The Next Generation of Planetary Atmospheric Probes

Howard Houben
Bay Area Environmental Research Institute
MS 245–3, Ames Research Center, Moffett Field, CA, 94035–1000, USA
houben@humbabe.arc.nasa.gov

Abstract

Entry probes provide useful insights into the structures of planetary atmospheres, but give only one-dimensional pictures of complex four-dimensional systems that vary on all temporal and spatial scales. This makes the interpretation of the results quite challenging, especially as regards atmospheric dynamics. Here is a planetary meteorologist’s vision of what the next generation of atmospheric entry probe missions should be: Dedicated sounding instruments get most of the required data from orbit. Relatively simple—and inexpensive—entry probes are released from the orbiter, with low entry velocities, to establish ground truth, to clarify the vertical structure, and for adaptive observations to enhance the dataset in preparation for sensitive operations. The data are assimilated onboard in real time. The products, being immediately available, are of immense benefit for scientific and operational purposes (aerobraking, aerocapture, accurate payload delivery via glider, ballooning missions, weather forecasts, etc.).

Key words: atmospheric probes, data assimilation.

1. Introduction

There are eight planetary atmospheres in the solar system exhibiting a wide range of behaviors, with variations on many spatial and temporal scales. The planets are conveniently grouped into pairs, making the solar system an excellent laboratory for the study of comparative atmospheres and associated scientific questions dealing with fluid dynamics, planetary evolution, climate change and predictability, and habitability. The challenge is to find methods to address the wide range of questions involved in a useful scientific manner. Entry probes working in coordination with orbital sounders can be an important part of this process.

Earth and Mars both have optically thin atmospheres that are primarily forced from below by the absorption of solar radiation at their solid surfaces. It is not surprising that these are the best studied atmospheres in the solar system. But many questions remain unanswered, particularly related to climate change and predictability. In the case of the Earth, there is a major international effort (with political ramifications) to predict changes in climate over the next hundred years. While the martian meteorology appears to be quite repeatable during the aphelion season [1], major planetwide dust storms occur on an irregular basis during some perihelion seasons [2]. Mars has received the most exploration attention of any planet besides the Earth, of course. But, the PMIRR instrument that was intended to be a dedicated atmospheric sounder was flown on two spacecraft that failed to achieve orbit (Mars Observer and Mars Climate Orbiter). A French mission with a microwave sounder (MAMBO) was canceled.

Venus and Titan have optically thick atmospheres, forced from above. Both appear to super-rotate with zonal wind speeds much greater than the rotation speed of the underlying solid surface. The source and mechanism of angular momentum transfer for these super-rotations are still uncertain. Both atmospheres are to be studied in great detail in the near future, by the Venus Express and Cassini-Huygens Probe missions, respectively.

Jupiter and Saturn have deep atmospheres (with no solid underlying surface). They have internal heat sources greater than the solar heating, and no large equator-to-pole temperature gradient. Yet, small scale meridional gradients in temperature are consistent with a prominent structure of (relatively clear) belts and cloudy zones with alternating easterly and westerly winds. (Whether these are related to a deep interior circulation is unknown.) Both planets have strong prograde equatorial jets. Large vortices like Jupiter’s Great Red Spot have fascinated observers for as long as telescopes have been available, yet the nature of the red chromophore is still unknown. The bulk compositions of Jupiter and Saturn (e.g., helium to hydrogen ratios) are of great astrophysical interest [3]. How these relate to the atmospheric values may...
provide a key to their interior structures and evolution. The Galileo orbiter’s atmospheric observations of Jupiter were curtailed because of the loss of the high gain antenna. Cassini made some observations of the jovian atmosphere during its flyby and will have the opportunity to make many observations of the saturnian atmosphere.

At first glance, the ice giants Uranus and Neptune appear to be smaller versions of Jupiter and Saturn. But they are characterized by relatively small internal heat fluxes and external forcing. That Neptune’s zonal jets are the largest measured on any planet is another mystery.

