An Enhanced Methodology for Satellite Data Assimilation in a Mars Atmosphere Reanalysis

[1,2]Steven J. Greybush, [3]Ross N. Hoffman, [2]Eugenia Kalnay, [3]Thomas Nehrkorn, [2]Yongjing Zhao, [4]Matthew J. Hoffman, and [5]R. John Wilson

[1]The Pennsylvania State University; [2]University of Maryland, College Park; [3]AER; [4]Rochester Institute of Technology; [5]GFDL

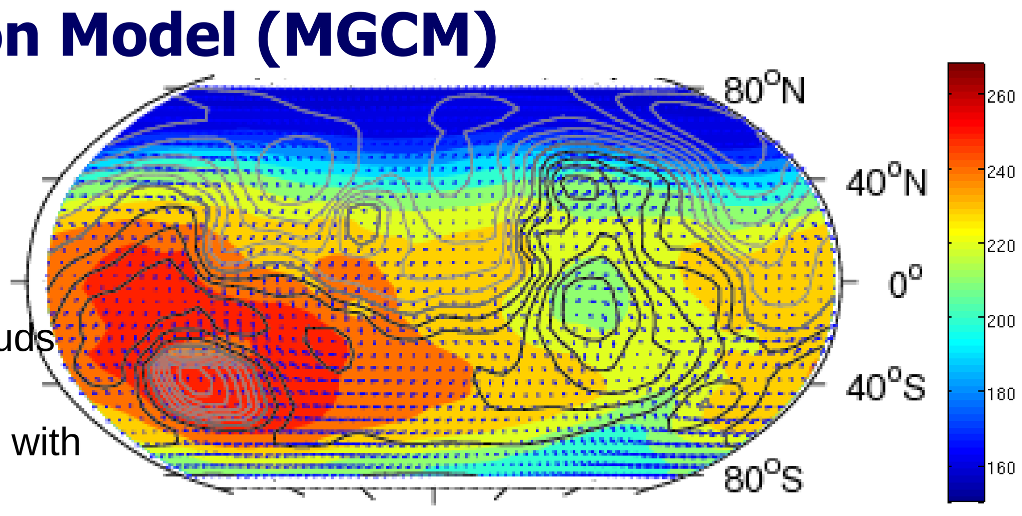
Author contact information:
Steven J. Greybush: sjg213@psu.edu

Mars Reanalysis Research Goals

- Created a 4-dimensional Mars **weather** and **climate reanalysis** by **assimilating** spacecraft **observations** into a **Mars** Global Circulation **Model**.
- Understand the characteristics and locations of any **temperature biases** between spacecraft data and the model, and improve **physical parameterizations** in Mars **models** to facilitate the match between observations and model output.
- Address scientific questions involving atmospheric **predictability**, origins of dynamical **instability**, **aerosol** (dust and ice) distribution, **traveling wave** activity, **thermal tides**, and genesis and decay of **dust storms**.

GFDL Mars Global Circulation Model (MGCM)

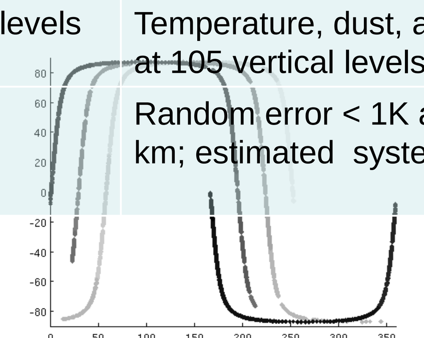
- Finite volume dynamical core
- 6°x5° (60x36) longitude-latitude resolution
- 28 vertical levels with hybrid p / σ vertical coordinate
- Radiatively active dust tracers; water ice clouds are optionally radiatively active.
- Options for interactive dust parameterization with lifting and sedimentation.



