Toward large-domain high-resolution continuously cycling data assimilation systems

Craig Schwartz

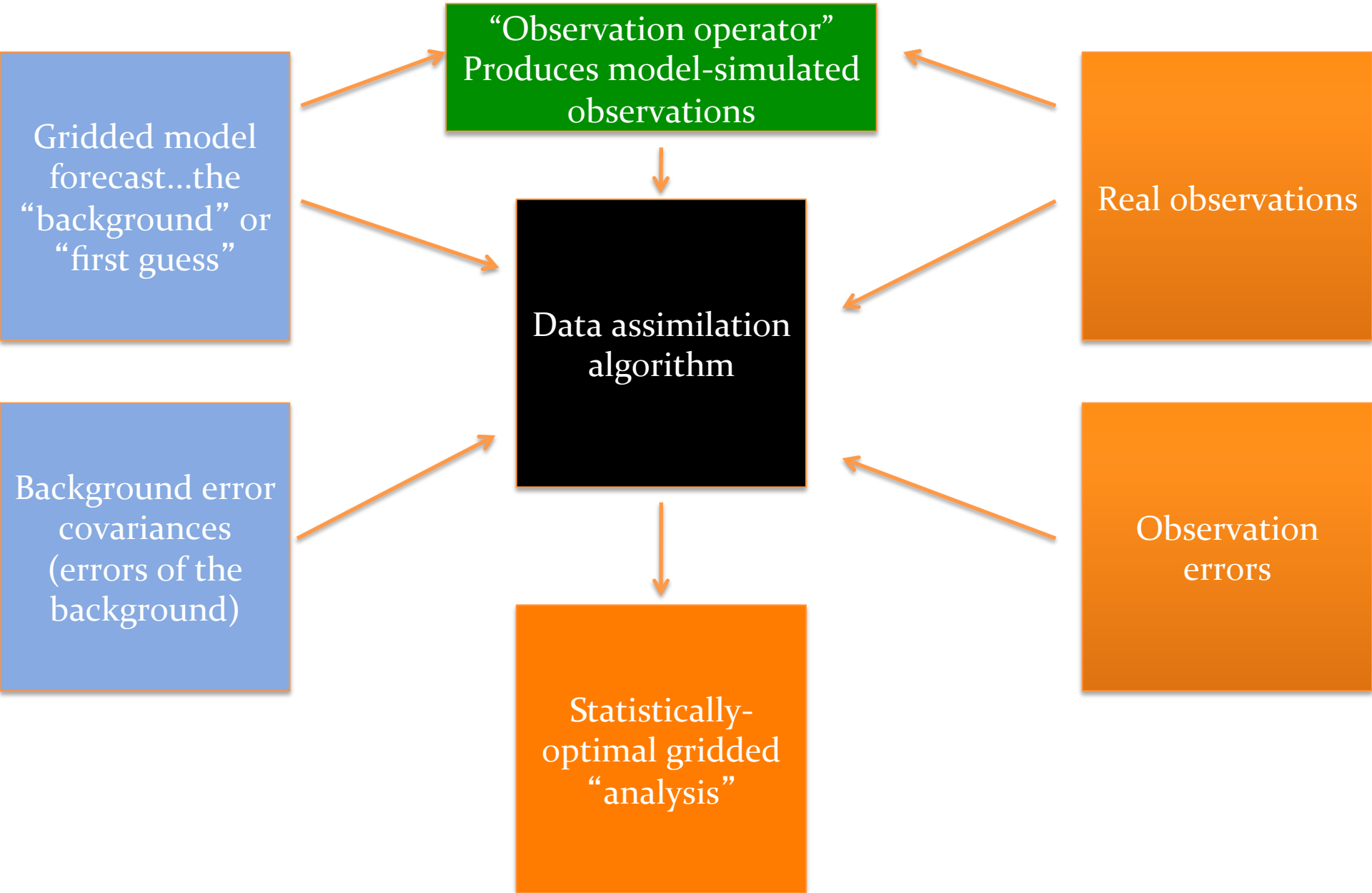
NCAR/MMM

Background

- Convection-allowing forecasts have traditionally been initialized from convection-parameterizing analyses
  - These forecasts have been good, but there are spin-up issues

- This work produces convection-allowing *analyses* over a large (e.g., 3000 km^2) domain
  - Many case studies have examined convection-allowing data assimilation over small domains

What is data assimilation?