Although there have been a number of spacecraft missions that have made observations of these atmospheres, none (besides the Earth) has been studied systematically from the point of view of a meteorologist. Such systematic study would require global coverage over a radiative time scale with an effective spatial resolution adequate to resolve the radius of deformation. (See the discussion and table in the next section.) We are receiving a large number of atmospheric measurements from the Mars Global Surveyor (mostly from the Thermal Emission Spectrometer, TES) and several instruments on the European Mars Express. It should be noted, however, that TES was primarily designed to look at martian surface composition. Its coarse resolution in the 15 micrometer carbon dioxide band allows the temperature structure of the atmosphere to be determined with a broad weighting function. (Somewhat better vertical resolution is obtained from the less frequent limb scans that sacrifice horizontal resolution.) On the other hand, the highly elliptical orbit of Mars Express gives only limited coverage of the planet over the course of a day. Both missions have taught (and are teaching us) a lot about Mars and about how best to study planetary atmospheres. No Mars Express data are in the public domain, however, and it is frustrating to think of doing meteorology (the value of which clearly degrades with time) at a remove of many months or a year. These missions also serve as excellent testbeds for the Mars Climate Sounder—which should be the first dedicated planetary atmospheric sounder in a suitable orbit for specifying the global meteorology—that is scheduled to be launched at the next opportunity (August 2005).

Of course, there have been a number of exciting probe missions to the planets. The Viking, Pathfinder, and Mars Exploration Rover missions all did entry, descent, and landing science. In spite of the quality of the data, these entry profiles are swamped in number by radio occultation temperature profiles from Viking and Mars Global Surveyor. That the Viking profile was relatively warm compared to Earth-based measurements and subsequent entry profiles has been remarked, but the implications are not clear [4]. The Pathfinder profile was measured at night and how much of its colder temperatures should be attributed to day-night differences is an open question. (Near-surface wind measurements deduced from the Viking entries show a spiral opposite in sense to what would be expected in an Ekman boundary layer, but this issue has also received little attention [5].) In general, the interpretation of probe measurements must deal with representativeness questions of this kind. They offer at best a one-dimensional picture of a highly variable four-dimensional dynamical system.

Similar problems plague the interpretation of the Galileo Probe results, at least as pertains to composition of the jovian atmosphere (a key goal for in situ science was the determination of the helium, nitrogen, and oxygen or water mixing ratios). The probe only detected water in small amounts at great depth. But it had entered a 5-micron hotspot (a gap in the normal planetary cloud cover where thermal radiation is much higher than for the bulk of the planet) [6]. How the special thermal environment of the hotspot affects the representativeness of its volatile components is still something of an open question. It depends crucially on the meteorological interpretation of these features [7]. (Interestingly, a similar question — of compositional and isotopic variations in space and time — has recently arisen in the Mars context where the partial condensation of carbon dioxide, the principal atmospheric component, has been seen to modify the winter polar composition [8].)

Venus has also been visited by a number of probes. The Pioneer Venus probes failed to make measurements all the way down to the surface and, as a result, were unable to shed light on the direction of the planet-atmosphere torque. The Vega Balloons were a bit of a departure, allowing direct (vertical and horizontal) wind measurements over a longer than normal probe lifetime, albeit at a single altitude [9].

The Huygens Probe is now poised to enter Titan’s atmosphere. Will it resolve questions of the origin and nature of the atmospheric super-rotation? Will it settle the issue of the existence of hydrocarbon lakes? Or of methane precipitation? A strict reading of past experience makes it unlikely that this will be the case. Great science (and more new ideas) will undoubtedly result from this mission, but a comprehensive meteorological picture is not going to emerge from the probe measurements alone. Undoubtedly, global observations from Cassini will be of great importance in the interpretation of the probe results. [A paper at this meeting by R. Young further addresses the need for synergy between probe and orbiter measurements.]