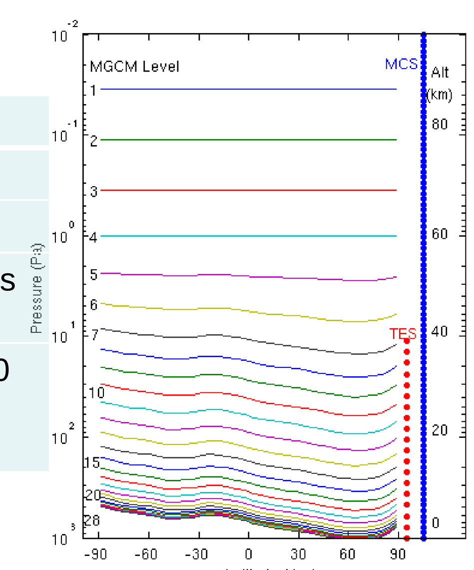
Background: NH Late Autumn ~3.5 km Temperature (Shading), Winds (Vectors), and Terrain (Contours)

Spacecraft Observations

Thermal Emission Spectrometer (TES)	Mars Climate Sounder (MCS)
Observations from 1997-2006.	Observations from 2006-present.
Nadir sounder.	Limb sounder.
Temperature retrievals at 19 vertical levels up to 40 km; column dust opacity.	Temperature, dust, and water ice retrievals at 105 vertical levels up to 80 km.
Observation error estimated at 3 K; original PDS retrieval characteristics not well known.	Random error < 1K at elevations below 50 km; estimated systematic error of 1-3 K.



Background Image: spatial coverage of TES observations (along MGS orbit) in 6 hour time period; MCS coverage is similar

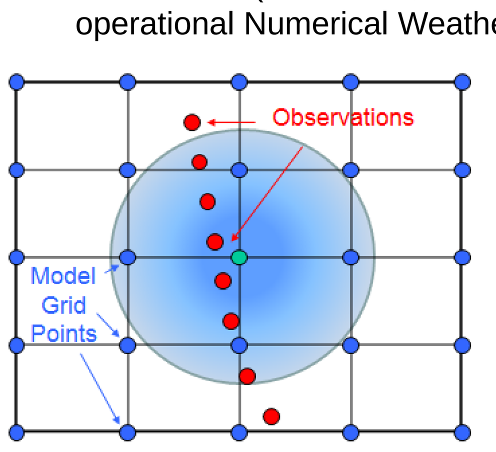


MGCM Levels and TES & MCS Observation Levels

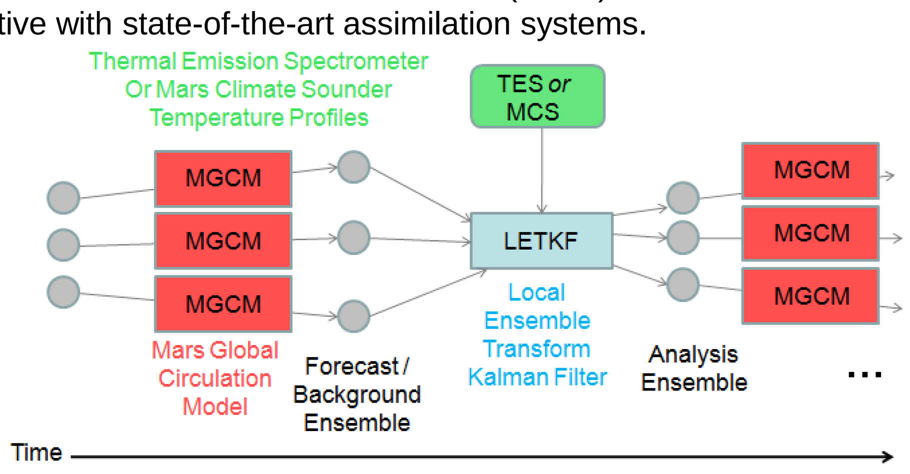
OSS retrievals for TES include error estimates and averaging kernel information, allowing us to more properly account for the retrieval process in the assimilation system.

Local Ensemble Transform Kalman Filter (LETKF)

The **LETKF** (Hunt et al., 2007) is an **efficient** implementation of the **Ensemble** Kalman Filter (EnKF) suitable for operational Numerical Weather Prediction, and is competitive with state-of-the-art assimilation systems.



Analysis at a given grid point is determined from the background at that point plus a weighted sum of observation increments within a **localization** radius.