A few data assimilation (DA) approaches

- Three-dimensional variational (3DVAR)
  - Background error covariances (BECs) typically fixed/time-invariant
  - May yield poor results when actual flow differs from that encapsulated within the fixed "climatology"
  - Computationally cheap

- Ensemble Kalman filter (EnKF)
  - Time-evolving, "flow-dependent" BECs estimated from a short-term ensemble forecast
  - More expensive than 3DVAR, but usually better

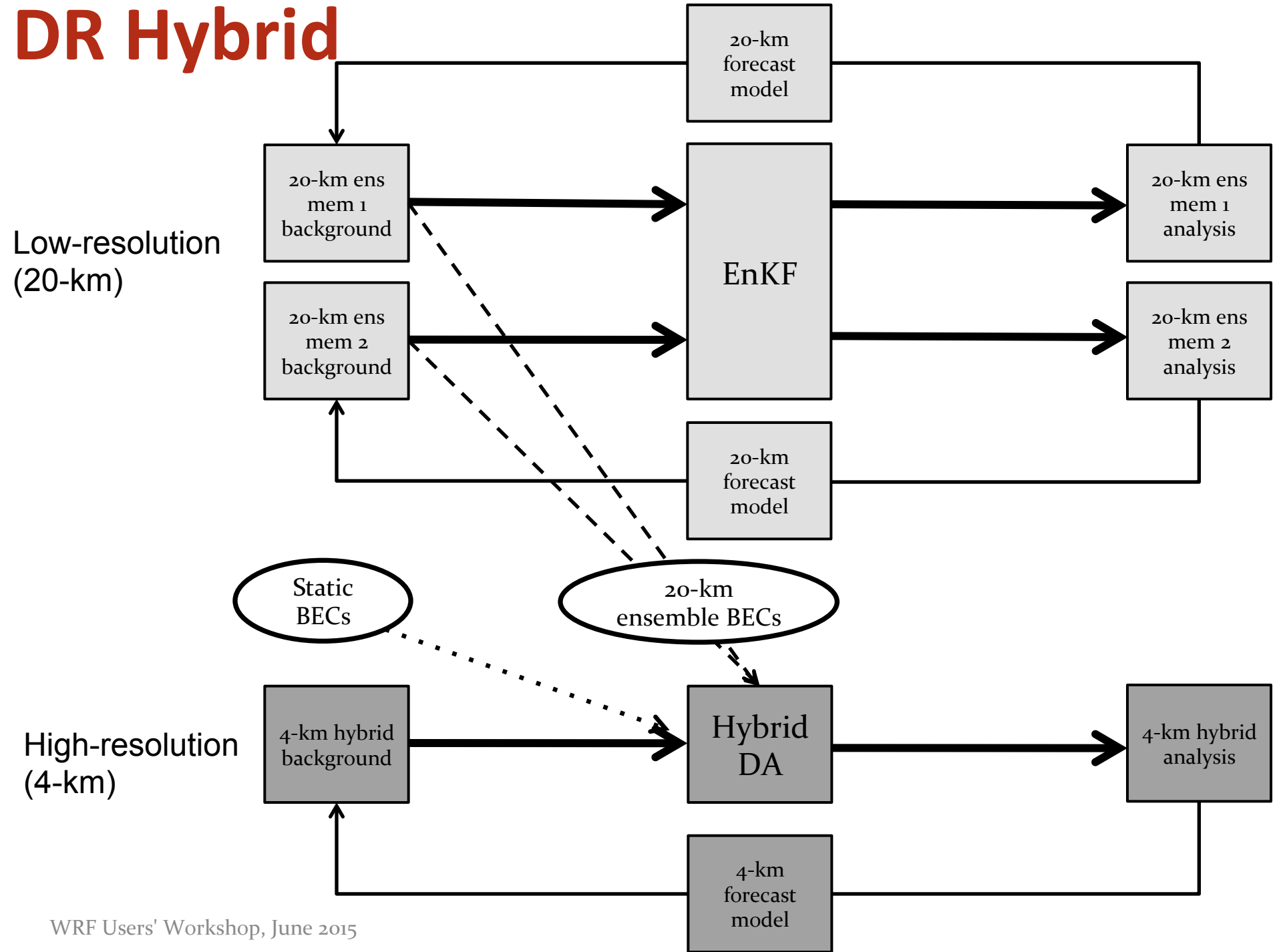
"Hybrid" DA

- "Hybrid" variational-ensemble DA systems incorporate ensemble background error covariances within a variational framework

- Combines static and flow-dependent background errors

- Can combine high-resolution background with a low-resolution ensemble in a "dual-resolution" (DR) configuration

DR Hybrid


20-km forecast model
20-km ens mem 1 background
20-km ens mem 2 background
EnKF
20-km ens mem 1 analysis
20-km ens mem 2 analysis
20-km forecast model
Low-resolution (20-km)
Static BECs
20-km ensemble BECs
4-km hybrid background
Hybrid DA
4-km hybrid analysis
High-resolution (4-km)
4-km forecast model

