JPL Publication 01-008, Rev. A

Biosignatures and Planetary Properties to be Investigated by the TPF Mission

David J. Des Marais
Ames Research Center
Moffett Field, CA

Martin Harwit
Cornell University
Ithaca, NY

Kenneth Jucks
Center for Astrophysics,
Smithsonian Institution
Cambridge, MA

James F. Kasting
The Pennsylvania State University
State College, PA

Jonathan I. Lunine
University of Arizona
Tucson, AZ

Douglas Lin
University of California
Santa Cruz, CA

Sara Seager
Institute for Advanced Study
Princeton, NJ

Jean Schneider
Observatoire de Paris
Meudon, France

Wesley Traub
Center for Astrophysics
Smithsonian Institution
Cambridge, MA

Neville Woolf
University of Arizona
Tucson, AZ

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

June 2002

ACKNOWLEDGEMENTS

This work was performed by D. Des Marais, M. Harwit, K. Jucks, J. Kasting, J. Lunine, D. Lin, S. Seager J. Schneider, W. Traub and N. Woolf as members of the Biomarker Subgroup of the Terrestrial Planet Finder Science Working Group (TPF-SWG). TPF-SWG is an activity sponsored by the Terrestrial Planet Finder Project Office at Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Support is also acknowledged from the following: awards by Jet Propulsion Laboratory to Ball Aerospace and Technologies Corp. (JFK, SS and WT), Lockheed-Martin (DJD, NW and JIL), TRW (DL) and Boeing/SVS (MH and JS); NASA via contract JPL 1201749 (WT), the NASA Astrobiology Institute and the NASA Exobiology Program (DJD and JFK); and the W. M. Keck Foundation (SS).

Abstract

A major goal of Terrestrial Planet Finder (TPF) mission is to provide data to the biologists and atmospheric chemists who will be best able to evaluate the observations for evidence of life. This white paper reviews the benefits and challenges associated with remote spectroscopic observations of planets; it recommends wavelength ranges and spectral features; and it provides algorithms for detection of these features. For the mid-infrared region, the minimum and preferred wavelength ranges are 8.5 to 20 μm and 7 to 25 μm, respectively. For the visible to near infrared, the minimum and preferred ranges are 0.7 to 1.0 μm, and 0.5 to ≈1.1 μm, respectively. Detection of O_2 or its photolytic product O_3 merits highest priority because O_2 is our most reliable biomarker gas. O_3 is easier to detect at low O_2 concentrations. O_3 might be detected in the UV range (at 0.34 to 0.31 μm), but potential interferences must first be evaluated. Even though H_2O is not a bio-indicator, its presence in liquid form on a planet's surface is considered essential to life. CO_2 is required for photosynthesis and for other important metabolic pathways. CO_2 indicates an atmosphere and oxidation state typical of a terrestrial planet. Abundant CH_4 can indicate a biological source, although nonbiological sources might be detectable, depending upon the degree of oxidation of a planet's crust and upper mantle. A planet's size is very important for assessing its habitability, as illustrated by comparing Earth and Mars. Planet size can be estimated in the mid-infrared range, but not in the visible to near infrared range. Both mid-infrared and visible to near-infrared ranges offer valuable information regarding biomarkers and planetary properties; therefore both merit serious scientific consideration for TPF. The best overall strategy for the Origins Program includes a diversity of approaches, therefore both wavelength ranges should ultimately be examined prior to launching the "Life-Finder" mission. The program must embrace the likelihood that the range of characteristics of extrasolar rocky planets far exceeds our experiences with our own four terrestrial planets and the Moon.

Table of Contents

Figures

Tables

I. Basic Science Goals of TPF and the Broad Diversity of Terrestrial Planets

Beyond the Earth yet within our Solar System, our search for life and evidence about the origin of life will likely be confined to Mars, Europa, and Titan, along with small bodies such as comets, asteroids, and meteorite fragments derived therefrom. These objects present a wonderful opportunity for detailed studies that is not possible when we make explorations outside the Solar System. But the compensating advantages of studies beyond the Solar System are the greater diversities of both of environment and of stages of development that are available for investigation. Secondly, the search for extrasolar planets with biospheres is a search for the broadest biological diversity possible, including the possibility of life having origins totally independent of our own.

We do not have available, in the present orbit of Mars, a more-massive planet that might have sustained a greenhouse atmosphere and carbon-silicate cycling, in contrast to Mars itself, but perhaps there is a system out there that has such a planet. We do not have the choice of seeing Earth without Jupiter being present, but probably there is a system like that out there. We do not have the choice in our solar system of seeing planets having mainly ice-free oceans, and in different stages of planetary development than Earth, but those may well be out there, too. Trying to make statistical arguments about what we do or do not find with three planets is quite limiting. Potentially there are also limits that are imposed by the possibility that Earth and Martian life have been linked by rocks transiting rapidly from one planet to the other. The benefits of an astrobiology program that includes both solar system observations and a search for and study of terrestrial planets in other systems is that it frees us from the limitations imposed by our own Solar System, and it addresses the profound question of life's cosmic ubiquity. Without both, too many extraordinarily interesting questions are left unexplored.

Therefore, in determining the appropriate wavelength range for TPF's spectroscopic capability, one must bear in mind that the range of characteristics of rocky planets is likely to exceed our experiences with the four terrestrial planets and the Moon. While the nearly (but not quite) airless Moon and Mercury arguably represent the lifeless end-member case of terrestrial planets, there are always surprises. For example, Mercury appears, based on radar data, to support small polar caps of water ice, and the origin of the water appears to be exogenic impact of icy material followed by molecular migration to the poles. Were such a body to be in a planetary system in which the orbital plane happens to be face-on to the Earth, could that water ice signature be detectable in the near-infrared range, and, if so, what would one conclude about the habitability of such an object? Habitability might be ruled out if the semi-major axis were too small (indeed, the planet might be missed altogether), but no laws of physics rule out a "Mercury" placed at the orbit of, say, Venus (0.7 AU). What would one conclude then?

Likewise, the point has been made in a number of papers and abstracts that, absent Jupiter in our planetary system, a rather wet terrestrial planet, possibly with large mass, might have formed in place of Mars, or beyond Mars in the orbital region of our present asteroid belt. We have absolutely no experience with such a body, and models of the stability of a dense greenhouse atmosphere and surface-atmosphere evolution are our only guides to such a case. What signatures would we look for in the case of such an object? How would we determine whether signs of habitability suggest equable conditions over geologic time versus, say limited periods in the distant or even recent past, except by assumption?

In some planetary systems, planets may not be coplanar, and in such systems terrestrial planets might be moved in and out of the habitable zone over long periods of time. Likewise, Earth-like planets evolving around stars of very different spectral type than the Sun, hence different spectral energy distribution, might evolve in unexpected ways (and, for K and M dwarfs, over much longer time spans than could our Earth.

The ill-fated namesake of Shakespeare's play Hamlet admonished his friend that: "There are more things in Heaven and Earth, Horatio, than are dreamt of in your philosophy" (Act 1, scene 5). Today, we might instead warn ourselves of the certainty that there are more kinds of Earths in the heavens than are dreamt of in our philosophy. TPF must be designed so that it can characterize diverse types of terrestrial planets with a useful outcome. This must be borne in mind while we cogitate over much narrower questions such as what the clouds were like on the Cretaceous vs. the modern Earth, etc. This does not mean that the discussions that led to this report were not useful or important—quite the contrary. They set the stage, with the valuable set of tools provided by the interdisciplinary research community, for thinking beyond the detection and characterization of a twin of our own Earth to the astronomical exploration of what must surely be a menagerie of wonderfully odd worlds.

Accordingly, the principal goal of TPF is to provide data to the biologists and atmospheric chemists who will be best able to evaluate the observations of a potentially broad diversity of objects in terms of evidence of life and environmental conditions (for example*, [Beichman et al., 1999] [Caroff and Des Marais, 2000]*). We must answer the following questions. What makes a planet habitable and how can that be studied remotely? What are the diverse effects that biota might exert on the spectra of planetary atmospheres? What false positives might we expect? What are the evolutionary histories of atmospheres likely to be? And, especially, what are robust indicators of life? The search strategy should include the following goals that collectively serve as milestones for the TPF Program.

TPF must directly detect and survey nearby stars for planetary systems that include terrestrial-sized planets in their habitable zones ("Earth-like" planets *[Beichman, et al., 1999]*). Through spectroscopy, TPF must determine whether these planets have atmospheres and establish whether they are habitable. We define a habitable planet in the "classical" sense, meaning a planet having an atmosphere and with liquid water on its surface. The habitable zone therefore is that zone within which starlight is sufficiently intense to maintain liquid water at the surface, without initiating runaway greenhouse conditions that dissociate water and sustain the loss of hydrogen to space *[Kasting et al., 1993]*. There are interesting potential examples where liquid water might exist only deep below the surface, such as the jovian moon, Europa (e.g. *[Reynolds et al., 1987]*), or on Mars. However, biospheres for which liquid water is present only in the subsurface might not be detectable by TPF. Thus a planet having liquid water at its surface meets our operational definition of habitability, which is that habitable conditions must be detectable. Studies of planetary systems will also reveal how the abundances of life-permitting volatile species such as water on an Earth-like planet are related to the characteristics of the planetary system as a whole [*Lunine*, 2001].

TPF must target the most prospective planets for more detailed spectroscopy, and it must determine whether biomarkers are present. A biomarker is a feature whose presence or abundance requires a biological origin. Biomarkers are created either during the acquisition of

both the energy and/or the chemical ingredients that are necessary for biosynthesis (e.g., leading to the accumulation of atmospheric oxygen or methane). Biomarkers can also be products of the biosynthesis of information-rich molecules and structures (e.g., complex organic molecules and cells). Life can be indicated by chemical disequilibria that cannot be explained solely by nonbiological processes. For example, a geologically active planet that exhales reduced volcanic gases can maintain detectable levels of atmospheric oxygen only in the presence of oxygen-producing photosynthetic organisms. Alternatively, an inhabited planet having a moderately reduced interior might harbor greater concentrations of atmospheric methane than an uninhabited planet, due to the biosynthesis of methane from carbon dioxide and hydrogen at cooler ($<120^{o}$ C) temperatures.

Although the "cross hairs" of the TPF search strategy should be trained upon "Earth-like" planets, TPF should also document the physical properties and composition of a broader diversity of planets. This capability is essential for the proper interpretation of potential biomarker compounds. For example, the presence of molecular oxygen in the atmospheres of Venus and Mars can indeed be attributed to nonbiological processes, but only through a proper assessment of the conditions and processes involved (for example, *[Kasting and Brown, 1998]*). On the other hand, a planet might differ substantially from Earth yet still be habitable. Accordingly, in order to assess thoroughly the cosmic distribution of habitable planets, we must understand both the processes of formation of planetary systems as well as the controls upon the persistence of habitable zones. This approach calls ultimately for observations of multiple planets within systems, including those that are uninhabitable.

The TPF program should create a path of continuous discovery of planetary systems and environments that leads ultimately towards locating and characterizing habitable and inhabited planets. For example, although we might find only uninhabited planets at the outset, these will enhance our understanding and guide our search for the more promising candidates. An ongoing parade of preliminary discoveries will build and sustain public interest, education and support. A strategy that seeks not only Earth-like planets but also the context required for habitable planetary systems *[Kasting, 1988]* to develop will probably lead us most directly to that first bit of evidence that we are not alone in the Universe.

II. Interpretation of Visible and IR Spectra of Terrestrial Planets

Introduction

The simple observation that a planet exists at some distance from a star will determine whether the planet is in the habitable zone of the star, but it will only give us a very rough estimate of the temperature. There are two temperatures of interest, the effective temperature (that of a blackbody having the same surface area and the same total radiated thermal power) and the surface temperature (at the interface between any atmosphere and the solid surface). In general, if there is a greenhouse effect present (e.g., from CO_2, H_2O, CH_4, or aerosols), then the surface temperature will be warmer than the effective temperature. The effective temperature is determined by the stellar brightness, the distance, R, to the star, the albedo, A, and whether the day-night temperature difference (D-N) is small (as for a rapid rotating planet or a planet with a massive atmosphere) or large (as for a slow rotator or a thin atmosphere). Our own Solar System offers the following examples:

Venus: R = 0.72 AU, A = 0.80 $\pm$ 0.02 *[Tomasko, 1980]* D-N is small;

Earth: R = 1.00 AU, A = 0.297 $\pm$ 0.005 *[Goode et al., 2001]*, D-N is small; and

Mars: R = 1.52 AU, A = 0.214 $\pm$ 0.063 *[Kieffer et al., 1977]*, D-N is large.

If we assume that we have a planet within this range of values, but with otherwise unknown values of R, A, and D-N, then the predicted effective temperatures will span a range of 202 K, which is almost uselessly large. If we assume that we know either R or A or D-N, then the predicted range drops to 123 or 159 or 169 respectively, showing that the most important parameter is R, followed by A, and finally D-N. If however we know R from physical observations with TPF, then the range of possible effective temperatures drops to 123 K, which is still rather large. Finally, assuming that we know at least two of these parameters, the range falls to 75 K if we know D-N, and 70 K if we know A, and of course to 0 degrees if we know all three parameters.

By definition, we can determine A if we measure both the visible and infrared flux, which argues in favor of a dual-capability TPF. We can determine D-N by several methods, as follows: measuring the infrared flux at two or more points in the orbit (looking at the day and night sides, respectively); measuring the visible flux at two or more points in a diurnal cycle to indicate rotation rate.

Regarding our search for life, learning a planet's surface temperature holds much greater value than its effective temperature. For example, both Venus and Earth have nearly identical effective temperatures (220 K and 255 K, respectively), but vastly different surface temperatures (730 K and $\approx$290 K respectively) owing to the different greenhouse gas column abundances. Visible and/or infrared spectra can help interpret these cases, but neither are able to penetrate clouds, therefore surface conditions may well be difficult to estimate. In the following discussion, we outline the facts that we might expect to learn from spectra.

Given the spectrum of an extrasolar planet, it is possible to derive the physical conditions and aspects of the composition at the layer where $\tau \approx 1$. Infrared observations of the continuum give a color temperature. By equating this with the physical temperature and using Planck's law, we can derive the planet size from the infrared emission, but infrared spectral band profiles are strongly affected by the thermal structure of the atmosphere. These spectral bands can be used to tell the presence or near absence of atmospheric constituents and to study the thermal structure, but they are poor for determining quantitative abundances.

The visible/near IR continuum does not give direct indication of the planet size because of the possible albedo range. Evidence from band studies is required to infer planet size, albedo and temperature. The intensities of visible/near IR absorption features are not affected by thermal structure; therefore they are good for abundance determinations, but their use to infer the planet's temperature is less direct. The combination of infrared and visible observations is very valuable. Neither region will give all of the information, and either region will require modeling to interpret. We have not yet explored the difficulties associated with modeling.

There is however a concern that Earth is a peculiarly easy planet to interpret from external observations. Slightly larger or smaller planets having slightly different insolation may well have more cloud cover and be much harder to interpret. Ultimately, additional observations with giant space interferometers at millimeter wavelengths will probably be needed to determine surface temperatures of some planets. However, if near-Earth conditions occur on some planets, we *can* find these and interpret them with much more limited observations. Testing of model spectra is needed before we can be sure that both visible and mid-IR observations could do this.

General Discussion

Observations of a very distant planet can only distinguish those planets whose habitability is apparent from observations of the reflected or emitted radiation. Planets with habitable surfaces that are hidden by deep, totally opaque material cannot be recognized. Planets with a small fractional area that is habitable cannot be detected. We are limited to exploring habitability for only those planets that ar*e easily demonstrably habitable*. Since habitability is assumed to be rare, we allow ourselves to make the error of assuming a planet uninhabitable when it is in fact habitable or inhabited. We are more concerned about making the other type of error—that is, to state that a planet is habitable or inhabited when it is not.

Planets with habitable surfaces hidden by clouds of small particles can only have their surfaces studied at wavelengths long enough to penetrate the clouds. For extrasolar planets, this cannot be achieved with current technology. Interferometer reflectors that are km in diameter must be deployed in space to measure planet surface temperatures. This will indeed be possible someday.