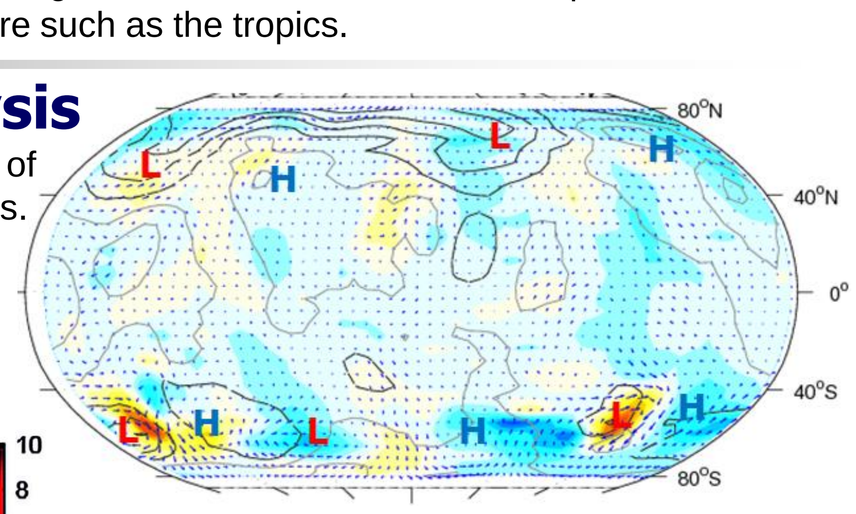
- Background, or forecast, **errors** are described by an ensemble of MGCM states, and **evolve with the flow** (an important advantage of ensemble data assimilation methods!). **Inflation** of the ensemble spread helps account for model error.
- Observation **errors** have both random and systematic components, and include instrument error and errors of representativeness. Gaussian **localization** (600 km in horizontal; 0.4 log P in vertical) of observation errors ensures that an observation's influence wanes away from the analysis grid point.

Methods for improving assimilation system performance:

- **4D-LETKF** considers observations at correct hourly timeslot.
- Superobservations combine nearby observations into a single value, and reduce the random component of observation error.
- Scaling of surface pressure increments ensures **mass conservation**.
- **Adaptive inflation** *(Miyoshi, 2010)* estimates the multiplicative inflation parameter using statistics of observation and forecast errors and ensemble spread, allowing inflation values to vary in space and time.
- **Empirical bias correction** *(Danforth et al., 2007)* of the MGCM based on analysis increments accounts for model error. Corrections (based on long-term differences between analyses and forecasts) are applied every analysis step, as if they were part of the model itself.
- **Varying** the **dust distribution** and water ice cloud strength among the **16 ensemble members** improves the ensemble spread of dynamically stable regions of the atmosphere such as the tropics.

Mars Weather and Climate Reanalysis

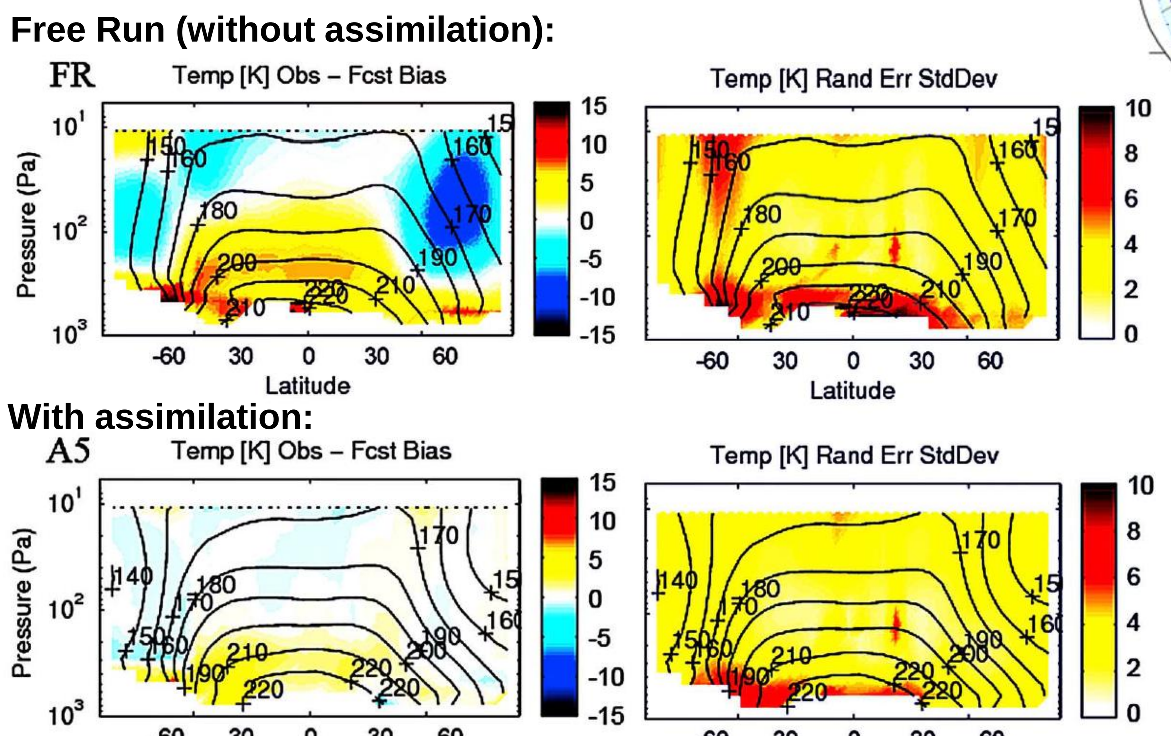
- Short term (0.25 sol) forecasts from analyses are superior to those of a freely running MGCM when compared to independent observations.
- Biases (imperfect aerosol) between the MGCM and obs are significant part of forecast RMSE.