DA experiments

- Full-cycling (6-hr period) between May 4 – June 30, 2013

- Four DA experiments (analyses every 6-hrs):
  - Pure 3DVAR ("3DVAR_20km") : 20-km cycling
  - Pure 3DVAR ("3DVAR_4km") : *4-km* cycling
  - SR Hybrid ("Hybrid_20km") : 20-km cycling
  - DR Hybrid ("Hybrid_4km") : *4-km* cycling

- Hybrid runs coupled to a 20-km, 50-member EnKF

- All assimilated identical conventional observations

Forecast initialization

- 0000 UTC analyses initialized 36-hr 4-km WRF forecasts

- 4-km initial conditions were *downscaled 20-km analyses* in the 20-km 3DVAR, and 20-km hybrid experiments

- *True 4-km analyses* initialized 4-km forecasts in the DR hybrid and 4-km 3DVAR experiments

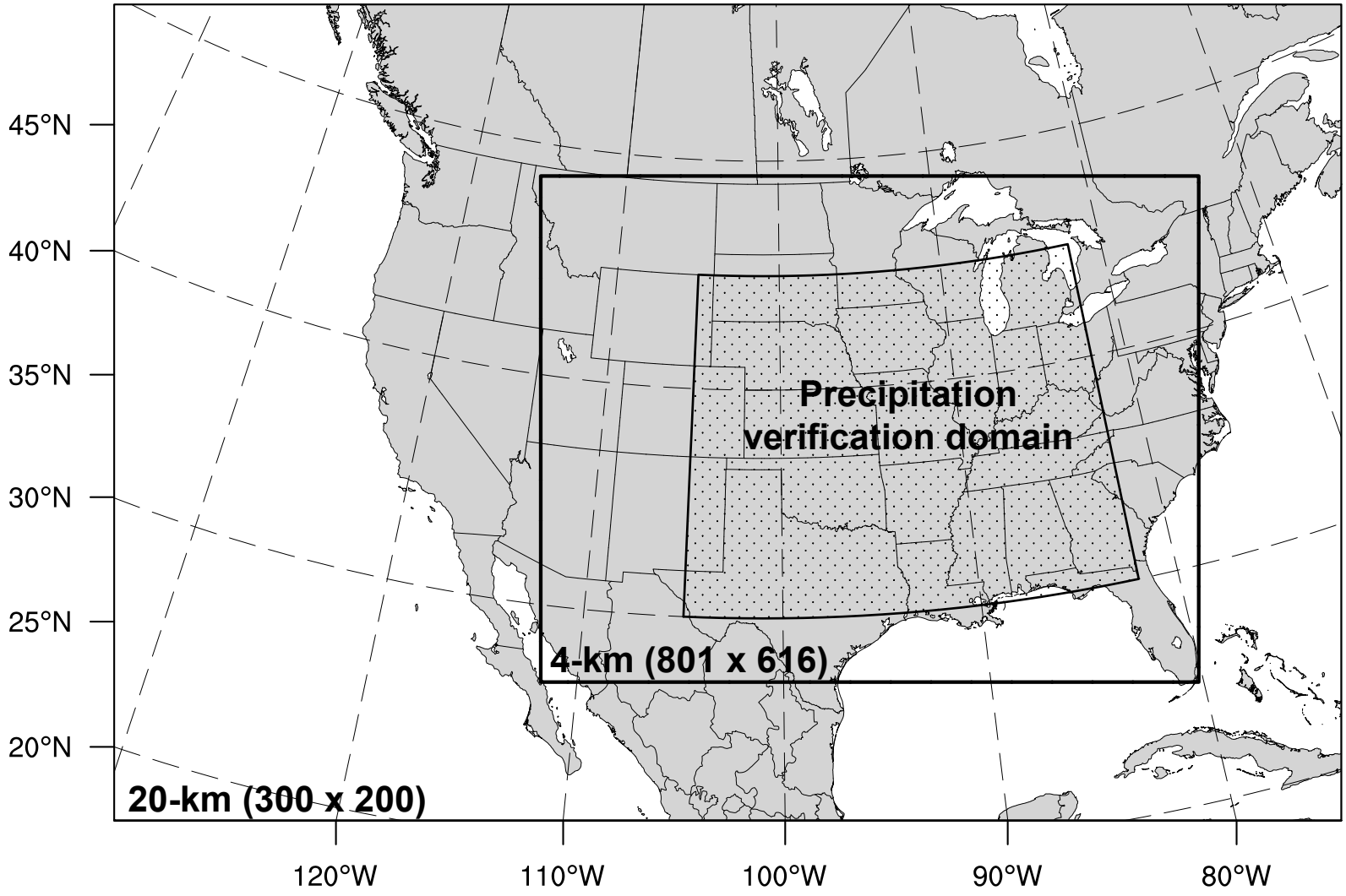
- Forecast differences between SR and DR hybrid experiments due to analysis resolution