This is an opportunity to re-evaluate how to study (and observe) planetary atmospheres. Now that an initial reconnaissance of the atmospheres in the solar system has taken place, and broad experience with entry probes has been obtained, it should be possible to design a mission architecture that will address and find answers to the outstanding scientific and operational questions associated with these bodies.
2. PLANETARY METEOROLOGY MISSION CONCEPT

For all their diversity, we believe that all of the planetary atmospheres obey the same fluid dynamic laws (as embodied in the Navier-Stokes equations, for example). Meteorology, then, is essentially a problem of determining and providing the correct global (planetary radius, rotation rate, gravity, composition, mass of atmosphere), boundary (topography, insolation) and initial conditions (i.e., the current state of the atmosphere) for the solution of these equations. However, we believe that the governing laws are chaotic (i.e., extremely sensitive to initial conditions) and, therefore, that in order to maintain an accurate description of an atmospheric state it is necessary to continually update our knowledge based on new observations. For this reason, the concept of sending a single probe into an atmosphere to make measurements over a very short period of time will always be flawed. It is impossible to know whether the measurements made at that time are, in fact, representative of any other time and place on the planet. Rather, we want sustained time coverage of the entire planet. This provides a meteorological context in which to interpret other detailed measurements that can be made in situ (and which are then quite useful in improving our knowledge of the atmospheric state). In addition, since the different meteorological variables are (frequently) related by balance relations which involve their gradients, we want a synoptic view of the planet (in which quantities at different locations are determined at the same time).

Our requirements are a bit daunting: four-dimensional coverage compared to the one-dimensional observations from traditional probe missions. However, there is a huge payoff, scientific and practical, from the more comprehensive approach. The goal of exploration, after all, is to discover things that are previously unknown. This requires a superabundance of observations, so that new parameters can be fit. But if we are able to gather all of this data — and process it in time — we will have a very accurate analysis of the weather system under investigation and many operational uses of this knowledge will be available.

Based on the equations of motion, we can estimate the time and length scales over which significant changes in the general circulation of a planet take place. Temporal changes are usually controlled by diabatic heating (from the absorption of solar radiation), and so the crucial time interval is the radiative time constant, the e-folding time for atmospheric temperature changes by radiation: \( \tau = \frac{\rho H c_p}{4 \sigma T^3} \). Values for all of the planets are given in Table 1 (as derived from the Planetary Data System’s Planetary Atmospheres Node). Because of the density factor in this relation, the values of \( \tau \) can vary widely through the altitude range of interest. We have chosen “tropopause” values as being typical of the region that is easily observed by remote sensing and by entry probes. The important length scale is the Rossby radius of deformation [10] that gives the scale of baroclinic waves in the atmosphere (i.e., of the storm systems that constitute the primary deviation from zonal symmetry). \( L = \left( \frac{g H}{\tau} \right)^{1/2} \). There is relatively little variation of this length scale from planet to planet. (See Table 1. For Venus and Titan, we choose values of \( \tau \) that correspond to the super-rotation rather than the solid body rotation rate.) The (polar) orbital period of a remote sensing instrument should be short enough so that it can see the whole planet with spatial resolution \( L \) within timescale \( \tau \), or \( P \sim \tau L/2R \).