If the surface temperature can be found, then wavelengths that do not penetrate to the surface can determine surface characteristics from features observed at some higher level in the atmosphere. If the observable upper levels of the atmosphere appear to be moderately close to saturation with water vapor, it is possible to determine approximate characteristics at ground level by assuming that the atmosphere is in convective equilibrium with a wet adiabat. On the other hand, if the upper atmosphere has a photochemical smog, one cannot predict the characteristics of the layer, and extrapolation to layers below is impossible. Both Venus and Titan are totally enshrouded by photochemical clouds.

It is possible to have an atmospheric layer of low humidity above a layer of high humidity if there is a temperature inversion (for example, the stratosphere above the tropopause). So while it seems that it is possible to infer that wet upper atmospheres imply a wet surface, it is not possible to imply that dry upper atmospheres imply a dry surface.

Determining the characteristics of a terrestrial planet around another star will generally be quite difficult, although in the lucky case that a planet is very similar to Earth it will not be hard to interpret. From the example of the Sun's terrestrial planets, we find that there is an abrupt onset of atmospheric gas at a particular size or mass (the size of Mars, for example). For the two larger planets we have one example of a planet roughly half cloud covered (Earth) and another example of a planet completely cloud covered (Venus). For small and intermediate cloud coverage, the surface characteristics can be found fairly well. For extensive clouds, we can only determine the characteristics of the cloud layer and above. It is not yet clear whether cloud cover

that is sustained by convective processes typically produces an atmosphere having extensive regions that are cloud-free.

We will consider three cases for trying to determine planet characteristics:

1) IR spectrum alone
2) Visible spectrum alone
3) Both spectra available

We shall start by considering information available from the continuum, and proceed from there to spectral bands.

Continuum observations

We begin by trying to determine the planet's orbit, so as to estimate the insolation, or energy input per unit area. Both visible and IR measurements can give the projected orbit. From this it should be possible to infer the true orbit. Even without that, it should be possible to determine the maximum elongation, and with the assumption of a circular orbit this elongation is a direct measure of the planet-star separation. With that information, we know the insolation, and can determine how the thermal environment of the planet compares with corresponding Solar System planets.

From observations in the mid-infrared range, we can expect to measure a color temperature from spectral shape. That temperature may be a temperature of the surface, of low clouds or of high clouds. By equating color temperature for selected spectral regions with physical temperature and using the observed flux, we can use Planck's Law to obtain the surface area of the planet. We can use the observed solar system relationship between mass, radius and thermal environment, to infer the likely planet mass, and to consider whether this is or is not likely to be a planet with a history of active tectonics. This is an important aspect of our understanding of habitability.

Although we have considered that the presence of rings or satellites might distort our answer, we expect IR continuum observations to determine the planet radius to ≈10%. With that size and temperature, we can calculate the total surface emission, and compare it with the insolation. That comparison will tell us the Bond albedo (Equations 1 and 2).

Let r_{pl} = planet radius (unknown)
A = albedo (unknown)
D = distance to planetary system
L^* = star integral flux at Earth
θ = angle of planet max elongation
T = observed color temperature
$F(\lambda)$ = observed thermal planet flux at wavelength λ
$B(\lambda , T)$ = Planck function at wavelength λ for temperature T
Then we have

$$(r_{pl}/D)^2 = F(\lambda) / \pi\, B(\lambda,T) \qquad (1$$

and

$$1\text{-}A = ((D/r_{pl})^2\ \theta^2\ 4\int F(\lambda)\ d\lambda)/L^* \qquad (2$$

From albedo we can tell whether the cloud cover of our planet resembles that of the Moon, Mars, Earth or Venus. However, in the past, the Earth has been through a cold phase in which it had a high albedo (due to ice) and a low surface temperature. It is not clear whether we could distinguish such a snowball Earth from a Venus-like planet. If the albedo is high, we will know that we are either determining cloud-top characteristics or we are seeing an ice-bound planet—it is very hard to decide which! If the albedo is low, it will seem likely that we are seeing a surface and are determining surface characteristics. If the albedo is intermediate, we can expect that we are underestimating the surface temperature, but not badly. These results need verification of their precision from, for example, observations of Mars and Venus and models of Earth.

The surface temperature can be determined from infrared observations only under a very limited range of conditions. For present Earth, one could obtain it moderately well, in principle, by making observations in the 8- to 12-μm "window" region. However, there is some absorption by water in the window region even at present. The minimum optical depth is ≈0.1 at 8 μm (*[Kasting et al., 1984]*); see Figure 6. The absorption in this region is often referred to as "e-type" absorption, meaning that it varies with the square of the water vapor pressure, P_v. Meanwhile, P_v varies according to the Clausius-Clapeyron equation: a 20 K increase in T increases P_v by a factor of 3.4- and ≈8-μm optical depth by a factor of 11. Hence, the window region becomes effectively opaque, and one would measure the emission from a region that is located between 2 to 3 water scale heights higher in the atmosphere. This would spuriously depress the measured "temperature." For planets warmer than this, the surface temperature cannot be determined by infrared observations. Of course, such a warm condition could precipitate a catastrophic greenhouse conversion from an Earth-like to a Venus-like condition.

Alternatively, if we consider a more massive planet than Mars but receiving Mars-like insolation, the surface temperature will only be Earth-like if the planet's greenhouse effect is much larger than that of Earth, in which case, again, the surface temperature will likely not be available to infrared observations. Or, if we have a small planet with a thin atmosphere but receiving a Venus-like insolation, then a moist planet will develop a black sphere temperature of ≈300 K. With greenhouse warming, the amount of water in the atmosphere will likely make the surface temperature indeterminate.

It is possible that an oxygen feature under either of these circumstances cannot be interpreted as a definitive sign of life because oxygen having a nonbiological origin can accumulate on a large ice-bound planet. The ice prevents it being absorbed by the surface rock. However, a large planet is likely to have volcanoes and volcanic exhalations with chemical compositions that are reducing and therefore would remove any nonbiological oxygen inventories. Therefore there may be a only very small range of planetary conditions that might produce a false positive answer for oxygen. Nonbiological oxygen might also become apparent on a hot planet undergoing a catastrophic greenhouse conversion *[Kasting, 1988]*. Water vapor is transported to the top of the atmosphere where it is easily dissociated, hydrogen is lost and oxygen retained until it reacts with surface rock. This, too, might give a false positive answer for the detection of

life. Both of these conditions are distinguishable by the planet being outside the habitable zone, and so we can probably discriminate between real and false positives.

Overall the first and best-known aspect of a planet from infrared observations is its size. From the size, insolation and integrated emission we can determine both the albedo and a temperature associated with the emitting layer. It may be possible from these three parameters to understand whether a region close to ground or a layer high in the atmosphere is being observed. If it is an upper layer, interpretation of surface characteristics is sometimes possible, but there are clearly cases where it is not.

For the visible, the process is more difficult, and it is necessary to use observations of bands (discussed below).

Observations of bands

Spectral band observations at low resolution are most informative regarding the abundances of materials. The strength of bands varies with both the amount of material and the atmospheric pressure. For likely planets, the breadth of individual lines in the band will vary linearly with the pressure, but the absorption from that line will vary with the square root of the abundance of material. If a curve of growth (which is the relationship between the abundance of a species and the intensity of one of its bands) is created, it will, over most of its range, satisfy a relationship that the band strength is a function of (Pressure x sqrt[amount of line producing substance]). In such a relationship, pressure-induced self-broadening is about twice as effective as foreign gas broadening. These two quantities cannot be separated unless the atmospheric pressure is very low, and Doppler broadening dominates. This circumstance will only occur for weak spectral bands in the visible region. (The pressure broadening is by a collision frequency, whereas Doppler broadening is proportional to the line frequency, which will be greatest in the visible.) For this reason it will in general be hard to quantify the meaning of an observed band strength.

It is even harder to interpret the strength of a band in the mid-infrared spectral range. There the absorbing material also emits. An absorption band exhibits a steadily higher opacity towards its center. Thus the band profile actually indicates the vertical temperature structure above the continuum-emitting region. Some features, such as central inversion or absence of inversion may be informative about the vertical temperature structure in the atmosphere. Thus, for example, the central inversion in the 15-μm band of CO_2 on Earth is caused by the temperature increase from tropopause to stratosphere which, in turn, is caused by O_3 absorbing sunlight. Such information could, for example, actually confirm that O_3 was indeed present.

Absorption by water poses an interesting challenge for interpretation. Water absorption is observed in the near-infrared spectra of cool giant stars and brown dwarfs. The bands are the same bands as seen in Earth's spectrum, but they are somewhat broader and therefore they modify the apparent shape of the continuum between bands. The bands imply a certain (pressure)(square root of water abundance) product. The water is likely to be near saturation if water is indeed reasonably abundant, but water vapor pressure varies dramatically with temperature (for example, from freezing point to boiling point on Earth it varies by a factor ≈180). Thus a fairly precise measure of temperature is needed to take advantage of information about the strength of water bands.

The widths of water and oxygen features hardly vary with increase of temperature and therefore cannot be used to infer temperature. It is more likely that water band strengths can be used to determine a reasonable range of temperatures and pressures. The amount of water is established by vapor pressure, and so there is probably some lower limit to temperature implied by keeping the atmospheric pressure reasonable. These relationships need investigation.

In the mid-infrared range, the short wavelength end of the rotational continuum overlaps with the long wavelength end of the 15-μm CO_2 band. Therefore the observed information is the depth of drop at this long wavelength edge. It gives information about the temperature where the water opacity ≈1, as seen from above. If we assume a saturated atmosphere, then, the lower the atmospheric pressure, the less will be the difference in temperature between the short and long wavelength edges of this band. The relationship between this band feature and the estimate of temperature, made from observations of the continuum, must be investigated further. Such observations might provide estimates of pressure at levels in the atmosphere where the water band is observable.

Determining Parameters from Visible- and Near-Infrared Bands

In addition to the bands, it is necessary to start with an approximate determination of the planet size. The visual luminosity of the planet and its annual phase variation can be used to derive the quantity $A(r_{pl})^2$. We require g, the acceleration due to gravity at the planet surface.

Let ρ be the planet mean density

$$g = 4/3\ Gr_{pl}\pi\rho \qquad (3$$

We can assume ρ ≈5 as a first approximation for a terrestrial planet. So g only varies with sqrt(A). That is, it is uncertain by ± a factor ≈1.7.

Then there are two kinds of molecules, ones that vary in mixing ratio (H_20 and O_3), and ones that have constant mixing ratio. For the gases having constant mixing ratios and that produce bands, we can use oxygen and CO_2 to estimate the total pressure, under the assumption that the mixing ratio is unity (that is, we assume that we have a pure O_2 or CO_2 atmosphere). This will give a lower limit to the pressure at the surface of view. Because the gas pressure varies linearly with T, we can assume some mean T ≈260 K (a reasonable approximation for Earth, Venus or Mars) because we chose to study only planets in the potentially habitable zone.

For H_20, we must assume the scale height (for terrestrial temperatures, it is about 1/5 the atmospheric scale height because of condensation). Thus using the just-determined surface pressure and the H_20 band strength, we can determine the vapor pressure of water at the surface. This can then be translated into a better temperature by assuming that the atmosphere is 50% saturated. Then this temperature can be fed back into the pressure determination and the process iterated.

Again it would be advisable to use this process to test against observations of Earth, Mars and Venus as well as simulations of a "smaller" Venus and "larger" Mars to explore the quality of the results. Indeed, it is a necessary step in this work that both visible and infrared processes be validated.

Effects of Dust Rings, Binary Planets and Dust Trails

The derivation of the albedo **A** of a planet from its radius $\mathbf{R_{pl}}$ and reflected flux $\mathbf{F_r = A\pi R_{pl}^2 \phi(t)}$ (where $\mathbf{\phi(t)}$ is an orbital phase factor) assumes that the planet is spherical. That assumption is not true in the following two cases: circumplanetary rings and a planet with a companion. The latter case can involve a satellite that is smaller than the planet, as with Earth, Jupiter and Saturn. However, the diversity of planetary systems that has already been discovered leaves open the possibility that some extrasolar planet companions may be as large as the planet itself ("binary planets"). Although the lifetime of rings may be short (10^4 years), one cannot exclude that they are continuously replenished. In these two cases, the planetary albedo that is estimated from its reflected flux (or its radius derived from an estimated or "guess" albedo) will be incorrect. "Counter-measures" have been proposed that address these potential sources of error *[Schneider, 1999]*.

These issues are presented more thoroughly in the appendix (Section VI B.). Here we briefly summarize key results. Planet-trailing clouds were discussed by Beichman et al. *[Beichman et al., 1999]*. For the cloud like the one that is currently believed to trail the Earth, the flux is about 10% of that of the Earth. The effects of this cloud on the analysis are small, indeed much smaller than was estimated by Beichman et al. *[Beichman et al., 1999]*. However, in systems having much more dust, the dust would compromise infrared observations as well as visible imaging for low angular resolution after apodization.

Observations in the visible range can distinguish between the effects of rings versus effects of binary planets or moons. Rings will likely express a distinctive shape in their phase effect. Eclipses might possibly reveal moons and binary planets. Such information would be helpful in determining the parameters of a moon-planet system.

In the mid-infrared range, the effects of a terrestrial-sized moon would be relatively unimportant, but a larger moon would result in spectral features whose strength varied throughout the year, and whose temperature similarly would appear to vary in this cycle. A planet in an eccentric orbit would show the temperature variation, but not the washing out of spectral features. High precision positional measurements that are performed as part of the study of planets might help to discriminate between the various effects and to infer a planet's mass.

Conclusions

We have examined the potential use of both visible/near infrared spectra and mid-infrared spectra to interpret observations of terrestrial planets at both wavelengths. Estimates of planet size and albedo can definitely be determined from mid-infrared observations. These parameters are much more difficult to determine from visible observations, but an iterative scheme has been developed for exploring this. Methods that utilize either wavelength range should be tested against actual observations and models.

Surface temperature determination is only possible if there is a planet with a substantial fraction that is cloud free. The presence of O_2 and H_2O is determinable from both spectral regions, but care will be needed to exclude false positives where detection of O_2 and O_3 is not a sign of life.

III. Infrared and Visible Spectral Features

Model

Model Earth spectra are calculated with the SAO code originally developed to analyze balloon-borne, far-infrared, thermal emission spectra of the stratosphere, and later extended to include visible reflection spectra *[Traub and Stier, 1976]*; *[Traub and Jucks, 2001 (submitted)]*. The spectral line database includes the large AFGL compilation plus improvements from pre-release AFGL material and our own sources. In a few cases, laboratory data are available but spectroscopic analysis is not, so here we use an empirical psuedo-line band shape. The far wings of pressure-broadened lines can be non-Lorentzian at around 1000 times the line width and beyond, therefore, in some cases (H_2O, CO_2, N_2) we replace line-by-line calculation with measured continuum data in these regions. Dust and Rayleigh scattering are approximated by applying empirical wavelength power laws that exert an appreciable effect in the visible blue wavelength range. Model atmospheres from 0 to 100 km altitude are constructed from standard models that are discretised to appropriate atmospheric layers, and additional radiative transfer methods used to ensure that line cores and optically thick layers are accurately represented. Integration over the spherical Earth atmosphere is approximated to a few percent accuracy by a single-point calculation at a zenith angle of 30 degrees, so the effective air mass is 2 in the infrared (outgoing emission) and 4 in the visible (2 incoming sunlight plus 2 outgoing reflected light). Clouds are represented by inserting continuum absorbing/emitting layers at appropriate altitudes, and broken clouds are represented by a weighted sum of spectra using different cloud layers.

The sections that follow present results for each atmospheric feature or species, in both the mid-infrared and visible-near-infrared spectral ranges.

Clouds (Figs. 1 and 2)

The infrared model Earth spectrum in Figure 1 illustrates the combined effects of spectral absorption features and clouds. The clear-atmosphere spectrum (with no clouds) is shown by the uppermost curve, illustrating that, with no clouds, we can penetrate deepest into the troposphere and therefore observe the warmest (brightest) emitting altitudes. If the planet is covered with high clouds, at about the altitude of the tropopause, we get the lowest curve which is essentially a blackbody at the tropopause temperature, plus superposed emission features from the higher-altitude warmer layers in the stratosphere. The curves in between these two extreme cases show intermediate altitude cloud covers. The heavy (middle) curve represents the combined effect of a weighted average of clear and cloudy patches to simulate roughly the present Earth. This range of spectra shows that clouds and gas species may dominate the mid-infrared spectrum. All infrared spectra are triangle-smoothed to a resolution of 25.