Level 20 (~3.5 km) Eddy Winds (vectors), Temperatures (shading every 2 K), Eddy Surface Pressure (contours)

Free Run (without assimilation):

With assimilation:



Plotted are ensemble mean temperatures [K] in NH Autumn, MY 25 for TES.

Synoptic maps of eddy fields provide insights on the traveling weather systems that help initiate dust storms. We demonstrated that data assimilation analyses using different initial conditions and aerosol assumptions converge about a unique synoptic state.

Analyses are also verified against independent Radio Science profiles.

Retrieval versus Radiance Assimilation

It is better to use radiances than retrievals because

Radiances are available sooner, have uncorrelated errors, and are independent of the prior.

Retrieved temperatures contain vertical smoothing, correlated errors, and prior information.

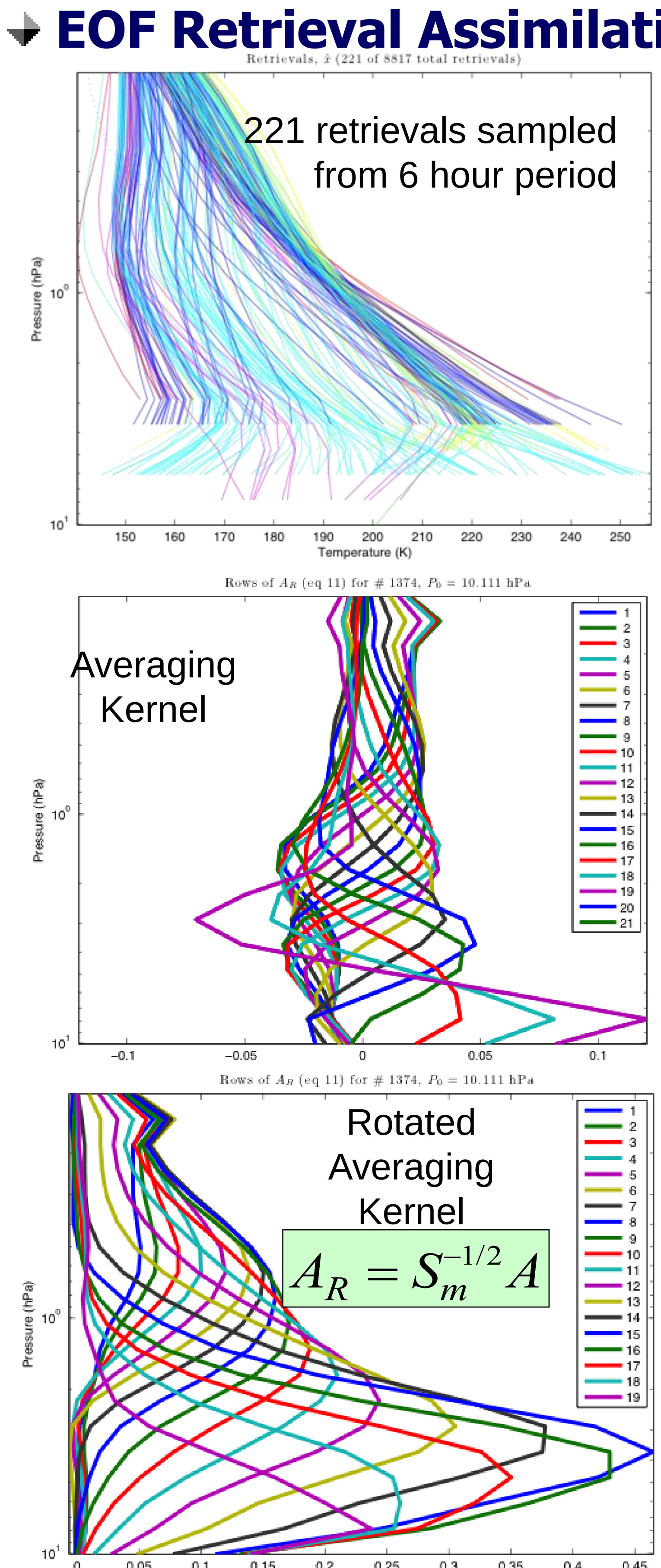
But it is better to use retrievals because