Table 1. Required periods for atmospheric sounders

<table>
<thead>
<tr>
<th>Planet</th>
<th>( R ), km</th>
<th>( \tau ), days</th>
<th>( L ), km</th>
<th>( P ), days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earth</td>
<td>6378</td>
<td>48</td>
<td>1862</td>
<td>7.0</td>
</tr>
<tr>
<td>Mars</td>
<td>3395</td>
<td>2</td>
<td>1391</td>
<td>0.4</td>
</tr>
<tr>
<td>Venus</td>
<td>6070</td>
<td>3</td>
<td>5631</td>
<td>1.4</td>
</tr>
<tr>
<td>Titan</td>
<td>2440</td>
<td>19</td>
<td>1145</td>
<td>4.5</td>
</tr>
<tr>
<td>Jupiter</td>
<td>71300</td>
<td>1650</td>
<td>1950</td>
<td>22.5</td>
</tr>
<tr>
<td>Saturn</td>
<td>60100</td>
<td>7590</td>
<td>1780</td>
<td>112.0</td>
</tr>
<tr>
<td>Uranus</td>
<td>24500</td>
<td>48000</td>
<td>1380</td>
<td>1350.0</td>
</tr>
<tr>
<td>Neptune</td>
<td>25100</td>
<td>44000</td>
<td>1930</td>
<td>1700.0</td>
</tr>
</tbody>
</table>

In fact, accumulating a lot of information from orbit is not difficult. TES has returned on the order of 200,000,000 infrared spectra over 3 Mars years. And only about 10% of the bandwidth is devoted to the atmospheric component of the observation. If the data can be made usable, there is a long list of applications that will benefit. Already, aerobraking has become an important part of the Mars payload delivery system. With increased confidence confidence in the forecasts of upper atmospheric densities, we can expect aerocapture missions and/or glider systems for accurate payload deliveries. The challenges of exploring a planet with as much surface area as the Earth will eventually lead to airplane or balloon missions that will require meteorological analyses and forecasts. And, of course, human exploration will depend on dust storm warning systems and the like. In order that this information be usable at Mars, it is highly desirable that not only data gathering, but routine data processing take place there. Already, the burden of aerobraking requires that a large Earth-based team make decisions about upper atmospheric densities. This would be unnecessary if the spacecraft could collect and evaluate the data autonomously. The case becomes stronger as we move to more distant planets where the two-way communication times are much too long to allow for continuous mission management from Earth. On the other hand, increased autonomy will allow for improved data products. Adaptive observations can be made to improve analyses in critical areas or where errors are large. Certainly, one can envision a small radiosonde network in support of human exploration, for example.

Clearly, then there is still a role for probes in providing the ground truth (from in situ measurements) that calibrates the more prolific orbital observations. Small mass (to increase the number of probes that can be utilized for a variety of purposes) and long lifetimes (to increase the data return) are clearly desirable. Because winds are more diagnostic than temperatures in determining the
meteorology, a possible approach is the deployment of small floaters that become entrained in the wind and can be tracked from the primary spacecraft. (There is clearly a challenge in this type of design as most probes would drift in the zonal direction, while the orbiters would presumably be highly inclined in order to obtain global coverage. So the tracking problem is operationally difficult, at least at the terrestrial planets. On the other hand, it may be possible to recover the probes after they become occulted by the planet, and to download stored data. Alternatively, perhaps a system of weather balloons would act as semi-permanent wind probes. If these include lidars for measuring atmospheric vertical structure, a very rich dataset could be obtained.)

The most important point to make is that the quantity of data is very important. Four-dimensional systems have many degrees of freedom and it is impossible to constrain them without adequate data. On the other hand, data quality is less important. Various measurements are physically linked by the equations of motion. So given enough observational data, one can add many constraints to that dataset. In practice, for example, most operational weather analyses and forecasts use initialized data (i.e., data that has been filtered and modified to give better operational results). It doesn’t matter very much if this data is noisy, as long as the statistics of that noise are known. It is then possible to modify it into the best form for operational use. Very accurate measurements of a few quantities (which are plagued in any case with the representativeness issues discussed above) are not nearly as useful. The trend in planetary probes — towards greater sophistication and expense to make a few localized measurements — flies in the face of this practical reality. A move towards cheaper, simpler probes is clearly in the cards.