The model Earth spectrum for the visible range in Figure 2 shows five curves representing spectra from the surface, each of three cloud layers, and a weighted average of these components. Because the clouds are assumed to be the same at all wavelengths, their main

effect is to make the absorption lines appear less deep. The visible spectra are smoothed to a resolution of 100 using a triangle function.

Water (Figs. 3 and 4)

Water has a present atmospheric level (PAL) of about 8000 ppm, corresponding to about 50% relative humidity at the 288 K model surface. Water vapor is concentrated in a few-km layer near its liquid water source at the surface, falls to a minimum of a few ppm at the tropopause, and increases to about 6 ppm in the stratosphere. Water in the stratosphere is produced both by upward transport from the troposphere and by oxidation of methane that also is tropospheric in origin. The optical depth increases approximately as the square root of abundance, because most water lines are saturated. The abundance of H_2O increases exponentially with temperature, but it is independent of ambient pressure. Thus the information that we derive from the H_2O bands is a mixture of the surface H_2O availability, surface temperature, vertical mixing activity, vertical temperature profile, and photochemical reactions.

The mid-infrared spectrum of H_2O is shown in Figure 3. The upper curve is for zero water abundance (0 ppm), and is of course a flat line because here we have an atmosphere-free planet and we see down to the 288 K surface. The main features are the rotational bands (about 15 μm and longer wavelength) and vibrational bands (about 5 to 8 μm). Adding tropospheric H_2O produces opacity at altitudes above the ground, hence cooler molecular kinetic temperatures and lower emitted flux levels, and thus spectral features that appear to be absorption features but are actually simply emission from lower-temperature layers of the atmosphere. Four potential bands are indicated where the assumed extremes of wavelength are set by instrumental spatial resolution at long wavelengths and by the weakness of emitted flux at short wavelengths. The best features may be the 17- to 50-μm rotational bands, which give information on water in the stratosphere as well as troposphere. The 6- to 7-μm region may be a second choice, since CH_4 and N_2O can overlap this region.

The visible spectrum of H_2O is shown in Figure 4. The spectrum shows relatively equally-spaced bands that are very strong at long wavelengths and weaker at short wavelengths. In the 0.5- to 1.0-μm interval alone, there are 5 significant H_2O features with a range of strengths. Although H_2O tends to dominate the visible spectrum, it does so in discrete bands which happen to be centered at wavelengths such that most other species are still accessible. On an Earth-like planet, the H_2O features in the 0.7-, 0.8-, and 0.9-μm regions are good candidates for detection. The strong, longer wavelength, bands in the 1.1-, 1.4-, and 1.9-μm areas are excellent candidates if this region is also instrumentally accessible.

Carbon dioxide (Figs. 5 and 6)

Carbon dioxide (CO_2) has a PAL of 355 ppm, uniformly mixed. The infrared spectrum (Figure 5) shows both the target species and also a background spectrum due to the present-Earth H_2O abundance. We believe that including the spectrum for H_2O provides a useful element of practicality because H_2O may well be a dominant spectral contributor. The main CO_2 band is centered at 15 μm and it is so strong that it is saturated for all mixing ratios shown. The central reversal in the 15-μm band at high abundances is caused by the Earth's stratospheric temperature inversion (which is in turn caused by absorption of solar UV by O_2 and O_3, and conversion to

heat). Weaker bands show up at 9 to 11 μm when the CO_2 abundance climbs to around 1%. Measurements of even higher abundances of CO_2, around 10%, may be possible, and these will encroach further into the 8- to 13-μm window.

Visible wavelength CO_2 features are shown in Figure 6, superposed on a background of present-Earth H_2O. The shortest wavelength feature of CO_2 is at 1.06 μm, and this will only become significant when the abundance is very high, on the order of 10%, which could well be the situation for an early-Earth atmosphere. The next shortest wavelength band is at 1.2 μm, in the wing of a H_2O band but nevertheless still appreciable at abundances of 1% and higher, also an early-Earth indicator. The next available band of CO_2 is at 1.6 μm, located between H_2O bands. Here the strength is significant for abundances of CO_2 only about three times greater than present on Earth. Thus if the full range from 1.0 to 1.7 μm is available, this spectrum provides estimates across a wide range of CO_2 abundances.

Ozone (Figs. 7, 8 and 9)

Ozone (O_3) has a PAL of 6 ppm in the stratospheric "O_3 layer" around 25 to 35 km, with a lower abundance tail extending down to the surface. In the infrared (Figure 7) the main feature is at 9 μm, with a weaker band at 14 μm. However, if CO_2 is present in even small amounts the 14-μm feature will be mostly blocked. The 9-μm band is highly saturated, and is essentially unchanged as the O_3 abundance varies from 1 to 6 ppm. This makes the 9-μm band a poor quantitative indicator of O_3, but, what is likely much more important is that it is an excellent qualitative indicator of the existence of even a trace amount of O_2. Silicate minerals and O_3 both have strong features in the 9-μm region, but there is little chance that the two could be confused because their spectral shapes are quite distinct. The closest match of a silicate feature to O_3 is that of the mineral illite, and even in that case the band shapes are readily distinguishable, based upon even our present knowledge.

Visible-wavelength O_3 has a broad feature extending from about 0.5 to 0.7 μm, the Chappius bands, shown in Figure 8. (A weak H_2O band falls in the middle of this broad O_3 feature, distorting its nominally triangular shape.) The abundance levels indicated in the figure represent the stratospheric O_3 layer, although the tropospheric abundance does contribute a small amount to the observed band. For quantitative analysis, the Chappius bands might offer advantages because their absorption depths increase roughly linearly with abundance, and in particular, they are not saturated at the present abundance level. The breadth of the Chappius bands is at once an advantage, because only low resolution is needed to detect them, and a disadvantage, because it is generally harder to distinguish broad features than narrow ones against a potentially complex or unknown background spectrum.

Ultraviolet absorption by O_3 is very strong in the Huggins bands (Figure 9), which start to absorb at about 0.34 μm and increase dramatically toward 0.31 μm and shorter wavelengths. At PAL these bands are opaque from about 0.32 μm shortward, and will be present for even very small amounts of O_3. Although there are perhaps more experimental difficulties working in the UV than at longer wavelengths, the extreme sensitivity of the Huggins bands makes this region an obvious potential target for a biomarker search.

Methane (Figs. 10 and 11)

Methane (CH_4) has a PAL of 1.6 ppm, uniformly mixed in the troposphere and decreasing in the stratosphere. In the mid-infrared range (Figure 10) the main CH_4 feature is at 7-μm wavelength, where it is overlapped by the 6-μm band of water and the adjacent bands of N_2O (both of which are in the background spectra in this figure). The CH_4 feature nevertheless produces a weak absorption even at PAL, however the combined effect of a falling Planck curve and the extra opacity contributed by water make this a possible but difficult observation. However, during its history, Earth has likely witnessed two types of high-CH_4 states: first, during the initial half of its life when the composition was primarily reducing, and second, during later "CH_4 burst" conditions where a 1% mixing ratio likely existed for some period. In both cases the CH_4 band is strong enough to be seen in the wing of the water-vapor band at 8 to 9 μm, therefore a detection may be possible.

In the visible to near-infrared-range, CH_4 (Figure 11) has five relatively unobscured absorption features between 0.6 and 1.0 μm, and two more at 1.7- and 2.4-μm wavelength. The features at 0.6, 0.7, 0.8, 0.9, and 1.0 μm have significant depth for high abundances in the range 1000 ppm to 1%, which is a range of great interest for the early Earth. The 1.7- and 2.4-μm features become significant at CH_4 levels of 100 ppm and above. Thus, none of the visible features are useful at present abundance levels, but they could be very significant for 100 times or more PAL.

Nitrous oxide (Fig. 12)

Nitrous oxide (N_2O) has a PAL of 310 ppb, uniformly mixed in the troposphere but disappearing in the stratosphere. In the mid-infrared (Figure 12), N_2O has a band near 8 μm, roughly comparable in strength to the adjacent CH_4 band but weak compared to the overlapping H_2O band. The background spectrum in Figure 12 includes both H_2O and CH_4. Fortunately, the absorption patterns of these three species are different, so in principle their contributions may be separable. Spectral features of N_2O will become progressively more apparent in atmospheres having less H_2O water vapor. An additional band at 17 μm may well be totally obscured by CO_2 therefore it should not be expected to be useful. There are no significant N_2O features in the visible range. Atmospheric concentrations of N_2O will decrease with decreasing O_2 concentrations.

Oxygen (Fig. 13)

Oxygen has a PAL of 21%, uniformly mixed. In the visible (Figure 13) there are three potentially significant features, at 0.69 (Fraunhofer B band), 0.76 (Fraunhofer A band), and 1.26 μm. Among these, the 0.76-μm feature is the strongest one, having appreciable depth at an abundance level of 1% and greater, making it potentially very useful as a biomarker. All three bands are essentially unobscured in the present Earth spectrum.

In the infrared there are no significant O_2 features. In the sub-millimeter region there are strong O_2 lines, however these wavelengths are much longer than those considered here.

Band list (Table 1)

The spectral bands discussed here are listed in Table 1, starting with the infrared bands within the 7- and 50-μm range (including the two continuum windows that are useful for temperature measurement). The visible bands between 0.3 and 2.5 μm are then listed (of which most lie between 0.3 and 1.1 μm). The table columns give the species name, the wavenumbers of the bands (minimum, maximum, and average), the spectral resolution (ave/(max-min)), and the corresponding wavelength values. The wavenumber entries were determined graphically by visual inspection of the figures, and thus are a function of both each species itself as well as any encroachment by neighboring spectral features. The FWHM values ("Full Widths at Half-Maximum intensity," or spectral resolutions) are intended to represent approximately optimum detection bandwidths, in the sense of a matched filter, based on the assumption that the spectra will be determined by only the species considered here. Therefore if significantly different spectra are anticipated, then higher spectral resolutions might be needed in order to permit unambiguous feature identifications.

Curves of growth (Table 2)

The average depth of each molecular absorption is listed in Table 2, as a function of mixing ratio. Each depth is calculated from the continuum levels indicated in the figures, that is, for H_2O, the depth is measured from a flat, featureless continuum and, for all other species, the depth is measured with respect to the nominal water continuum indicated in the figures. Thus, the depth values include an element of realism in the context of planet detection (concerning the presence of water, at least) that would not be present if the calculations were all done with respect to a featureless background. For each spectral band the integration is done between the nominal FWHM limits drawn in each figure and listed in Table 1. Thus, the absorbed area could also be written as the product of FWHM and depth. For example, the average depth of the O_2 Fraunhofer A band at 13105 cm^{-1} is 0.47 at PAL and in a cloud-free atmosphere.

IV. Spectral Features from the Planet's Surface

A planet's spectrum should vary due to seasonal changes, weather, and different surface types rotating in and out of view. On a timescale longer than that of a rotational period or typical change in weather, these effects should average out and not present any serious complications to detecting atmospheric biomarker spectral signatures. However, there are potential benefits associated with observing variations over time. For example, it may be possible to derive the planet's rotational period, existence of weather (time varying water clouds are indicative of water oceans), surface biomarkers, existence of water oceans, and perhaps even ocean or ice fraction. One might expect that the differing albedos and surface temperatures from different parts of an unresolved Earth-like planet will cancel each other. This is not correct mostly because of nonuniform illumination and viewing angle, and, in fact, a relatively small part of the visible hemisphere dominates the total flux from a spatially unresolved planet. Because of this, there is potential to detect surface features.

Flux variations in the optical wavelength range that occur on a rotational timescale and that are caused by different types of surfaces rotating in and out of view (for example, oceans including specular reflection, land, and ice cover) can be as high as 150 to 200% for a cloud-free Earth. Flux variations can be as high as 650% for an ice-covered planet with ice-free liquid oceans. These numbers are reduced to 10 to 20% in the case of Earth-like cloud cover. (For more details, see *[Turner et al., 2000]*.) This variation is a direct consequence of the large differences

in albedo between the ocean, land, and ice (<10% for ocean, >30 to 40% for land, >60% for snow and some types of ice). Discriminating between these different surfaces need not require a specific spectral signature, therefore the required observation time could be shorter than that for detecting spectral features.

It may be possible to detect surface biomarkers if they make up a large fraction of a planet's surface (for example, *[Schneider et al., 2000]*). An interesting example from the Earth is the "red edge" signature from photosynthetic plants at ≈750 nm where the reflectivity changes by almost an order of magnitude. This is much greater than the reflectivity change on either side of green wavelengths (less than a factor of 2) due to chlorophyll absorption. Photosynthetic plants have developed this strong infrared reflection as a cooling mechanism to prevent overheating which would cause chlorophyll to degrade. A simple calculation for the integrated reflectivity difference between a vegetation-free Earth and our own Earth (assuming that 1/3 of the Earth is covered with land and that 1/3 of the land is covered with vegetation) gives 2%. Considering geometry, time resolution on a rotational timescale, and directional scattering effects, this number should be considerably larger when a large forested area is in view. A detailed calculation is needed. (See *[Sagan et al., 1993]* for a detection of this photosynthetic vegetation signature from a small part of the Earth.) Some photosynthetic marine life also has a wavelength-dependent signature similar to land vegetation. In addition, phytoplankton blooms can cause a temporal change in large areas of the ocean. The ocean is very dark in the optical and has strong water absorption bands in the infrared, however, and so most of the reflected flux from the Earth does not come from the ocean even though it is a large fraction of the surface area. This means it will be difficult to detect the color difference due to spatial or temporal variability of photosynthetic marine organisms.

For a planet with nonzero obliquity, the seasonal flux variation might also be detectable. Total seasonal change for our Earth's global albedo is smaller than the expected rotational variation (≈7% from simulations of Earthshine by *[Goode et al., 2001]*, but this could be much different for planets with different obliquities or with different orbital inclinations. Rotational variation in the mid-infrared flux could also be detected, but for the Earth is expected to be lower than the optical flux variation because the surface temperature does not vary as much as the surface albedo across the Earth. In the mid-infrared the seasonal variation could be larger than the mid-infrared rotational variation because of seasonal temperature changes. Planets having uniform cloud cover such as Venus would show no rotational or seasonal change in the spatially unresolved flux.

A TPF architecture that is capable of detecting spectral features at high signal/noise will also be able to detect 10 percent changes in the flux. The opportunity to derive physical properties of the planet on a rotational or seasonal timescale from the TPF data is an important addition to the main goal of TPF to detect and characterize Earth-like planets. There is a large parameter space of possible physical characteristics of Earth-like extrasolar planets and a more careful study of time variation and surface features is recommended.

V. Wavelength Ranges and Prioritization of Spectral Features

Wavelength range

What would be the minimal wavelength coverage that we would hope to achieve with TPF in order to detect Earth-like planets and possibly life?

In the thermal infrared, it is absolutely essential to observe the 15-μm CO_2 band and the 9.6-μm O_3 band. O_3 is our best biomarker gas in the mid-infrared range. We also want to observe the entire 8 to 12-μm "window" region because that gives us our best chance at estimating the planet's surface temperature. Then, we need to have some measure of H_2O, which we can get from either the 6.3-μm band or the rotation band, which extends from 12 μm out into the microwave region. Finally, it would be helpful to observe the 7.7-μm band of CH_4 because this is potentially a good biomarker gas for early-Earth type planets. Thus, the minimum wavelength coverage would be 8.5 to 20 μm, and we would prefer to get coverage from 7 to 25 μm.

In the visible/near-IR, it is essential to observe the 0.76-μm band of O_2. The broadband 0.45- to 0.75-μm O_3 absorption will need about the same signal/noise value and, on a cloud-covered planet or one with a lower oxygen abundance, may well be easier to detect. We also need at least one strong H_2O band and suggest 0.94 μm, which we can obtain by going out to 1.0 μm. CO_2 is much more difficult to observe in the visible/near-IR. If the planet is CO_2-rich, our best bet is at 1.06 μm, which would require wavelength coverage out to at least 1.1 μm. This should not be a driver, though, because this band is weak, even in the spectrum of Venus. For terrestrial gas abundances, the shortest wavelength bands that show well are at 2.0 μm and 2.06 μm. Finally, the best CH_4 band shortward of 1.0 μm is at 0.88 μm, though it only shows at considerably greater abundance than terrestrial. Required wavelength coverage would be 0.7 to 1.0 μm, and we would prefer to see ≈0.5 μm (to look for broadband absorption by O_3) to ≈1.1 μm (to detect CO_2).