Using retrievals reduces complexity in the *H*-operator, reduces data volume, allows arbitrary cloud clearing and retrieval methods and makes the assimilation system more modular.

Using EOFs from the retrieval scheme can reduce data volume and reduces vertical interpolation errors.

The averaging kernel (AK) concept allows removing the influence of the prior (allowing for interactive retrievals) and rotating to a representation where the obs. errors are uncorrelated.

EOF Retrieval Assimilation



Averaging Kernel Method (after Rodgers, 2000)

$$\hat{x} = Ax + (I - A)x_a + G_y\varepsilon_y + \varepsilon_r$$

- Here $\hat{x}$ is the retrieval, x is the true state vector, and x_a is the prior. A is the averaging kernel. ε_y includes measurement error and forward model error which is unbiased with covariance matrix, S_ε. Representativeness error, ε_r, is assumed to be additive, unbiased, uncorrelated with ε_y and to have a covariance S_r. G_y is the sensitivity of retrieval to the radiances

$$G_y = \frac{\partial \hat{x}}{\partial y} = K^T S_\varepsilon^{-1} K + S_a^{-1}$$

Here S_a is the prior covariance and

$$S_q = K S_a K^T + S_\varepsilon$$

- Also $K = \frac{\partial F}{\partial x}$ is the Jacobian of the forward problem, F, evaluated at the solution $\hat{x}$. Finally

$$A = G_y K = \hat{S}(K^T S_\varepsilon^{-1} K)$$

$\hat{S}$ is the retrieval error covariance

Required Data

- Inputs to the retrieval
  - Prior mean and covariance and radiance error covariance

    x_a, S_a, and S_ε
- Outputs from the retrieval
  - Retrieval and Jacobian

    $\hat{x}$ and K
- OR, if the channel set is large since $A = I - \hat{S}S_a^{-1}$
  - Prior mean and covariance

    x_a and S_a
  - Retrieval and posterior covariance

    $\hat{x}$ and $\hat{S}$

Data Assimilation Interface

- Within the data assimilation, define the new observation as

$$\hat{y}_A = \hat{x} - (I - A)x_a$$

and the new observation operator by

$$y_A = Ax$$

We will now be comparing observed and simulated quantities with the same degree of smoothing. Now the observation increments

$$\hat{y}_A - y_A = G_y\varepsilon_y + \varepsilon_r$$

are unbiased if the ε_y are unbiased, and have covariance

$$S_m = G_y S_\varepsilon G_y^T + S_r$$

- If we scale by $S_m^{-1/2}$ we effectively rotate to a space where the observation errors are unbiased, uncorrelated, and have unit variance. The new observation is

$$\hat{y}_A = S_m^{-\frac{1}{2}}(\hat{x} - (I - A)x_a)$$

and the new observation operator by

$$y_A = S_m^{-\frac{1}{2}} Ax = A_R x$$

where

$$A_R = S_m^{-\frac{1}{2}} A$$

S_m is covariance of the errors of the new observations

Forecast as Prior (Interactive Retrievals)

- Since the observation increment

$$\hat{y}_A - y_A = G_y\varepsilon_y,$$

does not depend on the prior, we may use the ensemble background mean as x_a and the ensemble sample covariance as S_a.

This is important since the quality of the Mars retrievals depends heavily on the prior.

