A Hands-on DA Session

DARC, June-2025

What is Reanalysis Data?

- Reanalysis is our "best guess" of the past state: combining **models & observations**
- **Fills the Gaps:** It provides a complete, gridded dataset, even in areas with no direct measurements.
- **Consistent Over Time:** It creates a physically consistent record.

Getting Started: The Notebook

https://github.com/darc-reading/darc-training-2025/blob/main/reanalysis.ipynb

- **Follow Along:** Read the instructions in the markdown cells and run the code cells in order.
  - Task 1 - ERA5 as a Data Assimilation Product
  - Task 2 - Model-Observation Integration
  - Task 3 - Uncertainty via Ensemble Spread
  - Task 4 - Observing System Impact (ERA5 vs NCEP)

- **Think:** As you go, try to answer the **“ Questions/Discussions"** posed in the notebook.

Wrap-up

1. Obtain an analysis from the model and observations weighted by their respective uncertainties.
2. How key DA parameters such as localisation radius and error correlation length scale affect the analysis.
3. How an ensemble can provide information about analysis uncertainty and how ensemble size affects these estimates.
4. Changes in reanalysis when new satellite data is introduced

Questions

- Q1: How does the analysis represent a weighted average based on error covariances?
- Q2: In which regions does the analysis follow the model versus observations more closely, and why?
- Q3: How does increasing the localization radius affect the spatial distribution of analysis increments?

- Q1: How does the analysis represent a weighted average based on error covariances?

In Kalman filter-based methods, the analysis is a weighted average of the background (model forecast) and the observations.

The weights are determined by their respective error covariances

- Q2: In which regions does the analysis follow the model versus observations more closely, and why?

  Observation-dominated regions: Where observations are dense and accurate

  Model-dominated regions: Where observations are sparse or uncertain

- Q3: How does increasing the localization radius affect the spatial distribution of analysis increments?

Increasing localization radius allows observations to influence more distant grid points, spreads useful information further.
Too large radius can introduce spurious long-range correlations and physically unrealistic patterns