ENSEMBLE-BASED DATA ASSIMILATION WITH MAPPING DATASETS OF THE MARTIAN ATMOSPHERE. W. G. Lawson[1] and M. I. Richardson[2], [1]California Institute of Technology (1200 E. California Blvd, Pasadena, CA 91125; wglawson@gps.caltech.edu), [2]Caltech (mir@gps.caltech.edu).

What is Data Assimilation?: Data assimilation (DA) is the formalized approach to blending and organizing the information content of many, disparate observations by way of a dynamical model. In the case of atmospheric science, the model is a general circulation model (GCM), and it provides the rules (i.e., physics) for how one given state of the atmosphere evolves into the next. The state of the atmosphere contains the values of all prognostic variables at all discretized locations within the model domain (or equivalently, their spectral components). The set of prognostic variables generally contains those of the primitive equations for fluid evolution (momentum components, temperature, density, and pressure) potentially coupled with other interactive systems (e.g., hydrological cycle, radiatively-active dust, passive tracers). It is precisely these variables with which atmospheric scientists prefer to work, and it is ideal to have them specified at the regular, spatial resolution of a GCM (e.g., for computing fluxes). And it is precisely this ideal product that DA aims to generate by providing a time sequence of model states ("analyses") that agree with a relevant mapping dataset to within its stated uncertainty. If the model represents the time sequence of known observations well, then by its encoded laws of physics one can better believe its evolution of unobserved quantities.

Our Goal for Planetary Atmospheres: Many planetary meteorological / atmospheric instrument concepts and proposals since the early 1990s have emphasized the role that DA should play in the synthesis and analysis of the resulting data. By its very nature, DA has always been an enormous undertaking, both in its required computational power and person-power: early attempts within the planetary community were forced to make simplifying assumptions. However, with the advent of relatively cheap parallel computing clusters and the recent development of Monte Carlo (ensemble-based) DA algorithms within the terrestrial community [1], [2], [3], the prospect of implementing sophisticated (i.e., making less severe assumptions) DA has been placed within the grasp of the relatively small planetary atmospheric community.

We have embarked on a multi-year project to implement an ensemble-based DA system within a planetary context, namely Mars. One goal is to provide a road-map, written guidance, and a worked example, all of which will be made publicly available, for other planetary atmospheric groups who may want to implement DA in the future for their specific GCM of interest and their specific mapping dataset of interest (and their planet of interest). Another goal, a consequence of the first, is that having gone through the DA exercise for a planet that has been examined with earlier, more approximate methods, we will be able to gauge what, if any, advantages the more complicated approach gives over the earlier attempts. Another consequence is that we will have produced a DA product that should lend itself well to thorough analysis for the characterization of the Martian general circulation.

Prior Work: DA has seen three separate attempts for Mars, all using the MGS / TES dataset:

1. "Steady State Kalman Filter" – This technique was developed by observing, through a hierarchy of models, that the corrective structure (the "gain") of observations from a single polar orbiting satellite obtains an approximately constant shape in the frame of reference of the satellite [4]. Pre-calculating this steady state structure (from a training dataset) alleviates a large amount of computational work when implementing DA, thus making DA a tractable endeavor. However, it was observed that as models got more complex, there was potential for multiple different steady state structures – one for night and another day, one for dust storm conditions and another for non-dust storm conditions. When the approach was finally applied to a GCM using the *T(p)* nadir retrievals, the results were mixed [5].

2. "Analysis Correction Scheme" – This technique was judiciously adapted from an operational scheme used by the UK Met Office [6], [7]. This scheme also alleviates much of the computational burden of DA by effectively specifying the corrective structure of the observations, both in space and time; however, instead of pre-calculating these structures, they are specified *a priori*. Fortunately, the relevant dynamical length and time scales in the Martian atmosphere are not so different from those on Earth, so guidance for the shapes of these structures could also be taken from the Met Office. The Analysis Correction Scheme has been applied to essentially all of the nadir retrievals from the TES dataset, including those from aerobraking [8]. To date, the set of produced analyses has been: a.) verified against MGS / Radio Science retrievals [9], b.) used to study interannual variability of dust storms [10], c.) used to study atmospheric tides [11], and d.) used to infer that GCMs are very likely to be missing water ice cloud forcing effects in the tropics [12]. Overall, the analyses seem to be of good quality, and they are beginning to be used in the manner that DA products are designed.