- Control: Interpolate 0000 UTC GFS analyses directly onto the domain and run forecasts

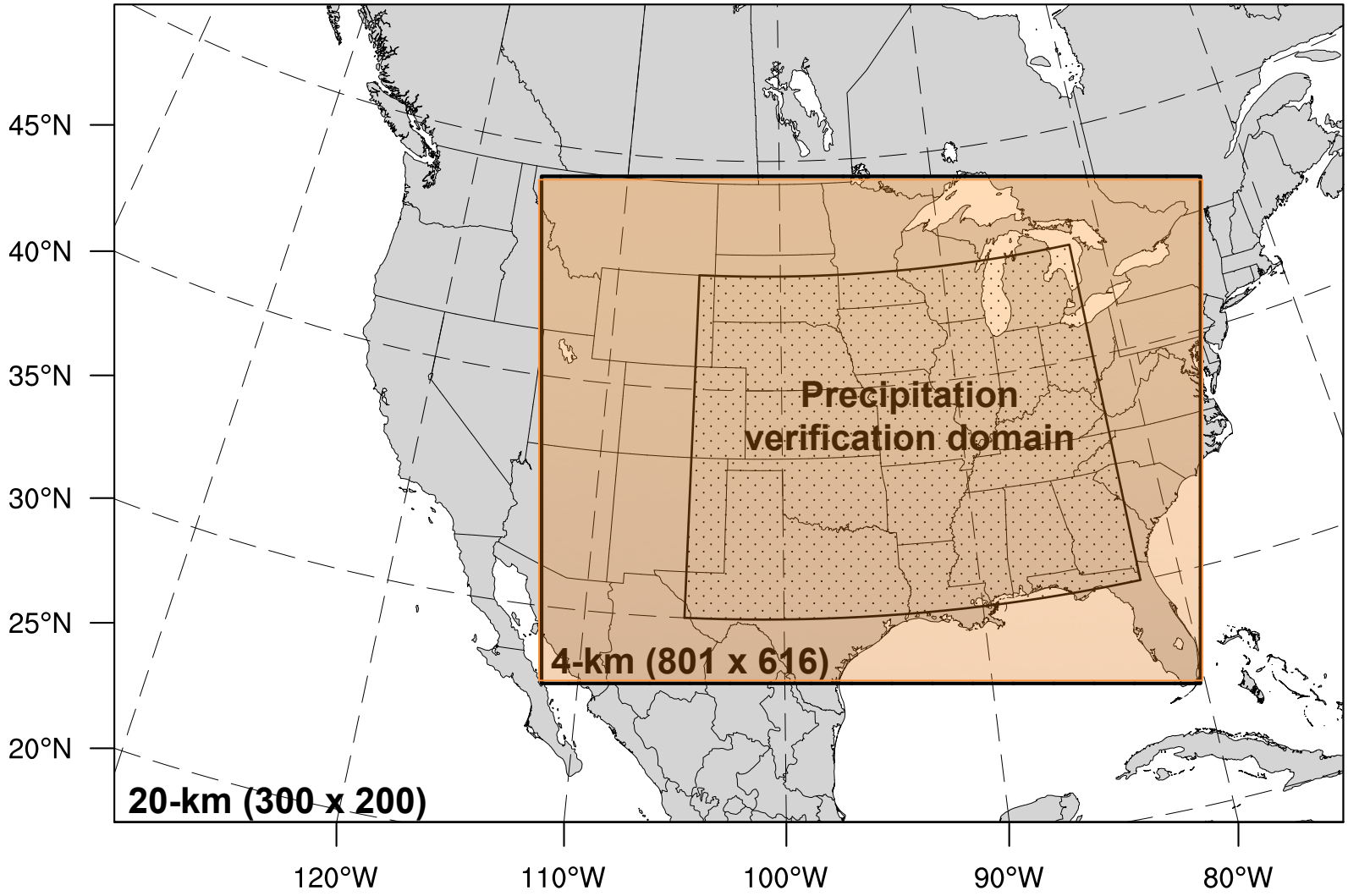
WRF settings and physics

- Forecast model: WRF-ARW (version 3.3.1)
- 57 vertical levels, 10 hPa top
- Physics:
  - Thompson microphysics
  - RRTMG longwave and shortwave radiation
  - MYJ PBL
  - NOAH land surface model
  - Aerosol, ozone climatologies for RRTMG
  - Tiedtke cumulus parameterization (20-km domain only)

