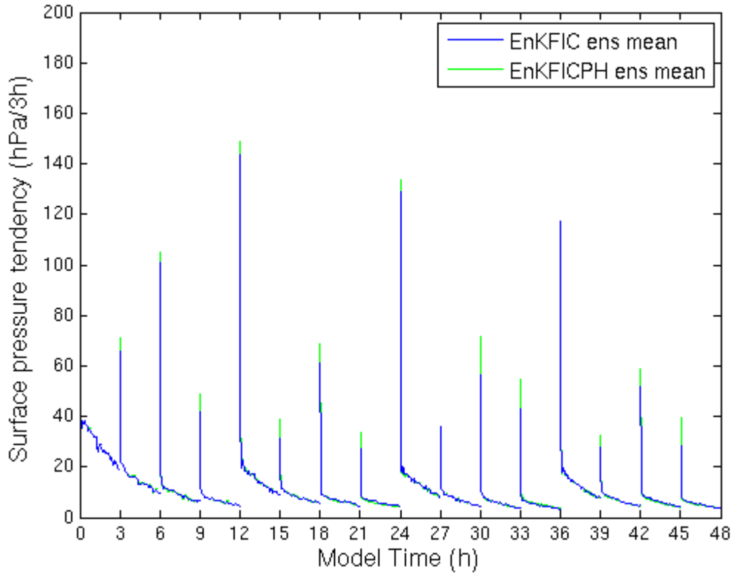


Evaluation of temporal smoothness and insertion noise for IC ensemble

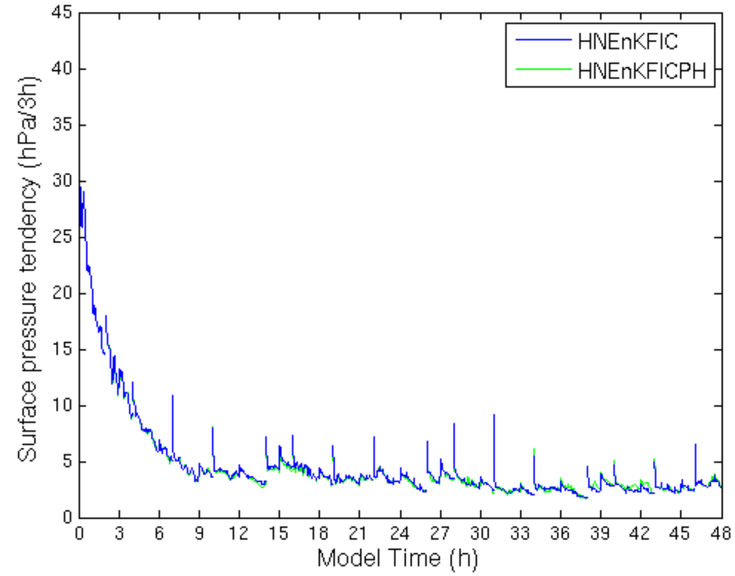


Evaluation of temporal smoothness and insertion noise for IC and ICPH ensembles

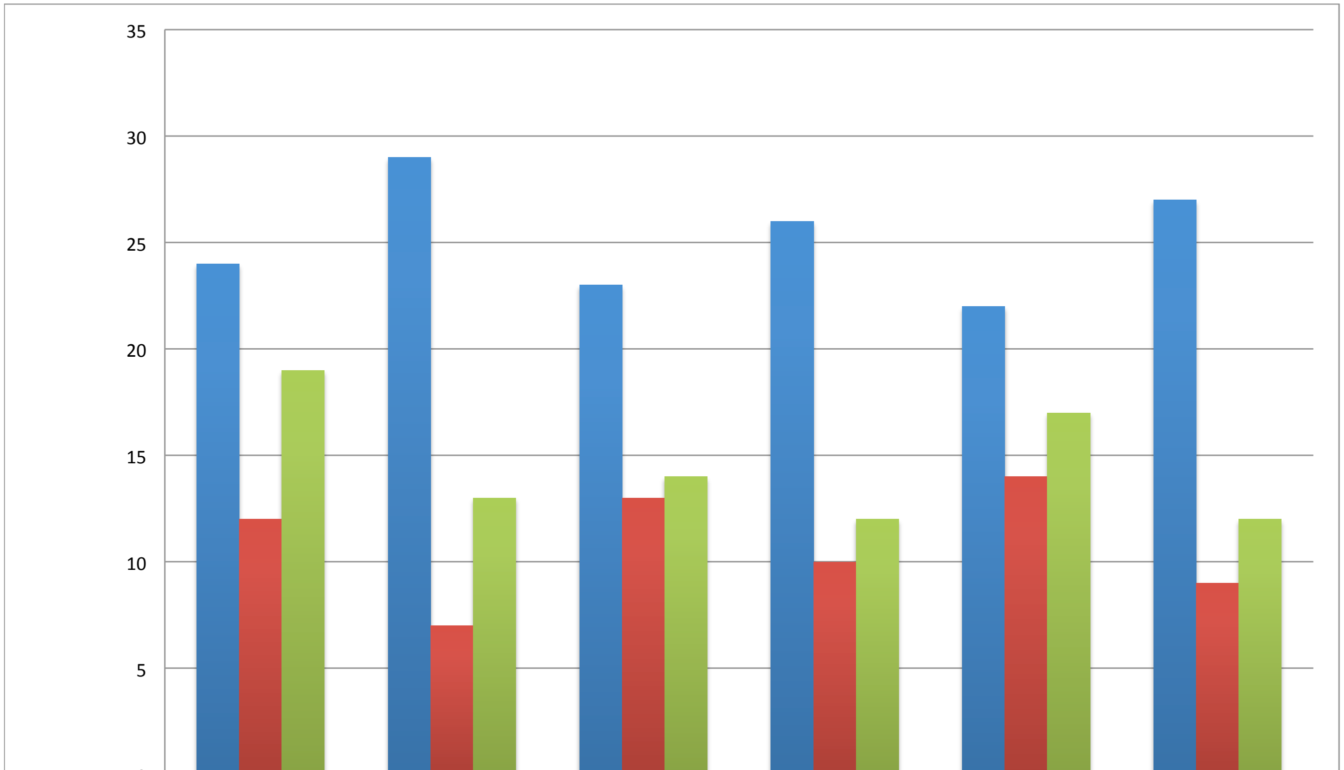
(a) EnKF



(b) HNEnKF



Evaluation using independent surface tracer data



	CTRL	FDDA	EnKFIC	HNEnKFIC	EnKFICPH	HNEnKFICPH
Hit	24	29	23	26	22	27
Miss	12	7	13	10	14	9
FalseAlarm	19	13	14	12	17	12

The composite statistics (hits, misses and false alarms) through the 24-h period

Evaluation using independent surface tracer data

Ordinal ranking	Experiment	Sum of misses and false alarms
1	FDDA	20
2	HNEnKFICPH	21
3	HNEnKFIC	22
4	EnKFIC	27
5	EnKFICPH	31
5	CTRL	31

Conclusions

- The HNEnKF has similar or better three-hourly forecast (prior) statistics than the EnKF and FDDA.
- The intermittent EnKF has much higher noise levels (the domain average absolute surface pressure tendency) than the continuous HNEnKF and FDDA following the data insertion.
- The noise levels in the EnKF are even higher when more observations are assimilated and when multi-physics ensemble members are included, while the HNEnKF shows little sensitivity to these factors.
- Verification against the independent surface tracer data shows the FDDA and HNEnKF have better statistics than the EnKF.
- Compared to the IC ensemble, the ICPH ensemble decreases the misses without increasing the false alarms for the HNEnKF, but degrades all tracer data statistics for the EnKF.