3. "4D-Var with a truncated model" – 4D-Var is a powerful DA technique employed by terrestrial operational forecasting centers. It essentially solves a very large four-dimensional least squares problem by way of a numerical model's so-called tangent linear model and its adjoint. It is notoriously computationally expensive when used with GCMs, particularly in the required overhead to get a GCM's adjoint, and it seems its implementation in planetary science was afforded by truncating the model used to a manageable number of degrees of freedom, perhaps as few as could be directly constrained by the available data [13]. 4D-Var is an encouraging direction for planetary atmospheric science, and its initial results were encouraging as well. Unfortunately, not much has been published with this method since.

Moving Forward: By necessity, the prior attempts outlined above were forced to make some simplifying assumptions that may not have been warranted. In particular, experience from modern terrestrial DA shows there is a huge benefit from evolving the corrective structures of observations (as opposed to pre-calculating or specifying their structures). It may be the case that this is not as important on Mars, but it is worth exploring these potential limitations, especially given that ensemble-based approaches makes such exploration feasible.

In addition, the development of an ensemble-based DA system for planetary use would serve at least three other purposes: 1.) it automatically offers flow-adaptive, time-varying uncertainty estimates for its produced time sequence of analyses, 2.) it is natively modular, which would ultimately serve the greater good of the community allowing for ease in "swapping out" one's particular GCM or mapping dataset, and 3.) it easily accommodates observations that have complicated relationships to a model's state, the most relevant example being radiances (related to a model's state through a radiative transfer model mimicking a given instrument's spectral characteristics and viewing geometry). In reference to this last point, all one needs is a given instrument's "Forward Operator," which is a standard component in any attempt to make retrievals from a dataset of radiances – one does *not* need the forward operator's tangent linear or adjoint code.

Progress Report: We have forged a relationship with the National Center for Atmospheric Research (NCAR) in order to become users and developers of DART, the Data Assimilation Research Testbed (http://www.image.ucar.edu/DAReS/DART). DART is a suite of publicly available, open source, developed and supported software built upon the modular and generalizable nature of ensemble-based DA. DART was created specifically with the goal of enabling DA for a great many applications, mainly by handling the common "overhead" shared by essentially all DA problems – this includes management of algorithms, computer communications, etc. Further, DART is aimed specifically at research applications, and thus offers a wide variety of tools to help users adapt DA to their idiosyncratic applications. For example, a user can try up to 8 different ensemble DA algorithms, including filters and smoothers. The main work required for typical users to begin using DART is to make their models and observations (via the appropriate forward operators) "DART-compliant." This requires generating several Fortran 90 module routines and interfaces that perform basic functions (e.g., one needs to know how to interpolate within one's model state).

Planetary atmospheric applications of DA present several challenges that are still subjects of active research within the DA community. Most notably, the main observations of use contained within mapping datasets are radiances. The terrestrial DA community reported a large benefit from assimilating direct radiance observations instead of retrieved products made from those radiances [14], [15]; however, using radiances in the newer, ensemble-based DA approaches is still a relatively unexplored problem. Among the reasons for these benefits is that radiances have more plausibly independent uncertainties, whereas retrieval processes inevitably introduce correlated uncertainty structures (e.g., if temperature is in error at one pressure level, it is likely in error with the same sign at an adjacent pressure level). Additionally, planetary datasets are typically quite sparse (compared to terrestrial data streams) and taken nearly continuously (as opposed to via a planet-wide agreement for coordinated, simultaneous release of radiosondes on a "synoptic time" interval). With the help of the NCAR / DART team (J. L. Anderson, N. Collins, C. Snyder, and Y. Chen), we plan to work through these research challenges and bridge the DA-gap even further so that other planetary scientists may implement DA as part of their own research programs in the future.

The PlanetWRF GCM is our most readily available model with which to work [16]. We have already made the Earth-version of PlanetWRF DART-compliant, and we have actually treated it as a numerical weather prediction model by assimilating hundreds of thousands of operational terrestrial atmospheric observations. As of writing this abstract, we have just made the Mars version of the model DART-compliant.

To begin our testing and experimentation with a planetary mapping dataset, we have selected the MGS / TES nadir observations due to their simpler viewing geometry. We gratefully acknowledge the help provided by the TES atmospheric retrieval team (M. D.