Computational domain



Computational domain



20-km ensemble covariances over orange region used for 4-km DR analyses

Selected data assimilation settings

- NCEP's Gridpoint Statistical Interpolation (GSI) data assimilation system:
  - GSI-3DVAR
  - GSI-Hybrid coupled to NOAA's GSI-based EnKF

- 50 ensemble members in hybrid/EnKF

- Hybrid: 75% of background errors from ensemble, 25% from the static contribution

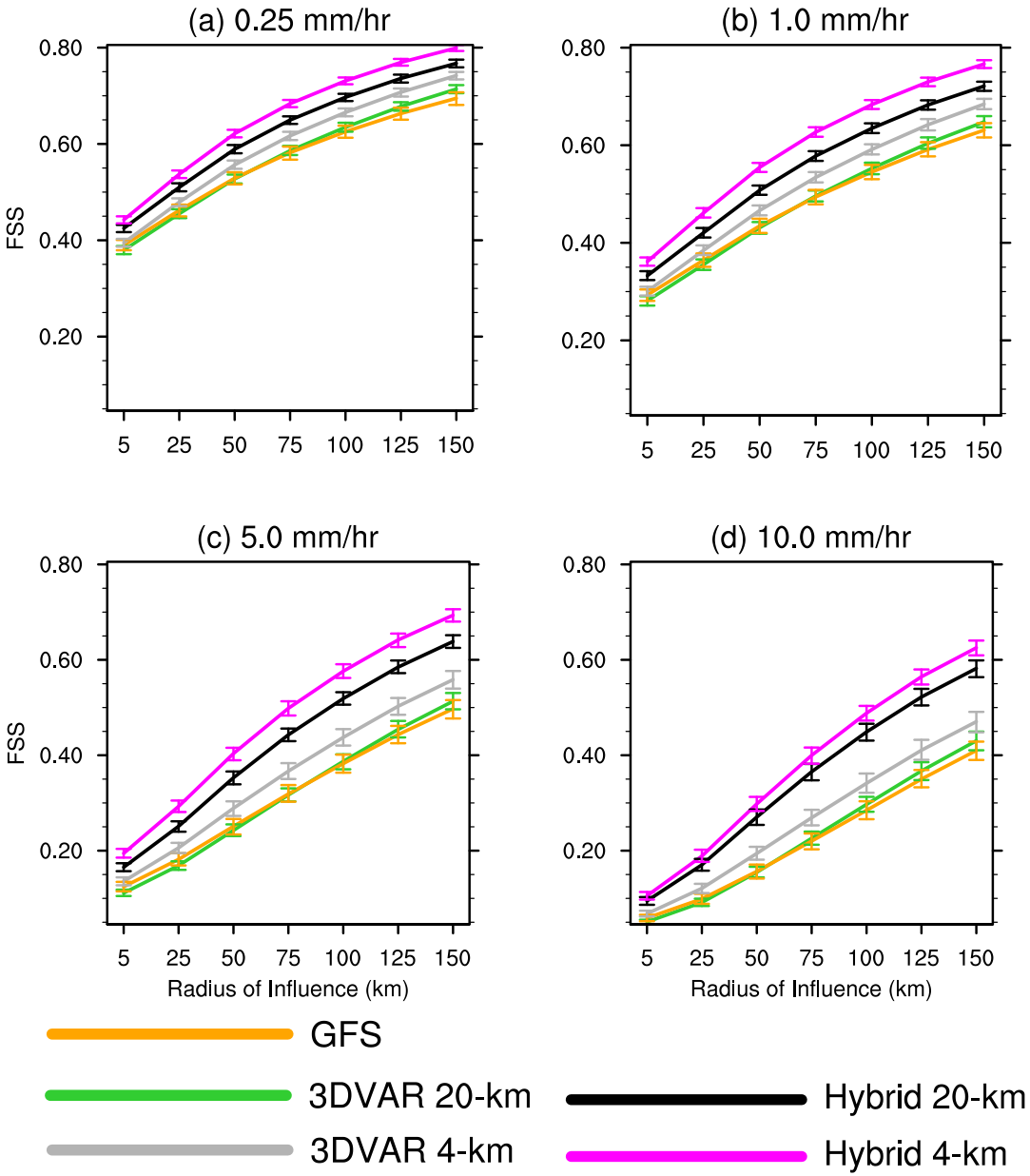
- Used posterior inflation for EnKF and localization in EnKF and hybrid

Precipitation verification

- Focus on bias-corrected 4-km hourly precipitation forecasts
- NCEP Stage IV observations as "truth"
- All precipitation statistics aggregated over 55 4-km forecasts
- Fractions skill score (FSS) quantifies displacement errors