O_3 might be detected in UV range (at 0.34 to 0.31 μm), however more studies are required to evaluate potential interferences.

Prioritization

Deciding which features in each wavelength regime are most important to observe is not a simple task. A general consensus, though, seems to have been reached regarding the following points:

1) Detection of O_2 or its photolytic product O_3 is our highest priority because O_2 is our most reliable biomarker gas. Possible "false positives" for O_2 have been *identified [Kasting, 1997] [Leger et al., 1999]*. One such pathological case involves the abiotic production of O_2 from H_2O photolysis followed by rapid hydrogen escape from a runaway greenhouse atmosphere. Another such case involves buildup of O_2 from the same abiotic process on a frozen planet somewhat larger than Mars (0.1 to 0.2 times Earth's mass). The frozen surface would keep O_2 from reacting with reduced minerals in the crust, while the mass range would preclude nonthermal escape of O atoms without creating enough internal heat to sustain volcanic outgassing of reduced gases. Both of these cases could presumably be identified spectroscopically. For a "normal" Earth-like planet situated within the habitable zone, free O_2 is a reliable indicator of life.
2) O_3 is as reliable a bioindicator as O_2. However, it provides somewhat different information. Because of its nonlinear dependence on O_2 abundance, O_3 is easier to detect at low O_2

concentrations. However, it is a relatively poor indicator of how much O_2 is actually present. It is difficult to say which type of information is more useful. A positive identification of either gas would be very exciting and significant.

3) Another category of important observable features includes water vapor and the 8- to 12-μm continuum. H_2O is not a bio-indicator, however, its presence in liquid form on a planet's surface is considered essential to life. Unfortunately, observations that show gaseous water vapor alone do not sufficiently constrain the conditions to determine whether the atmosphere is near saturation. Observation of the 8- to 12-μm continuum could, in some cases (such as present Earth), allow us to determine a planet's surface temperature, which is also an important constraint on life. However, the atmospheres of planets that are more than ≈20 K warmer than Earth (that is, >310 K) will be opaque in this region because of continuum absorption by water vapor. Such planets are also susceptible to a runaway or "moist" greenhouse effect *[Kasting and Brown, 1998]*, so the range of conditions where this occurs and the planet is also habitable will likely be limited. Planets with one-bar atmospheres and surfaces warmer than ≈340 K would lose their water relatively quickly *[Kasting and Brown, 1998]*. Cloud cover (as on Venus) could also obscure the surface and preclude the determination of temperature. Indeed, it would be difficult or impossible for observations to distinguish a planet like Venus (with a cold cloud layer) from a planet with a cold, ice covered surface. Thus, although there is an argument in favor of observing in the mid-infrared range, it is balanced by there being both O_2 and O_3 available in the visible. We consider that the technological issues (about which wavelength region is easier to observe with appropriate sensitivity) are more important.

4) The carbon-containing gases CO_2 and CH_4 can each provide useful information. In the mid-infrared range, CO_2 is the easiest of all species to observe. CO_2 is required for photosynthesis and for other important metabolic pathways. Furthermore, it provides good evidence that we are dealing with a terrestrial planet. It is hard to imagine how a moist, rocky planet from which hydrogen can escape could *not* have CO_2 in its atmosphere. Even if carbon was outgassed in a more reduced form, some fraction of it ought to be oxidized to CO_2 by reaction with oxygen formerly in water.

   The presence of CH_4 would provide interesting, but ambiguous, information. According to at least some models of atmospheric evolution (for example, *[Kasting and Brown, 1998]*) CH_4 is expected to have been present at mixing ratios of $\approx 10^{-3}$ in the Late Archean atmosphere (2.5 to 3.0 billion years ago). The high concentrations of CH_4 at this time would have indicated production by methanogenic bacteria. But CH_4 could have been produced abiotically as well. About one percent of the carbon released from midocean ridge volcanism today is in the form of CH_4. The rest is CO_2. If the early mantle was more reduced, most of the carbon released from submarine outgassing could have been in the form of CH_4, and abiotic CH_4 mixing ratios could have exceeded 10^{-4} *[Kasting and Brown, 1998]*. So, we would probably require additional information to decide whether a CH_4-rich atmosphere was really an indication of life.

5) Determining the size of the planet is an important component of our understanding of habitability. By equating color temperature for selected spectral regions with physical temperature, and using the observed flux, we can use Planck's Law to obtain the surface area of the planet. We can use the observed solar system relationship between mass, radius and thermal environment, to infer the likely planet mass, and to consider whether this is or is not likely to be a planet with current or past operative geological activity. Geological

activity is important for maintaining chemical resources for life as well as habitable conditions at the surface of the planet. Planet size can be estimated in the mid-infrared, but not in the visible to near infrared range.

The technological issues that determine which wavelength region is easier to observe with appropriate sensitivity are more important than any scientific priorities between these last three features.

Executive Summary

A chief goal of TPF is to deliver data to the biological and atmospheric science communities who can then evaluate these observations for evidence of environmental conditions and life. The spacecraft must directly detect and survey nearby stars for planetary systems that include terrestrial sized planets in their habitable zones ("Earth-like" planets). TPF should characterize a broad diversity of planets, but it must target the most prospective of these for more detailed spectroscopy. TPF should determine whether biomarkers are present.

This report reviews the opportunities and fundamental constraints that attend remote spectroscopic observations of planets; it recommends wavelength ranges and spectral features; it summarizes the rationale for these selections, and it provides algorithms for detection of spectral features. Mid-infrared spectra ("MIR": ≈5- to ≈50-μm range) can give the planetary albedo, the temperature of the observable emitting regions and therefore the planet's size. The potential for planetary rings, dust tails and moons to compromise estimates of planetary size must be kept in mind. Visible near-infrared spectra ("VIS-NIR": ≈0.4 to ≈2.5 μm) offer higher spatial resolution, are minimally affected by temperature and therefore are good for determining the abundance of atmospheric species. Conditions near the surface of a planet can be attenuated or totally obscured by clouds, water vapor, etc., thus some observations might actually indicate properties of the upper atmosphere. A careful consideration of environmental effects on spectra can reveal the thermal and pressure structure of the atmosphere. Observations in both the MIR and VIS-NIR ranges can give far more definitive results about conditions and composition than can a single range by itself.

Observations of the planet's surface and related temporal variations can, in principle, reveal the planet's rotational period, existence of weather, surface biomarkers, and existence of water oceans and/or surface ice. The opportunity to derive physical properties of the planet on a rotational or seasonal time scale would be an important addition to the characterization of Earth-like planets. Studies of planetary systems will also reveal how the abundances of life-permitting volatiles on an Earth-like planet are related to the characteristics of the planetary system as a whole.

Algorithms are provided to establish the detection of both spectral continua and spectral lines or bands. When the observable spectral continuum is used to detect the planet, the greatest possible fraction of the strongest parts of the continuum should be observed. Discrete spectral bands must be resolved against a continuum, therefore both must be defined in order to achieve a successful detection. Due to potential sources of noise, any result should be the sum of at least three independent measurements that, taken together, achieve a signal/noise value of at least 5σ (sigma). For each of the operational scenarios for the various proposed TPF architectures, the

appropriate data reduction process should be scrutinized and compared with the formulae proposed in this white paper to calculate the signal/noise value.

Model Earth spectra and related data were calculated with the Smithsonian Astrophysical Observatory code for both VIS-NIR and MIR. Spectra are presented graphically for the following: cloud abundance (no coverage, partial coverage, low and high altitude clouds), H_2O vapor (present atmospheric level [PAL]), CO_2 (PAL up to 1%, without and with water vapor background), O_3 (PAL), CH_4 (PAL up to 1%), N_2O (PAL, without and with methane and water background), and O_2 (PAL down to 1%). For each of the 37 spectral bands that were selected to represent these key species, a table lists the wavenumber, corresponding wavelength, and the spectral resolution required. Another table gives the average depth of each spectral feature as a function of the atmospheric mixing ratio of the gas.

The minimum required MIR wavelength coverage from 8.5 to 20 μm includes CO_2, O_3, and H_2O. The preferred MIR coverage from 7 to 25 μm adds CH_4 and H_2O rotation bands. The minimum required VIS-NIR wavelength coverage from 0.7 to 1.0 μm includes O_2, H_2O, CH_4, and O_3. The preferred VIS-NIR coverage from 0.5 to ≈1.1 μm adds CO_2 and the broadband absorption by O3. O_3 might be detected in UV range (at 0.34 to 0.31 μm), however more studies are required to evaluate potential interferences.

Detection of O_2 or its photolytic product O_3 is the highest priority because it is the most reliable biomarker gas. However, we must be cautious of "false positives" due to abiotic O_2 sources. Three additional features share equivalent priorities. First, water vapor bands can indicate the presence of liquid water, which is essential for life, but they can also indicate H_20 situated high in a Venus-like atmosphere or else an ice-covered planet. Second, the carbon gases CO_2 and CH_4 offer multiple benefits. CO_2 is required for photosynthesis and for other important metabolic pathways. The combination of CO_2 and CH_4 provide useful information about the planet's oxidation state; CO_2 can indicate a solar system-like terrestrial planet, and CH_4 might be a biomarker in cases where hydrothermal emission of CH_4 is relatively minor. Third, albedo and temperature of the observable emitting region of a planet can give its size, which is important for confirming the presence of a terrestrial planet and also indicating whether it is geologically active, an essential requirement for habitability. Planet size can be estimated in the mid-infrared, but not in the visible to near infrared range.

Both the mid-infrared and the visible to near infrared spectral ranges offer valuable information regarding biomarkers and planetary properties, therefore both ranges merit serious scientific consideration for TPF. The best overall strategy for the Origins program includes a diversity of approaches, therefore both wavelength ranges ultimately should be examined prior to launching the "Life-Finder" mission.

In determining the appropriate spectroscopic capabilities of TPF, we must anticipate that the actual range of characteristics of rocky planets is likely to exceed our experiences with the four terrestrial planets and the Moon. Therefore, the recommendations of this white paper advocates a design for TPF that moves beyond merely finding a twin of our own Earth, a design that helps us to embark upon the astronomical exploration of a much broader diversity of planets.

The molecular species and spectral bands used in this study are listed in Table 1. Two infrared continuum bands are also given, where in a cloud-free atmosphere, emission from the surface might be seen. Columns 3 to 5 give the nominally optimum wavenumber values (minimum, maximum, and average) for each band. Column 6 gives the corresponding spectral resolution. Columns 7 to 9 give the same information in terms of wavelength.

Table 1. Molecular Species and Spectral Bands Used in This Study

Band	Species	Sigma (cm^{-1})			Resol.	Lambda (μm)		
		Min	Max	avg		Min	Max	avg
1	H_2O	200.	300.	250.	2.	33.33	50.00	40.00
2	H_2O	300.	400.	350.	4.	25.	33.33	28.57
3	H_2O	400.	576.	488.	3.	17.36	25.	20.49
4	H_2O	1356.	1500.	1428.	10.	6.67	7.37	7.00
5	CO_2	587.	750.	668.	4.	13.33	17.04	14.96
6	CO_2	930.	990.	960.	16	10.75	10.10	10.42
7	CO_2	1046.	1102.	1074.	19.	9.56	9.07	9.31
8	O_3	1005.	1067.	1036.	17.	9.37	9.95	9.65
9	CH_4	1257.	1356.	1306.	13.	7.37	7.96	7.65
10	CH_4	11.50	13.56	1253.	6.	7.37	8.70	7.98
11	Cont.	804.	986.	895.	5.	10.14	12.44	11.17
12	Cont.	1082.	1226.	1154.	8.	8.16	9.24	8.67
13	H_2O	5080.	5580.	5330.	11.	1.79	1.97	1.88
14	H_2O	6740.	7480.	7110.	10.	1.34	1.48	1.41
15	H_2O	8580.	9050.	8815.	19.	1.10	1.17	1.13
16	H_2O	10320.	10930.	10625	17.	0.91	0.97	0.94
17	H_2O	12000.	12350.	12175.	35.	0.81	0.83	0.82
18	H_2O	13630.	14000.	13815.	37.	0.71	0.73	0.72
19	CO_2	4780.	5080.	4930.	16.	1.97	2.09	2.03
20	CO_2	6020.	6570.	6295.	11.	1.52	1.66	1.59
21	CO_2	8120.	8360.	8240.	34.	1.20	1.23	1.21
22	C O_2	9410.	9650.	9530.	40.	1.04	1.06	1.05
23	O_2	7840.	7950.	7895.	72.	1.26	1.28	1.27
24	O_2	13010.	13200.	13105.	69.	0.76	0.77	0.76
25	O_2	14380.	14650.	14515.	54.	0.68	0.70	0.69
26	O_3	15250.	19000.	17125.	5.	0.53	0.66	0.58
27	O_3	30000.	32000.	31000.	16.	0.31	0.33	0.32
28	CH_4	4040.	4570.	4305.	8.	2.19	2.48	2.32
29	CH_4	5610.	6190.	5900.	10.	1.62	1.78	1.69
30	CH_4	9790.	10280.	10035.	20.	0.97	1.02	1.00
31	CH_4	11040.	11390.	11215.	32.	0.88	0.91	0.89
32	CH_4	12420.	12850.	12635.	29.	0.78	0.81	0.79
33	CH_4	13660.	13900.	13780.	57.	0.72	0.73	0.73

Curve-of-growth values for each molecular band are listed in Table 2, first for the thermal emission infrared region, and second for the visible, near-infrared, and ultraviolet regions. For each species and abundance level, the average depth of each important spectral feature is listed, with the central wavenumber of the corresponding band noted at the top of each column. Refer to Table 1 for bandwidths and corresponding wavelength values. Entries range from weak lines (e.g., 0.029, or 2.9% average depth), to strong lines (e.g., 0.474, or 47.4% average depth).

Table 2. Curve-of-Growth Values for Each Molecular Band in this Study

Thermal						
H_2O	**250cm-1**	**350cm-1**	**488cm-1**	**1428cm-1**		
1ppm	0.090	0.043	0.072	0.048		
10ppm	0.213	0.124	0.025	0.161		
100ppm	0.341	0.279	0.068	0.427		
1000ppm	0.377	0.416	0.153	0.747		
10000ppm	0.339	0.436	0.226	0.850		
CO_2	**668cm-1**	**960cm-1**	**1074cm-1**			
100ppm	0.470	0.029	0.037			
350ppm	0.520	0.037	0.050			
1000ppm	0.548	0.054	0.075			
10000ppm	0.549	0.153	0.207			
O_3	**1036cm-1**	**710cm-1**				
1ppm	0.382	0.141				
3ppm	0.406	0.154				
6ppm	0.405	0.162				
N_2O	**592cm-1**	**1174cm-1**	**1290cm-1**			
100ppb	0.014	0.008	0.050			
310ppb	0.031	0.023	0.095			
1000ppb	0.063	0.063	0.160			
CH_4	**1306cm-1**					
0.5ppm	0.056					
1.6ppm	0.091					
5.0ppm	0.137					
100ppm	0.300					
1000ppm	0.330					
10000ppm	0.262					
Visible/near-ir/uv						
H_2O	**5330cm-1**	**7110cm-1**	**8815cm-1**	**10625cm-1**	**12175cm-1**	**13815cm-1**
10ppm	0.236	0.178	0.041	0.025	0.003	0.003
100ppm	0.586	0.485	0.172	0.124	0.025	0.024
1000ppm	0.901	0.830	0.500	0.401	0.118	0.130
10000ppm	0.990	0.988	0.885	0.795	0.379	0.441
CO_2	**4930cm-1**	**6295cm-1**	**8240cm-1**	**9530cm-1**		
100ppm	0.170	0.007	0.170	0.0002		
350ppm	0.309	0.030	0.309	0.0006		
1000ppm	0.443	0.065	0.443	0.001		
1pct	0.667	0.260	0.667	0.011		
10pct	0.714	0.566	0.714	0.062		
CH_4	**4305cm-1**	**\5900cm-1**	**10035cm-1**	**11215cm-1**	**12635cm-1**	**13780cm-1**
0.5ppm	0.005	0.003	0.063	0.001	0.0009	0.002
1.6ppm	0.009	0.012	0.011	0.002	0.0009	0.002
5.0ppm	0.111	0.039	0.025	0.004	0.0009	0.003
100ppm	0.462	0.298	0.039	0.060	0.010	0.010
1000ppm	0.587	0.630	0.315	0.417	0.032	0.073
10000ppm	0.627	0.814	0.881	0.818	0.267	0.455

Table 2. Curve-of-Growth Values for Each Molecular Band in this Study (con't)

O_2	7895cm-1	13105cm-1	14515cm-1
1%	0.023	0.150	0.025
10%	0.104	0.388	0.088
21%	0.153	0.474	0.124
50%	0.230	0.565	0.181
O_3	17125cm-1	31000cm-1	
1ppm	0.048	0.305	
3ppm	0.112	0.531	
7ppm	0.195	0.692	

Normalized infrared thermal emission spectral models of the Earth are shown in Figure 1 for five cloud conditions. Each spectrum is calculated as discussed in the text, for the present Earth's altitude-dependent, mid-latitude temperature structure and gas species mixing ratio profiles, and at a zenith angle of 30 deg to closely simulate a whole-Earth integration. Each spectrum is normalized to a blackbody spectrum at 288 K. A convenient way to interpret these spectra is to note that, at each wavelength, the emitted flux is that of a blackbody at a temperature that corresponds to the atmospheric temperature at an altitude where the optical depth is unity, where the optical depth is computed along a path from the top of the atmosphere downwards. For example, in the 20- to 100-μm region, the rotational band of water is so strong that we only see radiation from the relatively cold upper troposphere, and not from the hot surface or lower troposphere. The top curve represents a cloud-free atmosphere, with spectral features as explained in subsequent figures. The other curves represent complete overcast conditions with clouds at low, medium, and high altitudes, plus a curve for a nominally realistic mixture of these four cases. In the infrared spectrum, the main effect of adding clouds is to decrease the emitted continuum flux and reduce (and potentially invert) the relative depths of spectral features.

