# DATA ASSIMILATION OF MARS GLOBAL SURVEYOR METEOROLOGY.

Howard Houben, *Bay Area Environmental Research Institute (houben@humbabe.arc.nasa.gov).*

The wealth of atmospheric data from the Mars Global Surveyor (MGS) mission affords us an unprecedented opportunity to test our theoretical knowledge of the physical properties and dynamical state of the atmosphere of another planet. The primary tool for translating between observable quantities (like asynoptically measured infrared radiances) and meteorological variables (like winds and temperatures) that may be computed by a general circulation model (GCM) is data assimilation. This term encompasses a number of related techniques by which data are used to determine a statistical best-fit state of a predictive model. In the Martian context, where many significant components of the climate system have never been directly observed, data assimilation is an important part of the scientific discovery process.

The formulation of meteorological data assimilation is now becoming standardized [Ide et al., 1997]. In terms of operator notation, we represent the predictive dynamical model by $M$ and the observation operator (a forward model that projects from a model's state into the space of observable variables) by $H$. A vector of observations is given as $y$ and the model state control vector by $x$. The variational formulation of the data assimilation problem then seeks to minimize the cost function

$$J = \frac{1}{2}(x-b)^T B^{-1}(x-b) + \frac{1}{2}(HMx-y)^T R^{-1}(HMx-y), \quad (1)$$

where $b$ is a previously computed model background state, $B$ is the model forecast error covariance matrix, and $R$ is the observational error covariance matrix. The minimum of $J$ is obviously obtained at

$$\partial J/\partial x = B^{-1}(x-b) + M^T H^T R^{-1}(HMx-y) = 0. \quad (2)$$

Note that the matrix operator $M$ which represents the dynamical model need not be calculated explicitly. In practice, $M$ is a formally linearized version of the computer code—a tangent linear model—which is used to calculate a model state vector from a perturbation of the state at the previous step. (When the model is a GCM, the accuracy of the linearized dynamical core can be assessed quite precisely, since terms in the equations of motion have known linear or quadratic dependencies on the perturbations.) $M^T$ is similarly a formally adjointed version of the tangent linear code [Talagrand, 1997]. It is tested by verifying that the adjoint definition $\langle Mx, Mx\rangle = \langle x, M^T Mx\rangle$ is satisfied to machine precision. Equation (2) is solved by iterative techniques (like the conjugate gradient or quasi-Newton methods) that make use of the vectors $Mx$ and $M^T Mx$ but require no other information about the operator $M$ itself.

It makes sense to distinguish between the practical use of data assimilation in the terrestrial and Martian contexts. For operational forecasting at numerical weather prediction centers, there is usually a very well characterized, high resolution model available. The state of the atmosphere is also fairly well predicted from previous forecasts. The amount of new data at any forecast step is small compared to the number of model variables. So, the solution $x$ of equations (1) and (2) is generally quite close to $b$. The forecast error covariance matrix $B$ is typically used to impose a priori balances on the solutions (which suppress such unwanted features as gravity waves). This is necessary because the number of model degrees of freedom far exceeds the constraints imposed by the data.

With the abundant data from the MGS Thermal Emission Spectrometer (TES), on the other hand, the number of new independent observations per day is comparable to or greater than the dimensionality of a usable global Mars GCM, so the minimization problem of equation (1) is typically overdetermined. Nearby observations can be averaged to produce statistical estimates of the observational error covariance matrix $R$. It is not generally possible to impose a priori balances (because unbalanced components of the circulation—like atmospheric tides—and large variations in surface pressure are so important in Martian dynamics). However, these additional constraints may not be required because of the overdetermined nature of the problem. Also, the forecast errors may be large compared to observational errors and equation (1) then reduces to a $\chi^2$ minimization where

$$\chi^2 = (HMx-y)^T R^{-1}(HMx-y). \quad (3)$$

The influence of the forecast on the ultimate solution then comes about primarily because it is the first guess in the iterative solution procedure. As experience in the solution of equation (3) is gained, information about the structure of $B$ is also accumulated. Note that the control vector $x$ may contain model parameters and calibration constants along with the initial state of the atmosphere. So the assimilation can be used to improve knowledge of the model and the observing system.

A Mars GCM especially suited for the assimilation of MGS data has been developed [Houben, 1999]. It is a form of the baroclinic spectral model that has long been used in terrestrial meteorology [Bourke et al., 1977; Haltiner & Williams, 1980; Krishnamurti et al., 1998]. The model conveniently divides the terms in the equations of motion into linear and higher order terms. While this is done primarily to enable a semi-implicit integration scheme that allows for longer time steps, this treatment also speeds the production and testing of the required linear tangent model. The spectral formulation allows for an easy truncation of the model to low order (i.e., to zonal wavenumber 6) which corresponds to the observing pattern of the MGS orbit. With 16 Legendre functions in latitude and 16 vertical levels, the model has of order 10,000 prognostic variables that must be specified by the assimilation procedure.

The baroclinic spectral model uses a terrain-following sigma coordinate in the vertical direction. This allows the MOLA topography to be incorporated in the surface boundary condition. A bulk Richardson number-based boundary layer scheme parameterizes surface-atmosphere interactions.