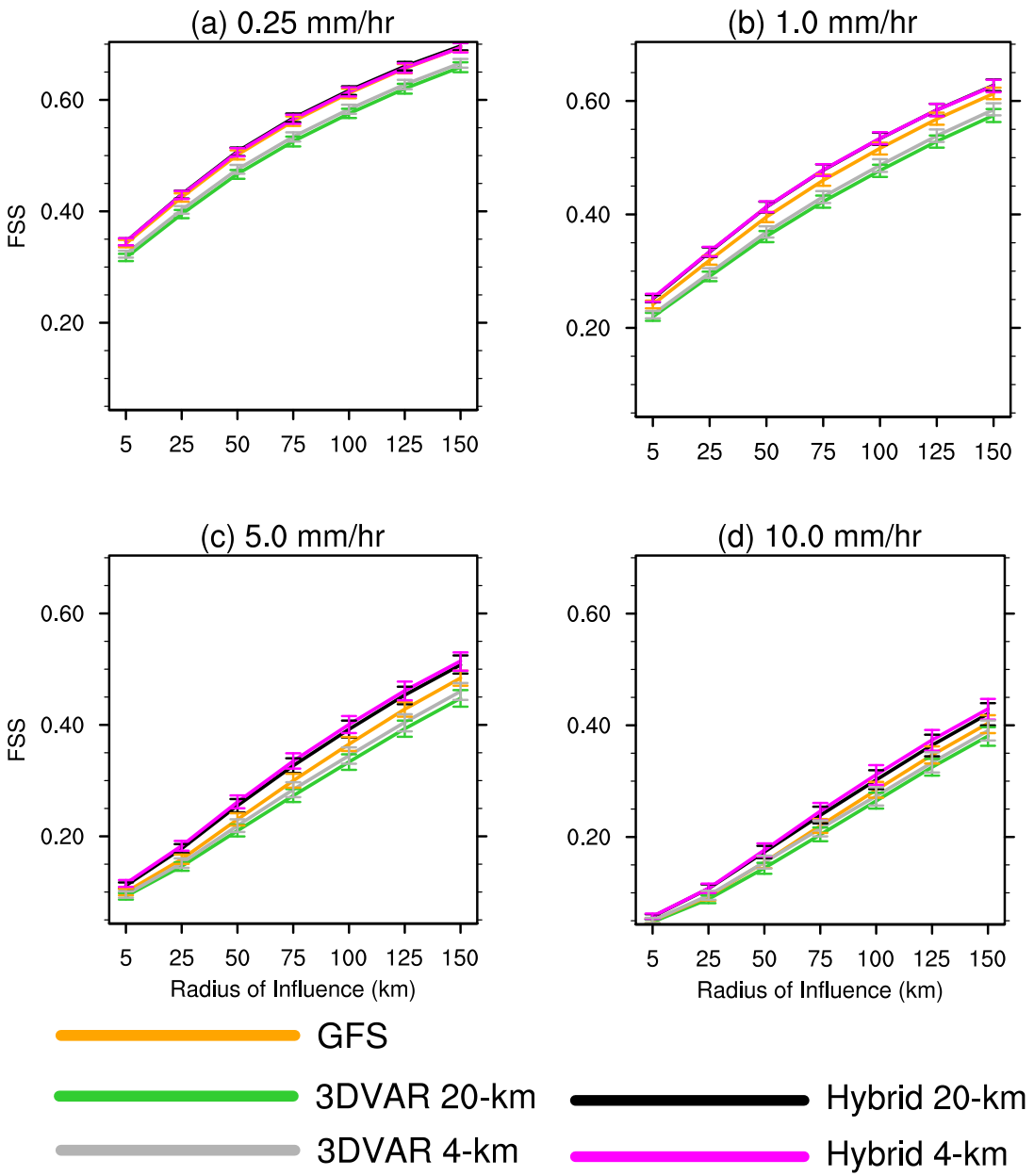
Precipitation verification: the first 12-hrs

- Fractions skill score (FSS) aggregated over the **first 12 forecast hours** and 55 4-km forecasts
- Precipitation forecasts bias-corrected