Smith, M. S. Kaelberer, and B. J. Conrath) in both getting us the appropriate TES nadir forward operator code and for further support to ensure we were using it correctly. We have already made the forward operator code DART-compliant. After working with the TES dataset and addressing some of the outstanding questions regarding radiance assimilation with an ensemble-based DA approach, we intend to work with the more complicated viewing geometry limb observations provided by the Mars Climate Sounder aboard MRO. We thank the MCS team (D. J. McCleese, J. T. Schofield, D. M. Kass, A. Kleinböhl, and W. Abdou) for their continued help and for making their limb viewing forward operator available to us.

Our presentation will give an overview of our project and a progress report on results obtained to date as of the meeting in November.

Teaser Comparison: Working with a GCM and radiance observations allows for model / data comparisons that are not often seen. The two figures below show one such juxtaposition, namely the view of the Martian atmosphere as seen by the TES 667.70 cm^{-1} channel in nadir viewing. The data included in Figure 1 comes from the MGS / TES dataset in its first Mars year of operation, only the data included within L_S = 140°-150°. The data has been separated according to two local time bins: the top panel includes observations with a local time (LT) between 0100 and 0300 and the middle panel includes those with LT between 1300 and 1500 (note the drop out of LT coverage near the poles). The data has been further binned to the horizontal spatial resolution of a typical GCM run, 64 longitude bins (5.625°) and 36 latitude bins (5°). The top two panels of Figure 1 show the mean values of the brightness temperatures (T_B) in each of the bins. The bottom panel shows the difference in T_B between the two LT bins. The afternoon T_B is evidently cooler over the equator and warmer in the mid-latitudes. While not the focus of this teaser section, the main cause of this pattern is likely the diurnal tide [17].

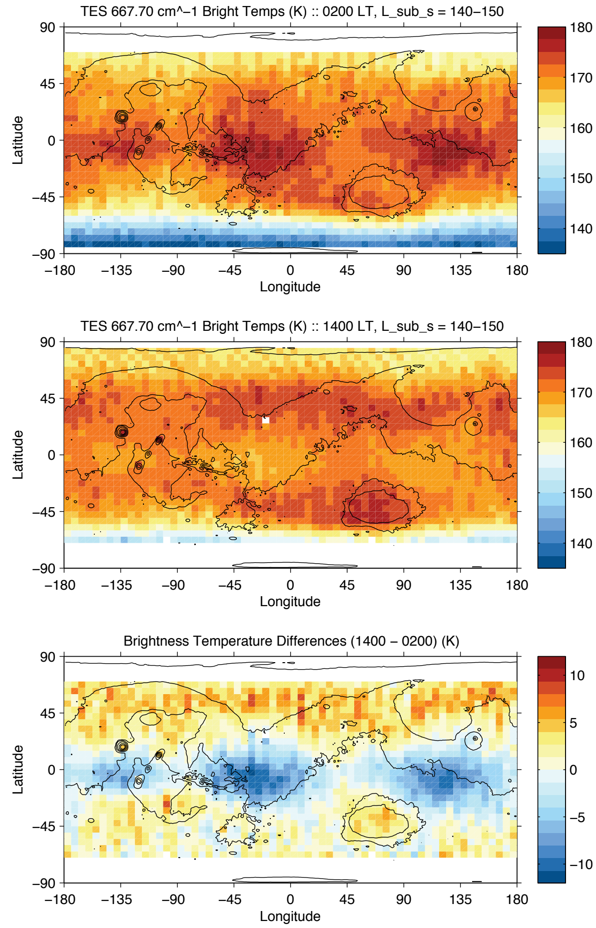

Figure 1: *a.*) shows mean brightness temperatures (K) of the TES 667.70 cm^{-1} channel over the period L_S = 140°-150° for observations with LT between 0100-0300. *b.*) shows the same as *a.* but for LT between 1300-1500. *c.*) shows the difference in brightness temperatures of panels *a.* and *b.*

Figure 2 shows analogous plots to those of Figure 1, except that instead of using actual MGS / TES data, we have applied the TES forward operator to the output from a typical MarsWRF GCM run. Note that no data assimilation has been performed as of yet, meaning no special attempts have been made to get the model to match the particular data presented in Figure 1. The model output was filtered into the same LT bins and examined over the same L_S range. We find that the model's predicted values for both 0200 and 1400 T_B are smaller than those observed. Similarly, the model's predicted range for the differences in T_B is less than observed. However, it is pleasing to see that the spatial patterns of these fields are rather similar, indicating that the model clearly has some some skill (and that the TES forward operator is likely working). It will be the job of DA to make the model's view of the TES 667.70 cm^{-1} T_B come into better agreement with the observations. DA will accomplish this by using the information contained within the observations to adjust the model's estimates of the mass, thermal, and aerosol fields of the atmosphere (i.e., the prognostic state of model). The changes necessary to make this agreement will give insight to the dynamical evolution of the Martian atmosphere.