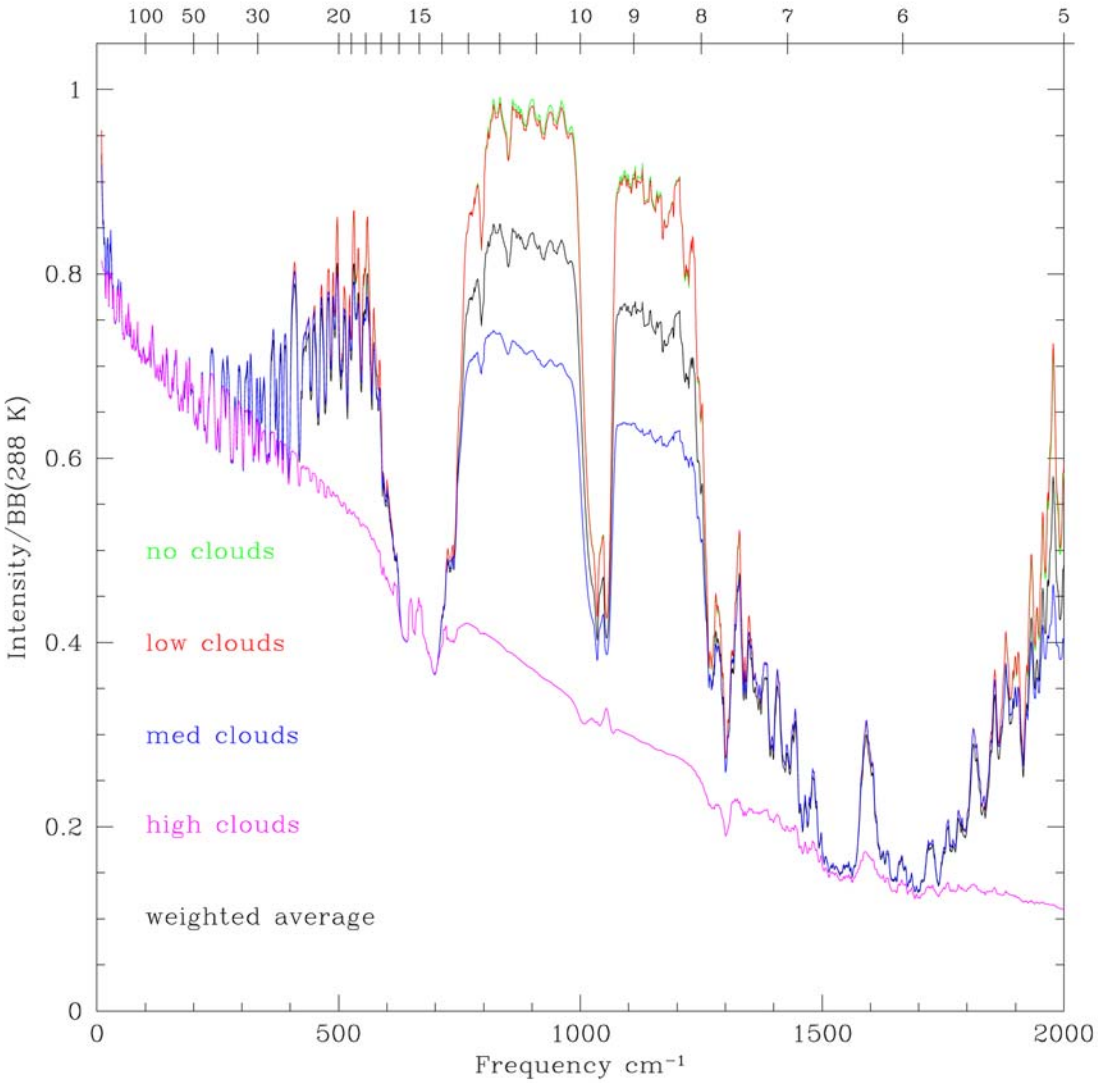

Figure 1. Normalized Infrared Thermal Emission Spectral Models of the Earth

Normalized visible reflection spectral models of the Earth shown in Figure 2 are for the same model atmosphere and cloud conditions as shown in Figure 1. The clear atmosphere case in the lower curve has the smallest continuum level owing to the fact that the surface reflectivity is low compared to clouds, and conversely, the overcast cloud cases have high continuum fluxes. The weighted average is a linear combination of the extreme cases, and was calculated for an overall albedo of about 39%, which is higher than the most recently available value of about 29.7% *[Goode et al., 2001]*. In the reflected spectrum, the main effect of adding clouds is to increase the reflected continuum flux and reduce the relative depths of absorption features.

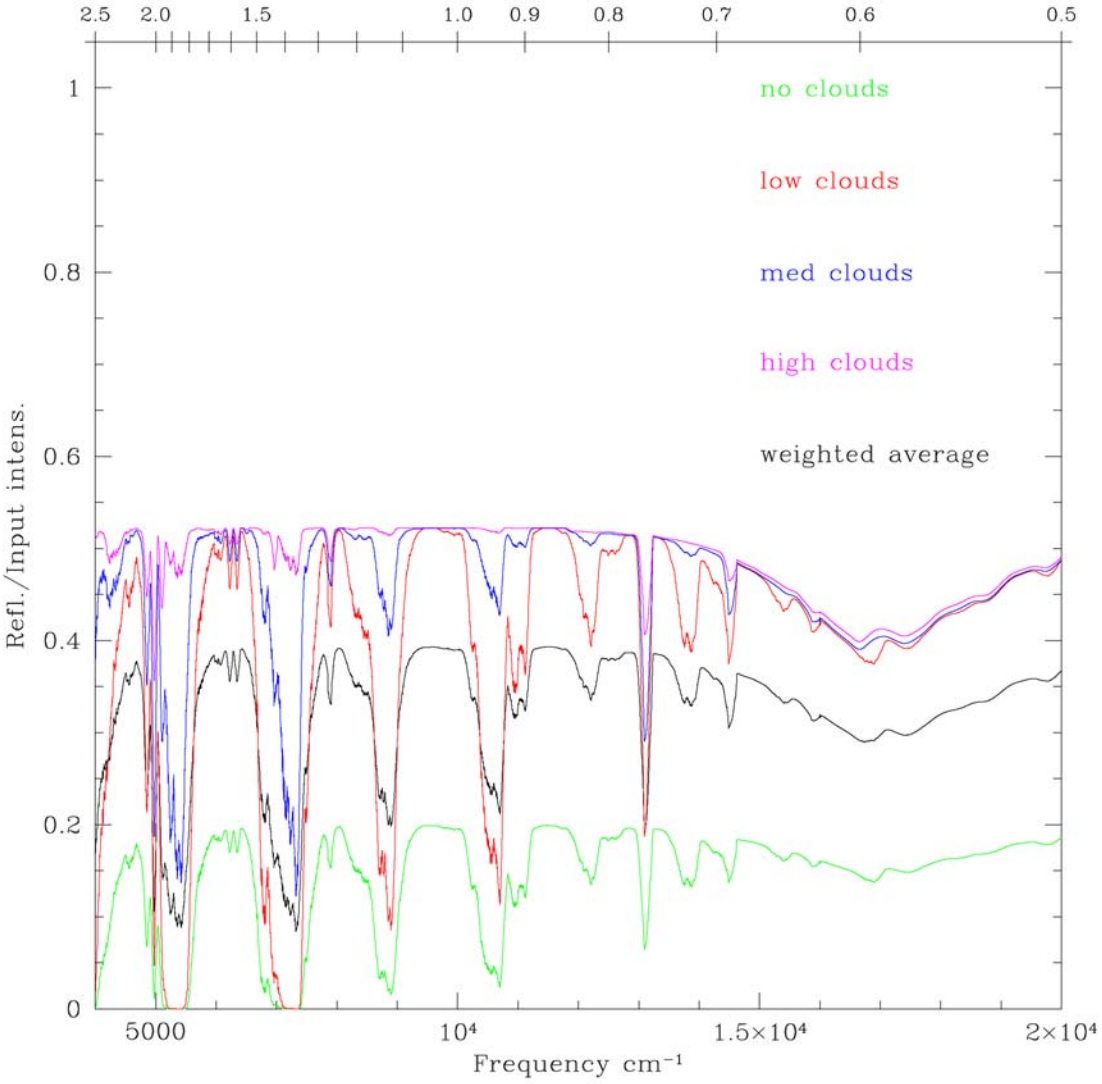

Figure 2. Normalized Visible Reflection Spectral Models of the Earth

The thermal emission spectrum of water is shown in Figure 3 for six humidity levels ranging from completely dry to essentially saturated. In each case the Earth's water vapor vertical mixing ratio profile is used, but scaled to give the surface mixing ratios indicated. Dominant features are the 15- to 100-μm rotational bands, and the 6-μm vibrational bands. Here and in subsequent figures, potential broad detection wavelength regions are indicated, chosen for their likelihood of producing a relatively clean detection of each species for an exo-solar terrestrial planet.

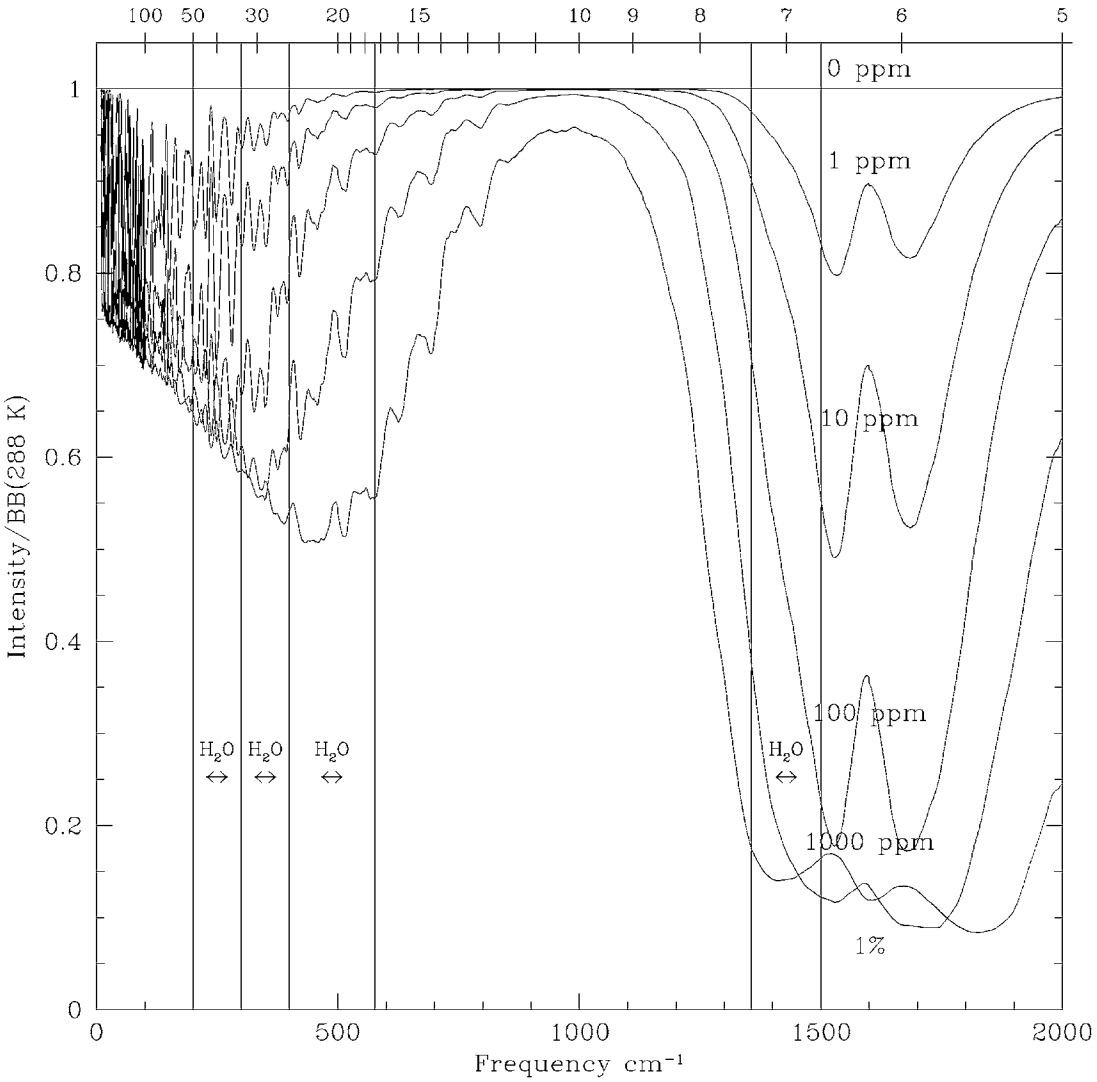

Figure 3. Thermal Emission Spectrum of Water

The reflection spectrum of water is shown in Figure 4 for five humidity levels, similar to the previous figure. Dominant features are at 0.7, 0.8, and 0.9 μm in the visible, and 1.1, 1.4, and 1.9 μm in the near infrared. Stronger features correspond to higher concentrations.