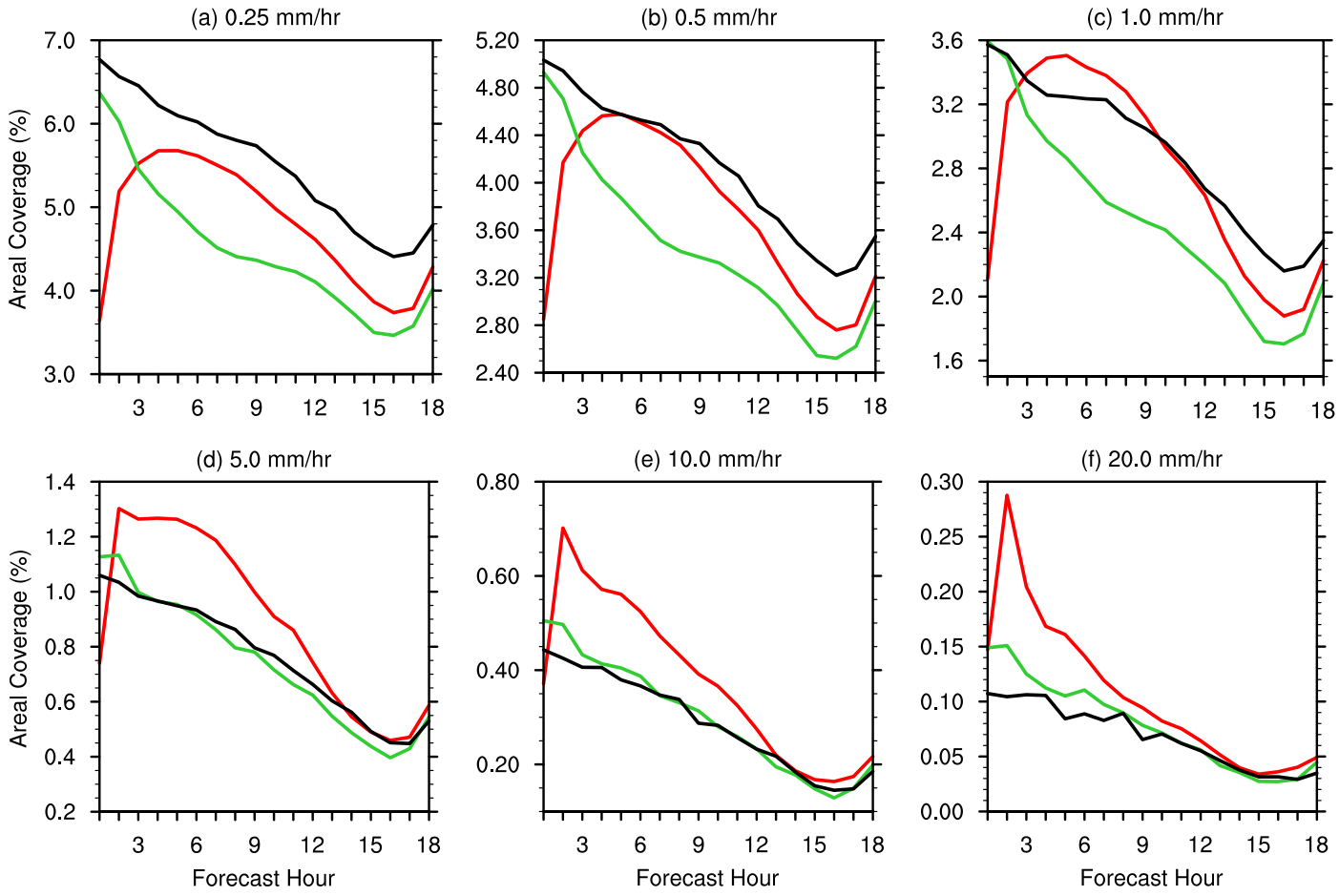
Precipitation verification: 18-36-hrs

- Fractions skill score (FSS) aggregated over **forecast hours 18-36** and 55 4-km forecasts
- Precipitation forecasts bias-corrected