3. NEW SOFTWARE TOOLS FOR PLANETARY ATMOSPHERES OBSERVATIONS

Based on a lot of technology transfer from the terrestrial meteorology community, the software tools for a new generation of planetary atmospheres observation missions are available. In fact, thanks to the limitation on resources that can be devoted to planetary exploration (as opposed to the high economic value of terrestrial numerical weather prediction), many of the planetary atmospheres tools are simpler to implement (and thus are better suited for the real time onboard processing possibilities that may be vital to remote operations). The tools are essentially 1) a versatile predictive modeling capability; 2) robust data assimilation techniques; and 3) adequate filtering methodologies.

The EPIC model [11] is a hybrid coordinate model that has been used to model a large number of planetary atmospheres [12], including the 5-micron hotspots that are so crucial to the interpretation of Galileo data [7]. The choice of vertical coordinate is suitable for modeling the deep atmospheres of the giant planets (for which the model was designed), but can also be used to simulate an atmosphere with a solid lower boundary. EPIC incorporates all of the physics that is needed for a weather forecasting model. And the code has been made available for all interested users via the Planetary Data System. This is just one example of a general circulation model that is well-tested and readily available to be applied to the analysis of planetary atmospheric data.

Data assimilation combines the forecast model and observations to produce an analysis of the state of the atmosphere that is consistent with both the governing equations of the model and the data [13]. The analysis can be thought of as a weighted average of the model predicted state and the measured state [14]. Determining the optimal average requires that both the model forecast errors and the observational errors (which include not only instrumental noise but representativeness errors) be known. The latter can clearly be inferred from the statistical properties of the data themselves. Model errors are another matter. While there are a number of approaches to determining forecast errors for well-tested terrestrial numerical weather prediction models (e.g., by making ensemble forecasts or from knowledge of the intrinsic co-variability in the atmosphere) [15, 16], planetary atmospheres and general circulation models are not so well known. Therefore it is preferable to deal with the the statistics of the residuals of predicted observables (like infrared radiances), rather than the presumed model errors in wind forecasts, say. We have shown that use of the resulting innovation covariance matrix [17, 18] leads to a viable and efficient formulation of the data assimilation problem.

The resulting observation space assimilation procedure greatly reduces the calculations required to make meteorological analyses, while putting the emphasis on producing a quality controlled version of the original data. The algorithm cycles (i.e., each new observation leads to an updated estimate of the atmospheric state which is used for the analysis of the following observation) and so is suitable for real time operation.

Given the estimate of the true value of the observed quantities, it is still necessary to find the atmospheric state (temperatures, winds, etc.) consistent with these measurements. This is best done by the four-dimensional variational (4DVAR) technique if an adjoint form of the predictive model is available. (The adjoint determines the sensitivity of the model’s final state to the initial state and therefore gives guidance as to how to adjust the initial state in order to arrive at a desired final state.) For example, to test the viability of 4DVAR in the martian context, we have formulated a Martian general circulation model with an adjoint version [19]. The dynamical core of this model is based on the baroclinic spectral formulation that has long been used in terrestrial numerical weather prediction [20-22]. The model includes realistic topography and a diurnal cycle, but is treated as imperfect when doing assimilations. In practice, this means that the diabatic forcing (a complicated function of the highly variable atmospheric dustiness on Mars) is assimilated rather than predicted by the model.
Will the answer to the assimilation problem derived by this technique be unique? Not necessarily, as a primitive equation model has a large number of free modes which are useful for matching observed conditions, but which do not correspond to real motions in the atmosphere. These rapidly varying modes, usually called gravity waves in the atmospheric modeling community, must be filtered out of the solutions in order to obtain reliable forecasts. This initialization process has been the subject of much study in numerical weather prediction [23]. It is frequently achieved in terrestrial modeling by imposing a balance condition between the model wind and mass (temperature) fields. However, it is difficult to impose such a condition on the Martian atmosphere as that would filter out atmospheric tides which are known to be important and which do not satisfy a balance relation. Instead, we have recently developed a digital filtering technique [16] that minimizes the ill effects of fast gravity waves.