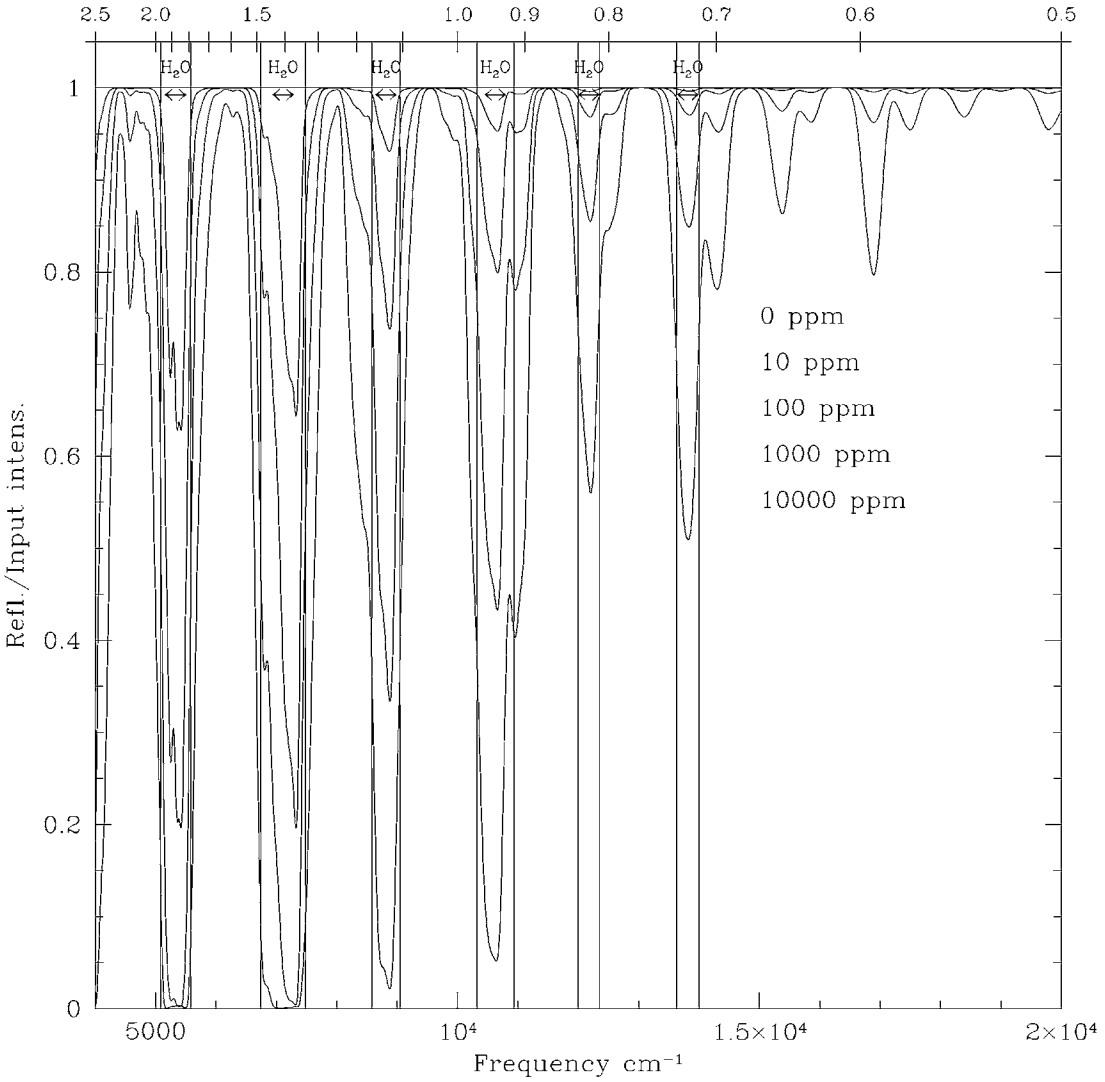

Figure 4. Reflection Spectrum of Water

The thermal emission spectrum of CO_2 is shown in Figure 5 for four abundance levels. A nominal level of water vapor is assumed to be simultaneously present. The 15-μm feature of CO_2 dominates the spectrum, but at high abundances the weak 9- and 10-μm bands also begin to show up. Note the inversion in the core of the 15-μm band at high abundances, due to increased opacity in the warm stratospheric layer.

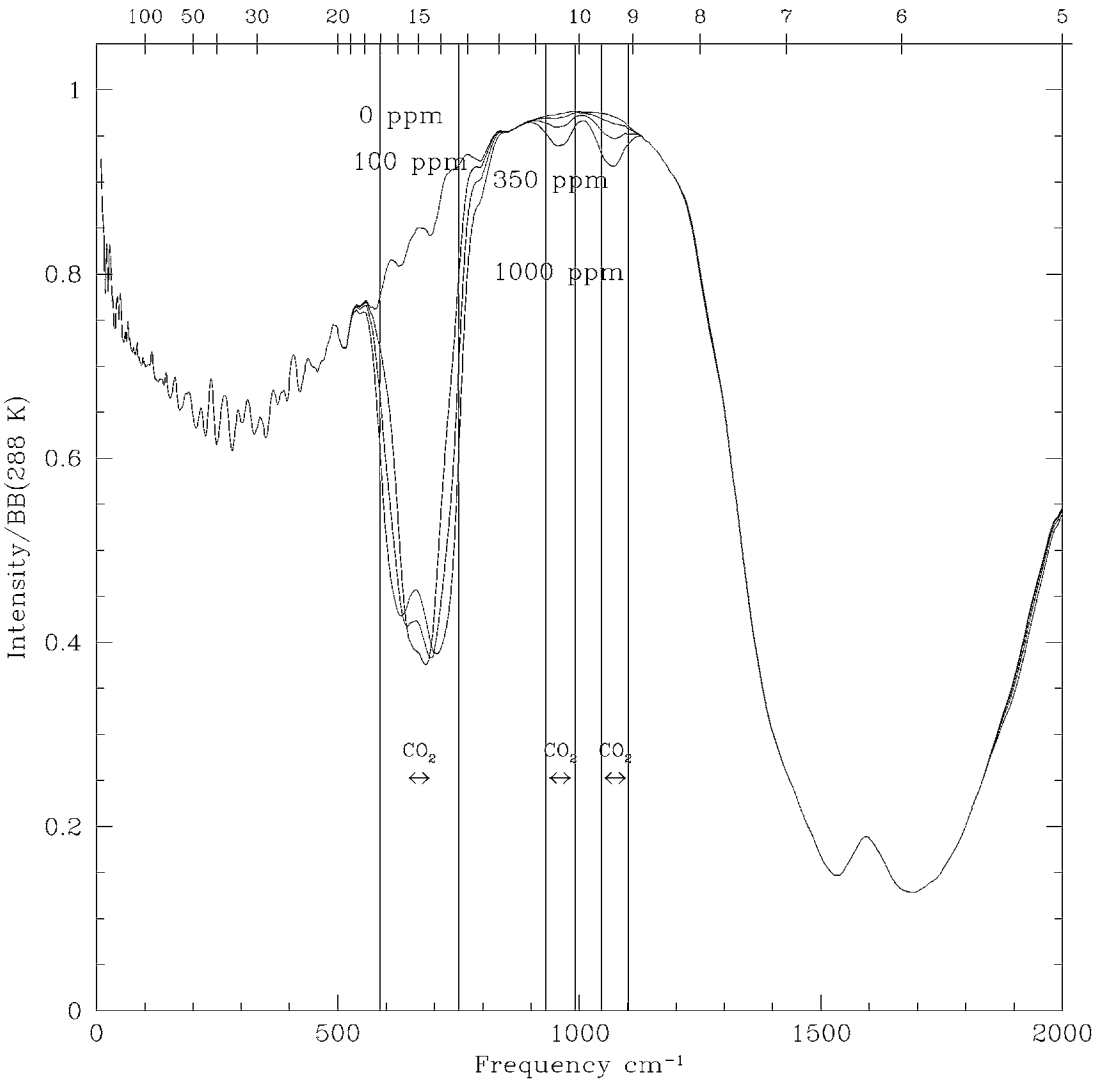

Figure 5. Thermal Emission Spectrum of CO_2

The reflection spectrum of CO_2 is shown in Figure 6 for six abundances. As in the preceding and following figures, a background abundance of water vapor is also included. At high abundances, CO_2 could be detected in the 1.06- and 1.3-μm bands, but to see lower abundances the 1.6- and 2.0-μm bands would be needed.

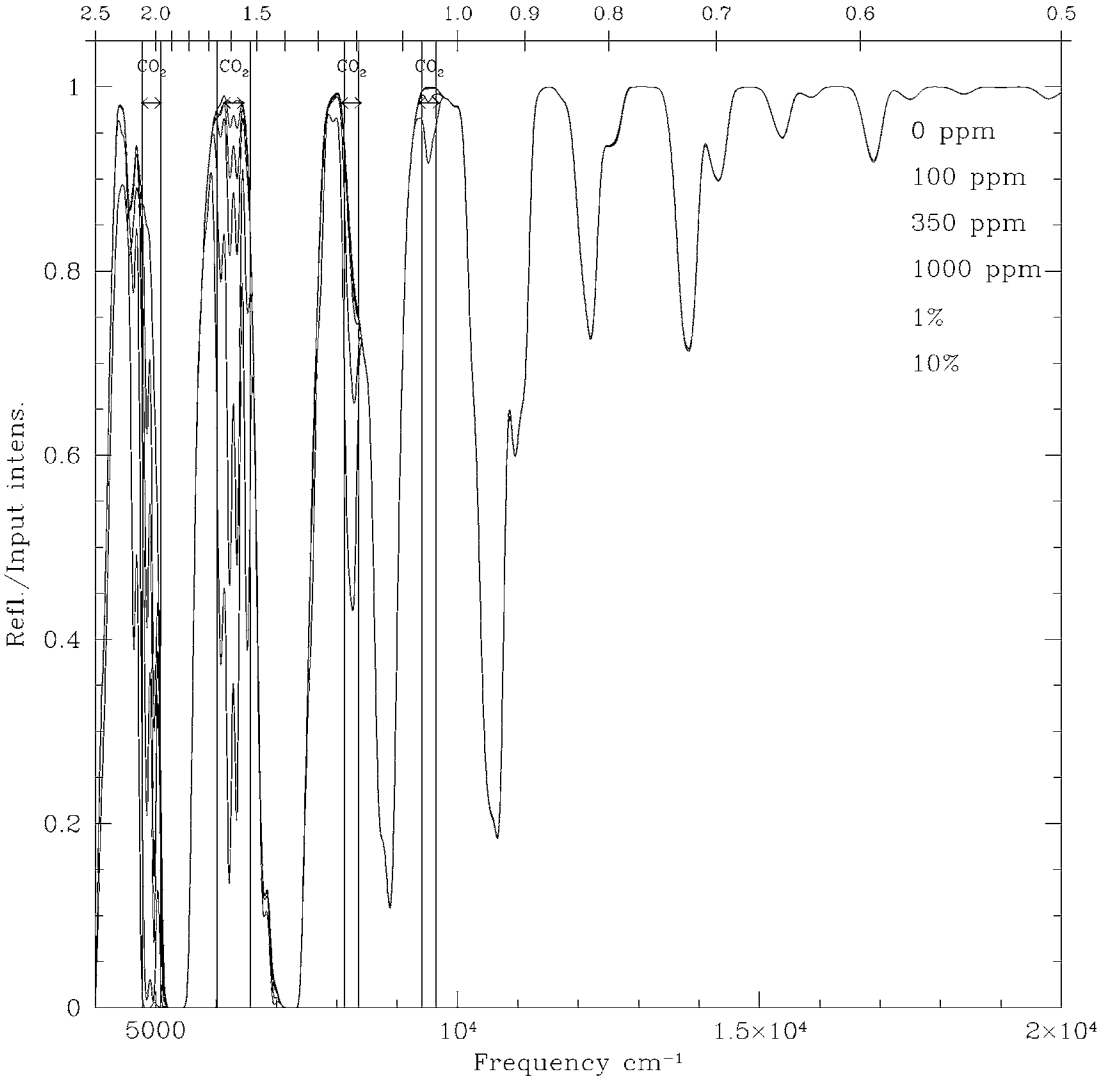

Figure 6. Reflection Spectrum of CO_2

The thermal emission spectrum of ozone is shown in Figure 7 for four abundance levels, where the present Earth's mixing ratio profile, dominated by the stratospheric ozone layer, has been scaled, and the value indicated is that at the peak of the ozone layer. The 9-μm feature is saturated for all indicated abundances, and only the 14-μm feature, which would be hidden by any CO_2 present, is on the linear part of the curve of growth.

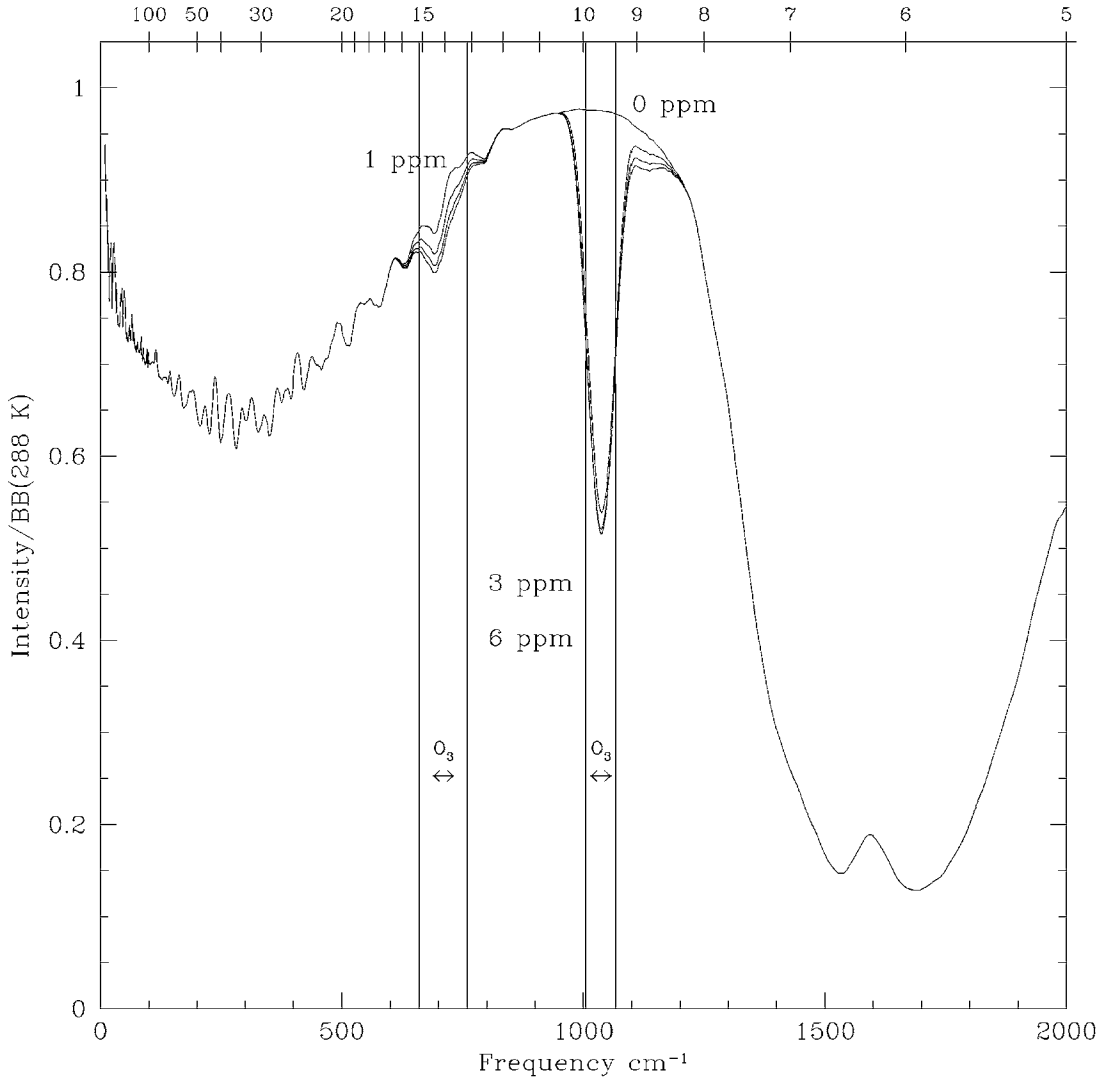

Figure 7. Thermal Emission Spectrum of Ozone

The visible reflection spectrum of ozone is shown in Figure 8 for four abundance levels, for conditions as given in the previous figure caption. The broad 0.6-μm band intensity responds linearly to variations in mixing ratio.