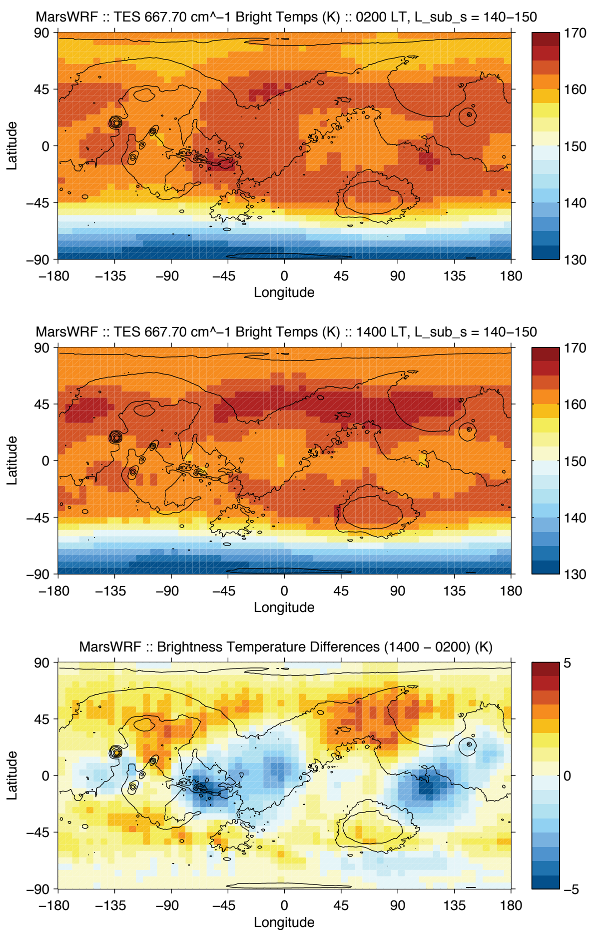

Figure 2: The same information is plotted here as in Figure 1, except that the data plotted is from having applied the TES forward operator to the MarsWRF GCM. Otherwise, the data has been filtered in LT and binned in the same fashion.

Acknowledgments: We thank NASA's Applied Information Systems Research Program for their support.

References: [1] Evensen G. (1994) *JGR-Oceans*, *99*, 10143-10162. [2] Anderson J. L. et al. (2005) *J. Atmos. Sci.*, *62*, 2925-2938. [3] Houtekamer P. L. et al. (2005) *Mon. Wea. Rev.*, *133*, 604-620. [4] Banfield D. et al. (1995) *J. Atmos. Sci.*, *52*, 737-753. [5] Zhang K. Q. et al. (2001) *JGR-Planets*, *106*, 32863-32877. [6] Lorenc A. C. et al. (1991) *Q. J. R. Meteorol. Soc.*, *117*, 59-89. [7] Lewis S. R. et al. (1996), *Planet Space Sci.*, *44*, 1395-1409. [8] Lewis S. R. et al. (2007), *Icarus*, *192*, 327-347. [9] Montabone L. et al. (2006), *Icarus*, *185*, 113-132. [10] Montabone L. et al. (2005), *Adv. Space Res.*, *36*, 2146-2155. [11] Lewis S. R. and P. Barker (2005), *Adv. Space Res.*, *36*, 2162-2168. [12] Wilson R. J. et al. (2008), *Geophys. Res. Lett.*, *35*, L07202. [13] Houben H. (1999), *Adv. Space Res.*, *23*, 1899-1902. [14] Eyre J. R. et al. (1993), *Q. J. R. Meteorol. Soc.*, *119*, 1427-1463. [15] Andersson E. J. et al. (1994), *Q. J. R. Meteorol. Soc.*, *120*, 627-653. [16] Richardson M. I. et al. (2007), *JGR-Planets*, *112*, E09001. [17] Banfield D. et al. (2003), *Icarus*, *161*, 319-345.