The base model uses a simplified Newtonian relaxation radiative forcing scheme (with a parameterized diurnal cycle), as has been used in climate modeling studies of Mars [Haberle et al., 1997] where the speed of integration is more important than the detailed simulation of atmospheric eddies. This is clearly not appropriate for data assimilation where the amplitudes and phases of wave modes are of the utmost importance. It would be theoretically possible to specify the diabatic forcing of our GCM with a full radiative transfer code (and, of course, a comprehensive boundary layer scheme). However, this would require detailed knowledge of the distribution of atmospheric tracers like dust, water vapor, and associated clouds. Even the extensive MGS observations are not adequate to specify all of these parameters based on current models. (Development of tracer assimilation models is an active field of research.) Instead, we have chosen an innovative approach to determining the proper diabatic forcing for our GCM.

Rather than specifying the diabatic forcing of the Martian atmosphere, we have included it (along with the initial temperature, streamfunction, divergence, and surface pressure) as part of the control vector $x$ that is solved by the assimilation procedure. This inclusion of forcing in the control vector is equivalent to using the model as a weak constraint on the assimilation (or the assumption of an imperfect model) [Courtier, 1997]. This is quite appropriate for our current state of knowledge of the Martian atmosphere. Without this approach, we would be confronted with the problem that while the diabatic forcing is poorly known, the atmosphere responds rapidly to the specified forcing. It would therefore be difficult to steer the model towards the observations (rather than back to the original model state).

The TES Team have kindly provided temperature retrievals for one detector for 630 sols during the MGS mapping year. (For convenience, a "sol" is a period of 12 MGS orbits, a bit shorter than the standard definition.) In order to obtain good characterizations of the data, temperature profiles were averaged over thirty second intervals (during which time MGS travels approximately 1.5 degrees in latitude) and standard deviations were computed (to be used as statistical weights). This dataset has been used (with our baroclinic spectral model) in the $\chi^2$ variational assimilation scheme described above (equation [3]). Analyzed temperature fields agree with the retrievals to about 2 K. One-sol are accurate to about 3.5 K on the sunlit parts of the planet (where TES retrievals are thought to be more accurate. The question of how well the assimilation of diabatic forcing works is addressed by the biases in the 1-sol forecasts. Since the total radiative forcing is on the order of 10 K/day, small biases in the forecasts (less than 1 K) indicate that the assimilated forcings are reasonable.

Given the success in assimilating TES retrieved temperatures, a forward model for the prediction of TES radiances based on baroclinic spectral model temperature profiles has been developed. The model resembles that used by Conrath et al. [2000] (utilizing correlated $k$'s, etc.), but the constraints for retrieval of the spectra come from the data assimilation cost function rather than any ab initio assumptions. The radiances in the wings of the 15 micrometer band are quite sensitive to surface temperature, so a sophisticated ground temperature prediction model is required. Thus albedo and thermal inertia are included in the control vector for radiance assimilations. Agreement with observations is obtained to within about 1 erg/cm$^2$/s/sr/cm$^{-1}$. Based on the successful development of these data assimilation techniques for use with MGS data, a wide range of studies comparing the results from different experiments and examining the resulting meteorology are possible.

References


Bourke, W., B. McAvaney, K. Puri, and R. Thurling, 1977. Global modeling of atmospheric flow by spectral methods, in *General Circulation Models of the Atmosphere,* ed. J. Chang, *Methods in Computational Physics,* **17,** pp. 267–324, Academic Press, New York.

Conrath, B. J., J. C. Pearl, M. D. Smith, W. C. Maguire, P. Christensen, S. Dason, and M. S. Kaelberer, 2000. Mars Global Surveyor Thermal Emission Spectrometer (TES) observations: Atmospheric temperatures during aerobraking and science phasing, *J. Geophys. Res.,* **105,** 9509–9519.

Courtier, P. 1997. Dual formulation of four-dimensional variational assimilation, *Quart. J. Roy. Meteor. Soc.,* **123,** 2449–2461.

Haberle, R. M., H. Houben, J. R. Barnes, and R. E. Young, 1997. A simplified three-dimensional model for Martian climate studies. *J. Geophys. Res.,* **102,** 9051–9067.

Haltiner, G. J., and R. T. Williams, 1980. *Numerical Prediction and Dynamic Meteorology,* Second Edition, John Wiley & Sons, New York, 477 pp.

Houben, H., 1999. Assimilation of Mars Global Surveyor meteorological data, *Adv. Spa. Res.,* **23,** 1899–1902.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997. Unified notation for data assimilation: Operational, sequential, variational, *J. Meteor. Soc. Japan,* **75,** 181–189.

Krishnamurti, T. N., H. S. Bedi, and V. M. Hardiker, 1998. *An Introduction to Global Spectral Modeling,* Oxford U. Press.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992. *Numerical Recipes in FORTRAN, the Art of Scientific Computing,* Second Edition, Cambridge U. Press, 963 pp.

Talagrand, O., 1997. Assimilation of observations, an introduction, *J. Meteor. Soc. Japan,* **75,** 191–209.