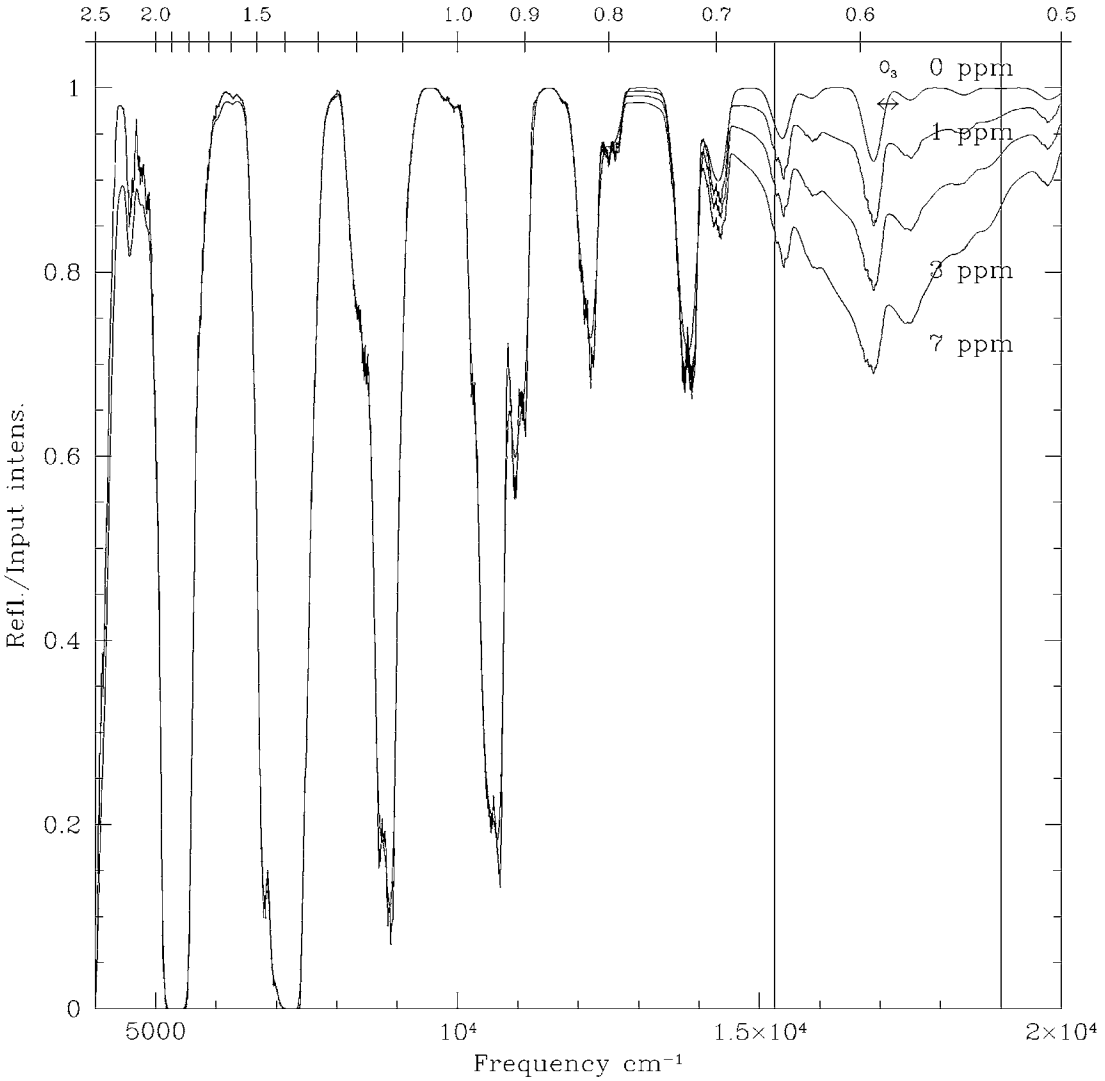

Figure 8. Visible Reflection Spectrum of Ozone

The ultraviolet reflection spectrum of ozone is shown in Figure 9 for three abundance levels in the stratospheric ozone layer, as discussed in the two preceding figure captions. A wavelength band over which the intensity is roughly linear is indicated, around 0.33 μm. The core of the ozone band is at a yet shorter wavelength and is completely saturated, but could be used for ozone detection at much lower abundances, if no other absorption feature were present.

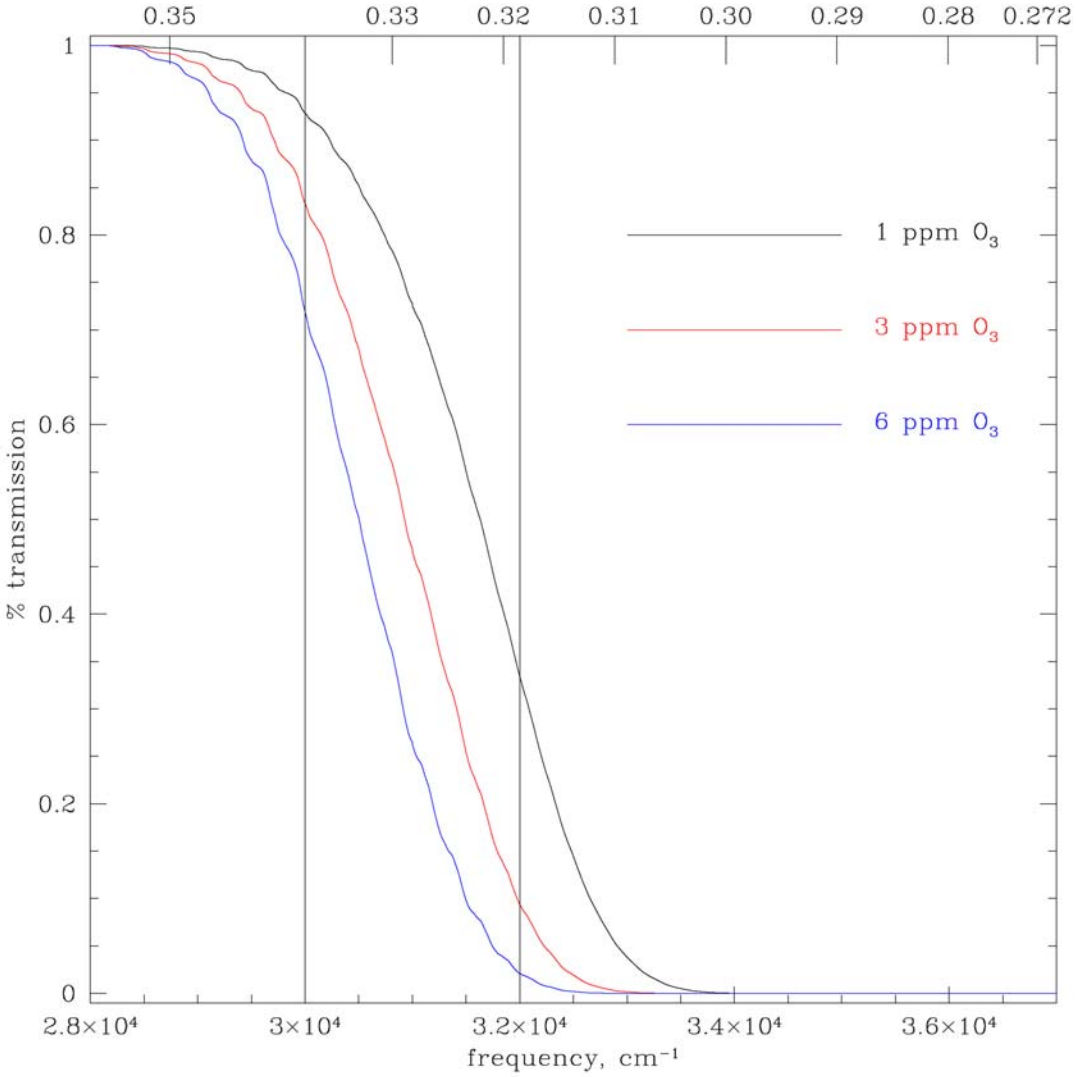

Figure 9. Ultraviolet Reflection Spectrum of Ozone

The thermal emission spectrum of methane is shown in Figure 10 for seven mixing ratios. The 7-μm band would show up more distinctly if it were not for the presence of water in the background, which in this case is a serious potential spectral contaminant. Present terrestrial methane levels will only produce a weak feature on the side of the dominant water band, but high methane levels, such as may have been present in the early Earth's atmosphere, could produce a stronger, more easily detectable feature.

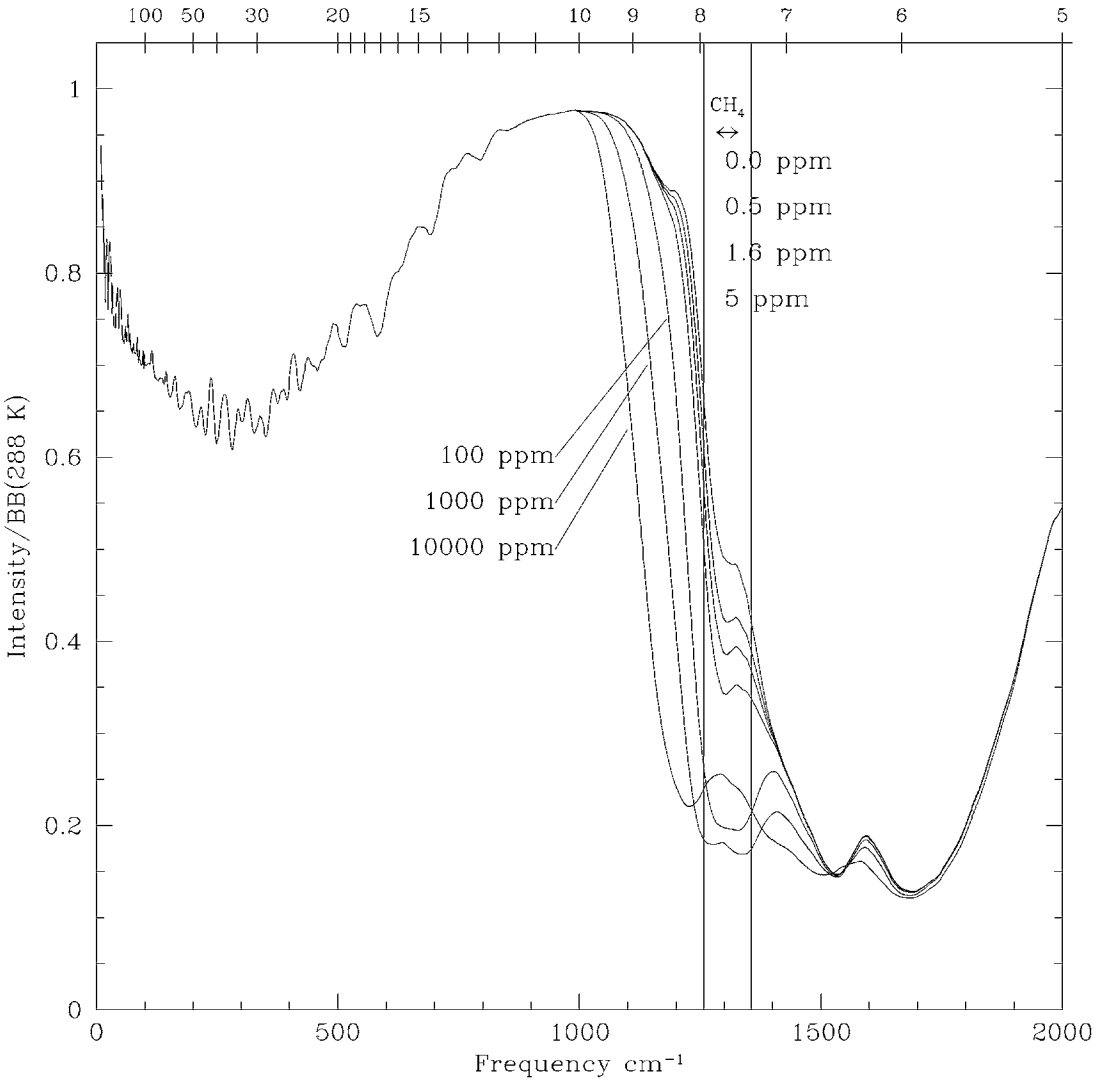

Figure 10. Thermal Emission Spectrum of Methane

The visible reflection spectrum of methane is shown in Figure 11 for seven abundance levels. At low abundances, methane would be detectable only at the longest wavelength, around 2.4 μm. However at moderate and high abundances methane could be detected in the 0.7-, 0.8-, 0.9-, and 1.0-μm bands.

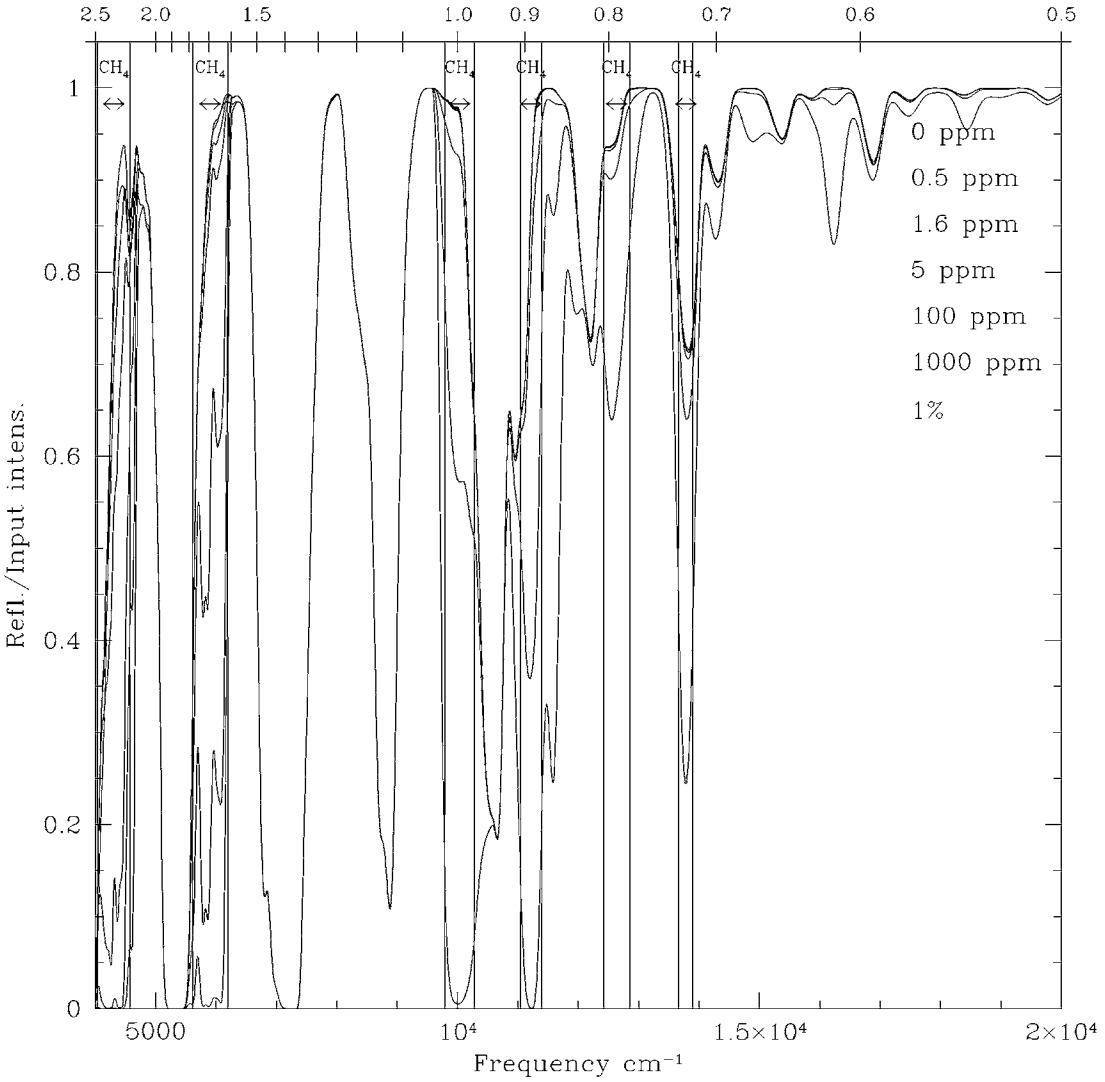

Figure 11. Visible Reflection Spectrum of Methane

The thermal emission spectrum of nitrous oxide is shown in Figure 12 for four abundance levels. As with methane, nitrous oxide only becomes detectable in the 8-μm region at high abundances, and it will be partially blended with methane as well as the dominant water feature. (There is no corresponding visible wavelength feature of nitrous oxide.)