Implementation

The LETKF H-operator has been generalized:

First, interpolate the model temperature to the retrieval location and pressure levels

Second, if using EOFs remove an overall mean temperature profile

Third, applies weights to each temperature to compute observation space quantities

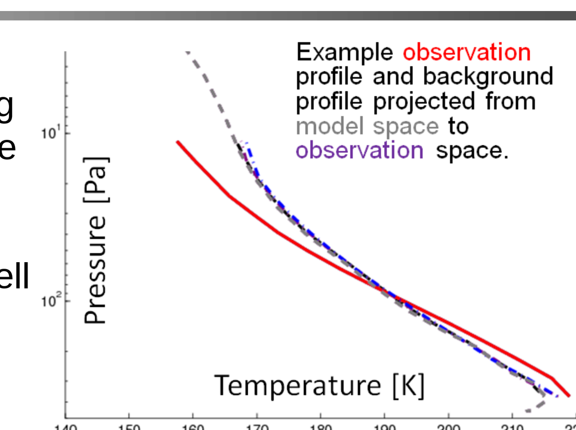
These weights are part of the observation data structure

For example, in the original assimilation strategy the weight is 1 at the observation level and 0 elsewhere; or we can apply a box car average; or any other linear combination of the temperatures

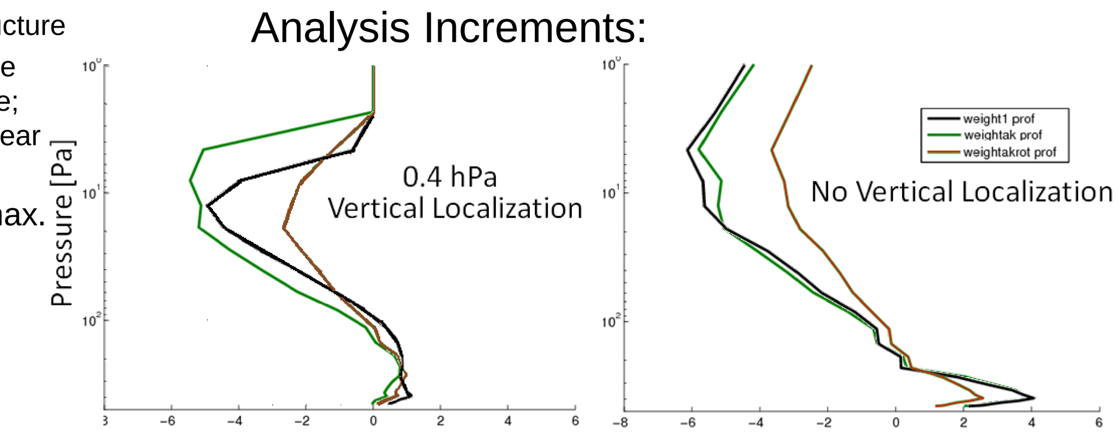
Fourth, for localization assign to vertical level of max. weight.

We consider in our methodology:
- Representativeness errors.
- Levels in the retrieval pressure grid below the surface.
- Superobservations.

Here, we provide a proof of concept by assimilating a single OSS temperature profile in the **traditional way**, in the **averaging kernel framework**, as well as using the **rotated averaging kernel**.



Analysis Increments:



Summary and Ongoing Work

- Convert standard retrievals into "observations" with expected errors that should be zero mean, uncorrelated, and unit variance, and independent of the background or prior.
- Define a corresponding obs-function (or H-operator) that is a weighted sum of the temperatures on the radiative transfer model vertical grid.
  - No changes to the assimilation method are needed, except to interpolate to the radiative transfer model vertical grid and to calculate the weighted sum.
- Projecting onto EOFs used by the retrieval can reduce the number of observations
- We plan to compare to radiance assimilation using the OSS forward model, as well as conduct "interactive retrievals" using forecasts from data assimilation as a prior.
- Based on ideas in Rodgers' book, "Inverse Methods for Atmospheric Sounding: Theory and Practice".

For further details:

Hoffman, R. N., 2010: A retrieval strategy for interactive ensemble data assimilation. *arXiv*, (1009.1561v1 [physics.ao-ph]), 1–13, http://arxiv.org/abs/1009.1561.

Greybush, S. J., E. Kalnay, K. Ide, T. Miyoshi, T. McConnochie, M. J. Hoffman, R. N. Hoffman, and R. J. Wilson, 2012: Ensemble Kalman Filter Data Assimilation of Thermal Emission Spectrometer (TES) Profiles into a Mars Global Circulation Model. *J. Geophys. Res. Planets*, 117, E11008, doi: 10.1029/2012JE004097.

Acknowledgements: NASA grant NNX07AM97G, NNX11AL25G; MCS Team, UK Mars Group, Takemasa Miyoshi, Kayo Ide, Janusz Eluszkiewicz, Dan Gombos, Timothy McConnochie, Luca Montabone