Digital filtering of gravity waves sidesteps the issues of evaluating in advance what the balance relations for a given planetary atmosphere should be. The modes that are filtered do not represent physical reality in any case, and so manipulating their amplitudes has a small effect on the ultimate quality of the results. What is most important in all of these procedures is the minimum foreknowledge of the given system that is required to produce high level atmospheric dynamics products. Since our goal is exploration, it is important that we do not assume that we have good preknowledge of the conditions in a planetary atmosphere.

To implement the onboard processing of data for a planetary mission, it is important that observational teams produce forward models for their instruments. (In any case, this would be a desirable part of the instrument design process.) Such models need to be constructed properly so that linearization and adjointing can be implemented and lead to robust codes in preparation for the 4DVAR analyses. Only an instrument whose signal is, in principle, invertible is going to be a reasonable flight prospect, so it is to be expected that such robust codes can be produced. Once available, the codes and the onboard processing relieve the instrument teams of the responsibility for producing standard data products, so they can devote their efforts to higher level scientific concerns.

The upcoming Mars Climate Sounder (MCS) will provide an opportunity to test this software and this modeling approach (though not yet the onboard processing aspects of the proposal). It is likely that at least some of the other atmosphere-observing instruments orbiting the planet at this time will overlap with MCS and provide an excellent opportunity for validating its products.

4. NEW HARDWARE TOOLS FOR PLANETARY ATMOSPHERES OBSERVATIONS

The goal of this paper is to challenge the experienced probe community to come forward with a new set of hardware components that will match the available software techniques in providing all elements needed for future planetary atmospheres missions.

Clearly, a new generation of sounders will be needed. Most planetary spacecraft to date have depended on infrared spectrometers to determine vertical structure. But the wave of the future seems to be moving towards microwave sensors. The ability to measure individual lines in the microwave region leads to greatly enhanced sounder capabilities. Among these are the ability to penetrate clouds and dust, the ability to measure isotopes and tracers, and probably the ability to determine horizontal winds from Doppler shifts.

The new entry probes will provide crucial observations, but many fewer measurements than the sounders. So it is appropriate that they be less expensive and significantly lighter. It might be possible for very light probes to float in the atmosphere, significantly increasing the data return and the decreasing the cost/benefit ratio for these mission components. The exciting new role of the entry probe will be to make targeted observations. Thus, it is desirable that the orbiter be capable of carrying a large number of entry probes with an accurate release mechanism.

5. CONCLUSIONS

The World Meteorological Organization has recently begun the THORPEX program to improve the accuracy of terrestrial weather forecasts. Among the objectives of this program are [24]

- the incorporation of model uncertainty into data-assimilation systems;
- developing adaptive data-assimilation and target-observing strategies;
- improving the assimilation of observations of physical processes and atmospheric composition; and
- the introduction of interactive procedures that make the forecast system more responsive to user needs.

The planetary exploration program is now in a position to take advantage of the terrestrial weather forecasting experience in the design of planetary atmospheres missions with these valuable capabilities. The tools that are becoming of greater importance to weather forecasters can be used as readily by spacecraft at other planets. They will lead to greater autonomy and adaptability of missions. This will in turn lead to better quality high-level
scientific products and real-time availability of data for operational purposes.

The Mars program will be the testbed for these technologies. We have already learned that most of the required information to determine the state of the atmosphere can be obtained from low orbit, even with instruments that are not specifically designed for atmospheric observations. However, to make the most of the data, assimilation techniques must be used to assure that the retrievals of structure and winds are physically consistent. This same assimilation allows supplementary data to be used for calibration and validation. They also allow targeted in situ observations to provide the crucial ground truth for the remote sensing instruments.

The design of a new generation of small entry probes that can be released by the orbiters to provide ground truth measurements is the crucial next step in the development of this exploration program.

REFERENCES