Areal coverages of precipitation

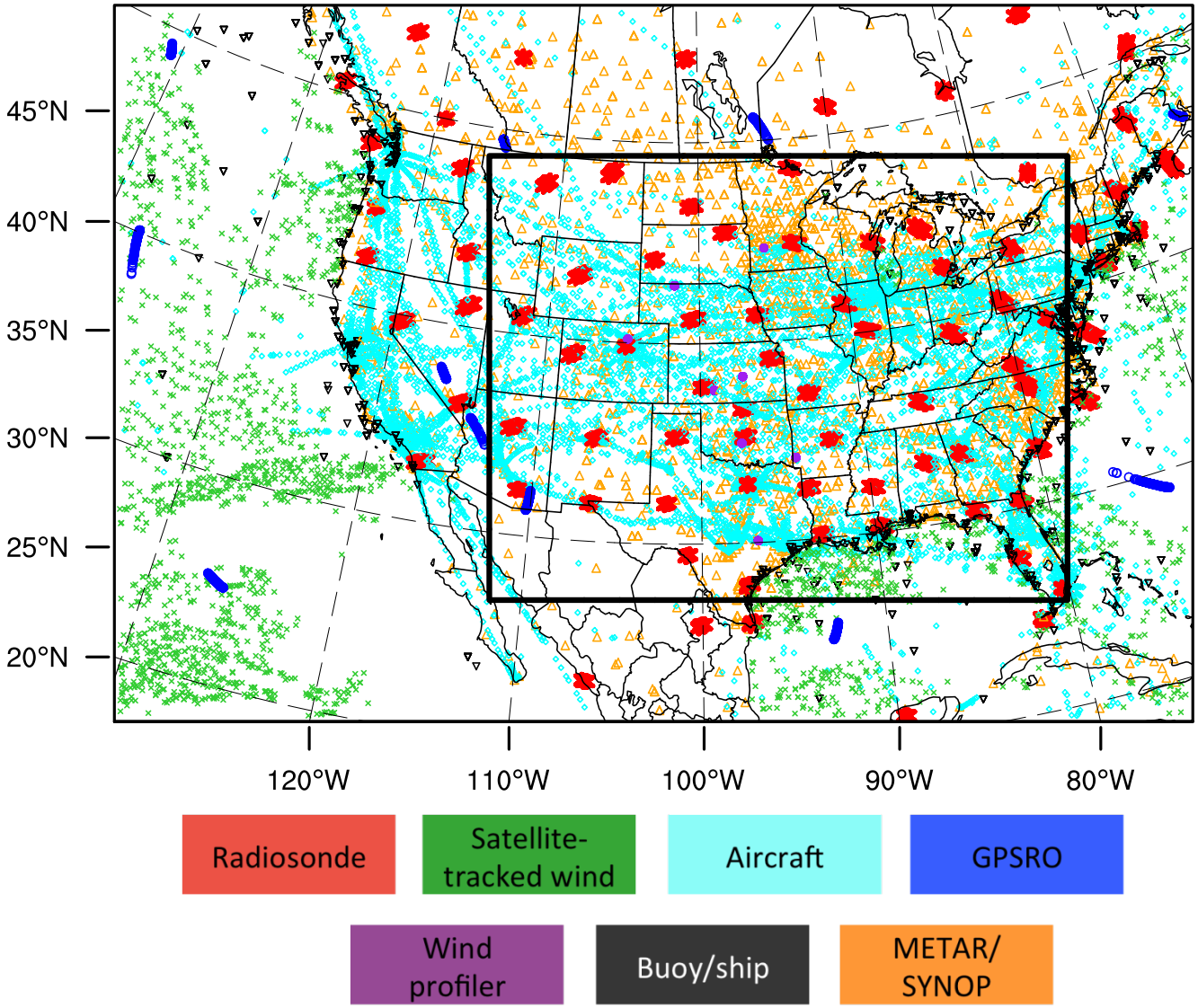
- Aggregate fractional coverage of precipitation exceeding certain thresholds aggregated over all 55 forecasts

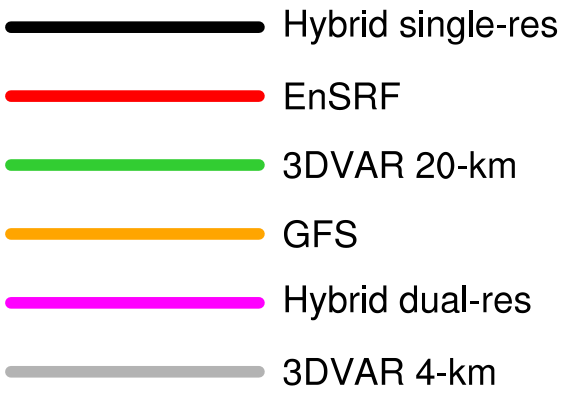


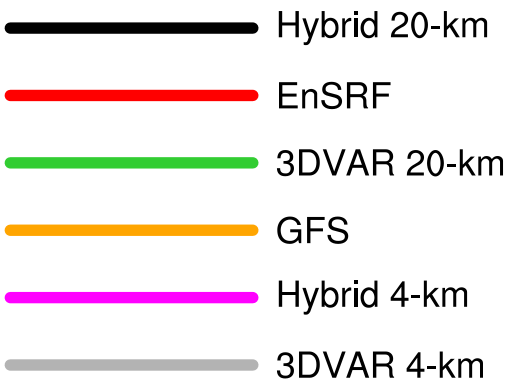
Summary

- DR hybrid DA systems are practical
- Biggest impact from high-resolution analyses at earlier times, but some benefit persists to longer times
- DA method more important than analysis resolution
- The DR hybrid system described here does not produce convection-permitting ensemble analyses

Observation snapshot (0000 UTC 26 May)




Hybrid single-res
EnSRF
3DVAR 20-km
GFS
Hybrid dual-res
3DVAR 4-km


Hybrid 20-km
EnSRF
3DVAR 20-km
GFS
Hybrid 4-km
3DVAR 4-km


GFS
3DVAR 20-km
3DVAR 4-km
EnSRF
Hybrid 20-km
Hybrid 4-km