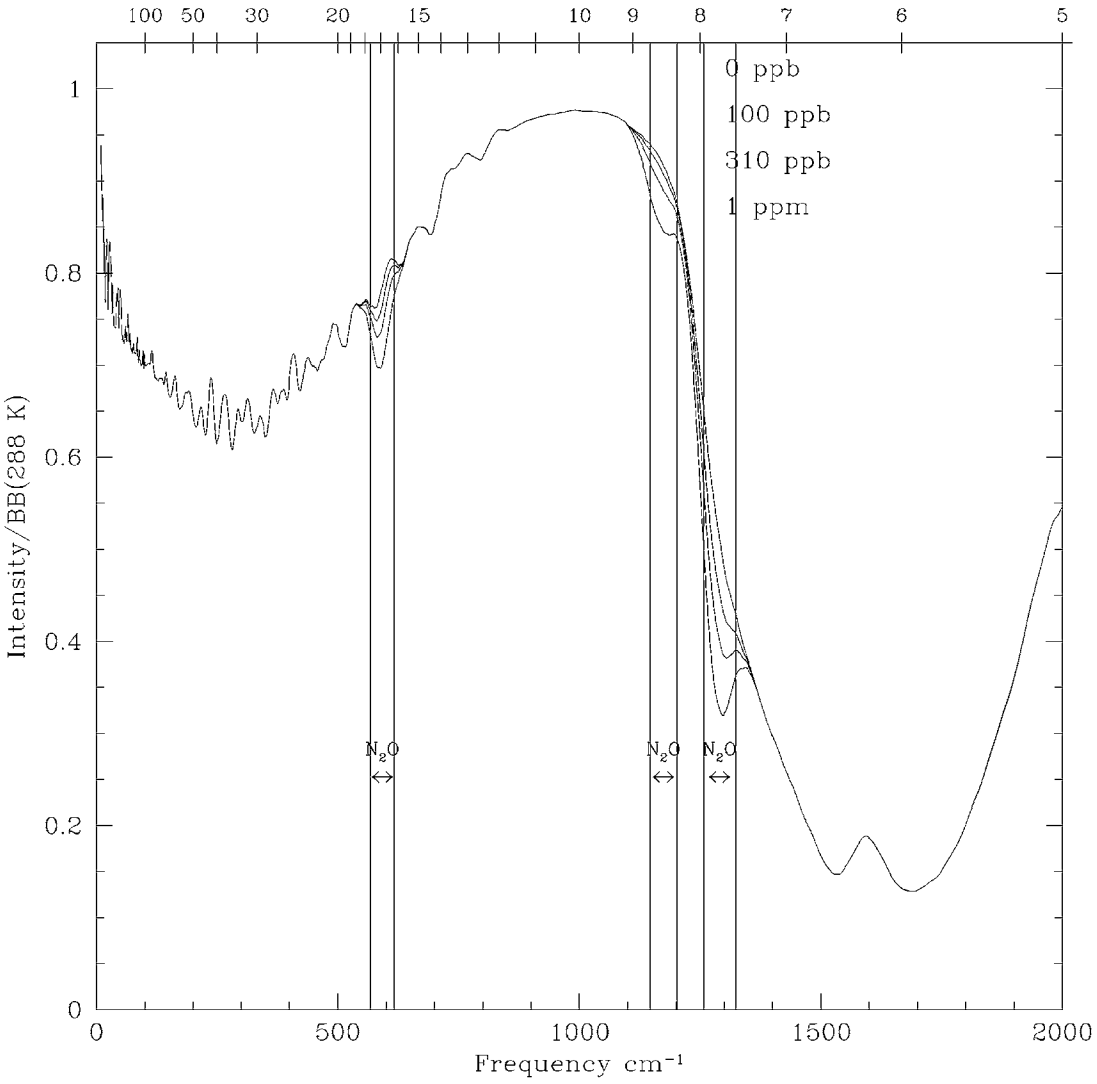

Figure 12. Thermal Emission Spectrum of Nitrous Oxide

The visible reflection spectrum of molecular oxygen is shown in Figure 13 for five abundance levels. Candidate detection bands are the 0.7- and 0.76-μm bands, which are spectrally relatively narrow, but which are roughly linear in their response to oxygen abundance. At 0.76 μm, there is no known interfering feature. (There is no corresponding thermal emission infrared wavelength feature of oxygen.)

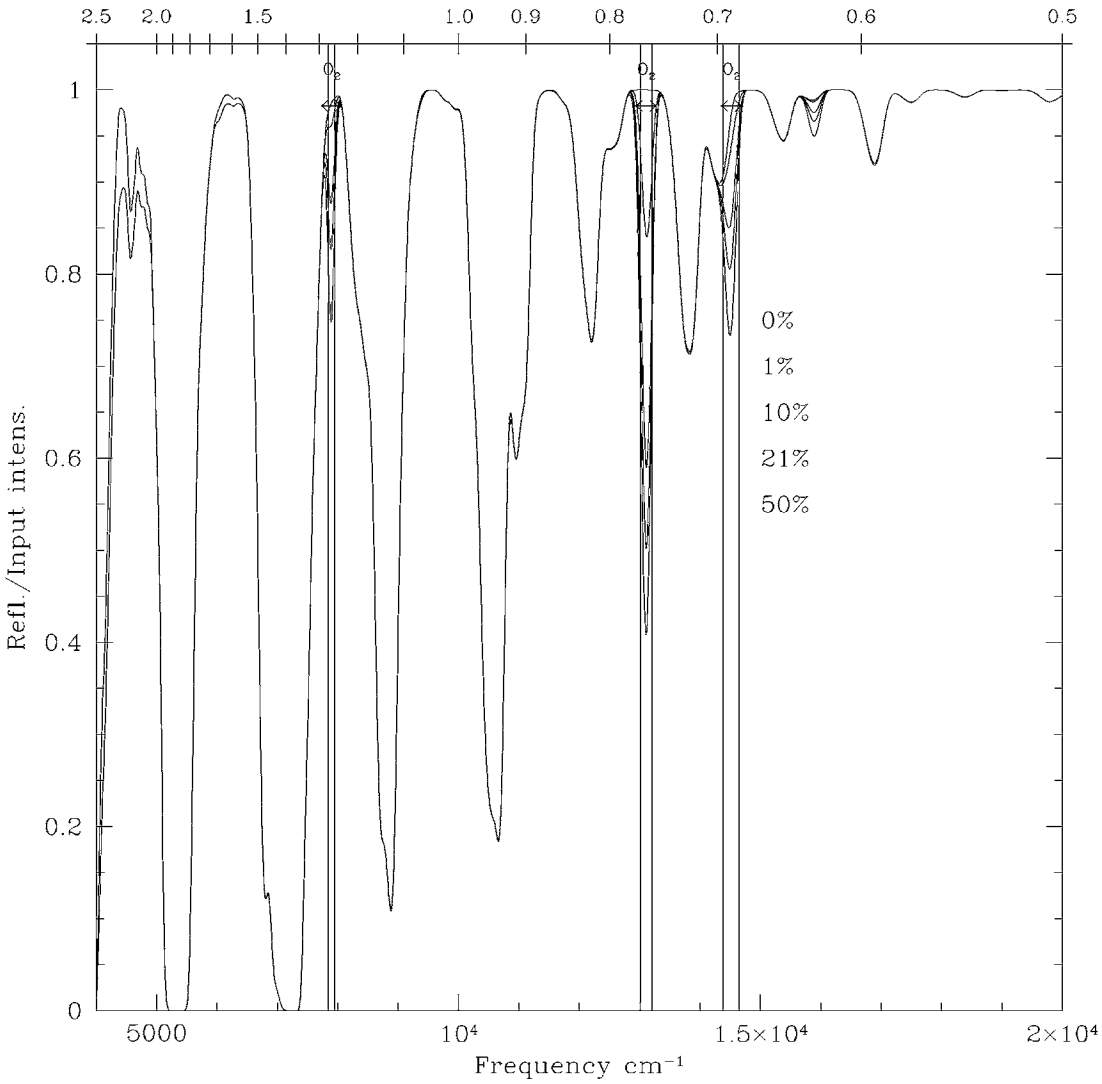

Figure 13. Visible Reflection Spectrum of Molecular Oxygen

Figure 14 shows (a) Illumination of a planet at different orbital phases around its parent star, and the corresponding light curve seen by a distant observer viewing the planet along its orbital plane. (b) Illumination of an orbiting planet with a circumplanetary ring inclined to the orbital plane, and the corresponding light curve as a function of orbital phase (solid line) contrasted to the light curve for the planet alone (dashed line).

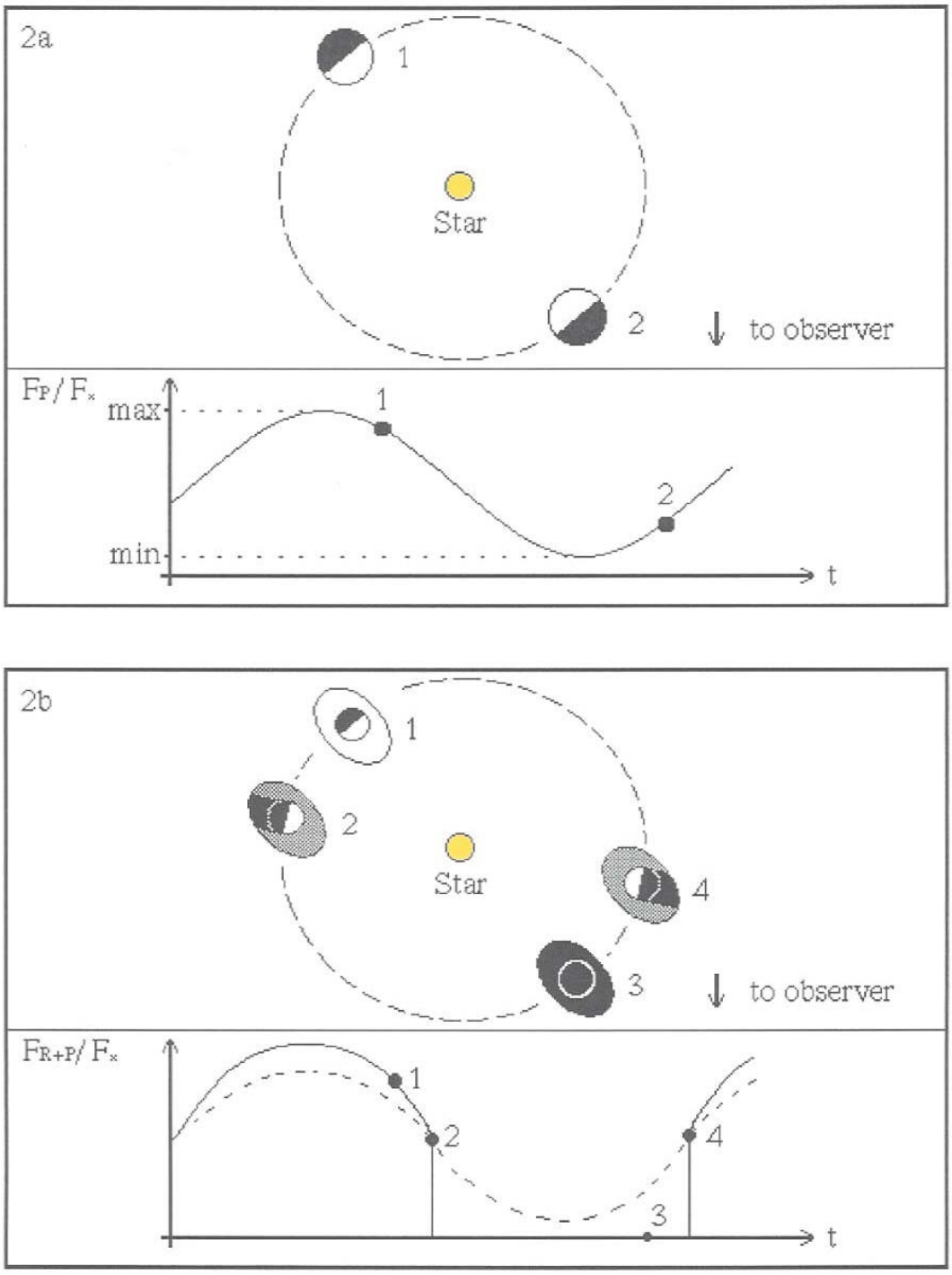

Figure 14. Illumination of a Planet (a) Shown at Different Orbital Phase and (b) Shown With a Circumplanetary Ring

VI. Appendices

A. Algorithms for Spectral Detection

Overview

The analytic formulae that are used to determine whether devices meet a particular detection criterion are among the simplest and least controversial issues in designing devices for observation of planetary systems and extra-solar terrestrial planets. The formulae are part of a suite of items that are needed to determine whether a device does or does not satisfy a detection criterion. The components of this suite can be separated as "optical" and "mechanical".

The optical suite has the following components:

a) An optical diagram

b) A scenario of how the observations are made that shows how the terrestrial planet is found, its presence confirmed and its spectrum obtained.

c) An analysis of performance losses through expected characteristics of the optical components and system associated with a set of specifications that achieve certain levels of performance.

d) A formula that calculates the required time that is necessary for the observations.

e) A list of optical items where the performance exceeds the current state of the art, and an analysis of how the performance degrades if these requirements are stepped back at various levels down to the current state of the art.

It is clearly impossible to produce a formula that steps through all the items of discussion for this suite, but it is possible to indicate how the formulae, item (d), can be used in a consistent way.

The two types of features

There are two types of features to be observed for an extra-solar planet:

a) The continuum, and

b) Spectral lines or bands

The processes for obtaining these observations are quite different for the various devices currently being considered for TPF. There seem to be two main kinds of device:

a) A device that produces an optical image directly, and which then moves the entrance slit of a spectrograph onto the planet and adjacent sky to observe spectral features.

b) A device that produces the spectrum of certain areas of sky and modulates the areas of sky so as to permit angular resolution within each area, and then each spectral element is separately demodulated to determine the image that accompanies the spectral measurement.

The Continuum

Presence of an object. The first reason for observing the continuum is to see the object. The previous suggestion for observing the continuum was to observe with $\lambda/\Delta\lambda = 3$, and with signal/noise (S/N) = 5. Certainly from experience, S/N = 5 is a much safer limit than S/N = 3. On the other hand, if we insist on greater than S/N = 5, we are likely to miss real planets. (In all of this we are discussing gaussian statistics.) If instead we are in a very low photon detection rate (which seems unlikely to us), then Poisson statistics will apply, and instead of looking for a particular S/N ratio, we will be looking for an equivalent frequency of spurious detection.

In order to be sure of the presence of an object, one collects together the total signal and compares it with the total noise. For an object that is expected to be strong in some parts of the spectrum and weak in other parts, it is recommended that wavelengths at which the planet is weak and/or the background strong be removed from consideration. It is permissible to exclude, for example, short wavelengths where the thermal emission of a planet far from its star is expected to be weak, or long wavelengths where a planet closer to its star is also expected to be weak. However, the optimum way to detect planets is to add together as much of the stronger parts of the continuum as possible.

If we assume that TPF is likely to observe about an octave of spectrum, then an observation with R = 3 and S/N = 5 per spectral element is actually a detection with S/N = 5sqrt(3) or S/N = 8.6. That is an unnecessarily high threshold. Consider that we are looking at a picture with ≈100 pixels where a terrestrial planet might be found. There will be perhaps 200 pictures, and if we want the probability of a spurious signal on any one of them to be less than 1%, then we are asking for the Error Function to be less than 1/2,000,000. Now with Gaussian statistics, the probability of exceeding a 5σ error is ≈1/2 10^6, therefore, unless there are non-Gaussian errors, a 5σ detection criterion should have a wide enough safety margin. This statistic is appropriate for any detection method of which we are aware.

It is not recommended that the 5σ result be obtained in a single observation. There are too many risks of intermittent or spatially located noise producing a spurious result. Instead it is recommended that the result be the sum of, say, three independent measurements. Thus if the planet is imaged directly, it would be possible to make three separate measures at three different roll angles. Other types of devices might use, for example, three spectral bands. The mean signal/noise for any individual observation among 3 should be >2.9. Three such results combined produce a 5σ answer.

If the planet contributes n photoelectrons per unit time, and the background, detector noise equivalent input, star leak, local and system zodiacal flux contribute m photoelectrons per unit time, then, for a 5σ result

$$nt = 5\mathrm{sqrt}(nt+mt)$$

$$t = 25(n+m)/n^2 \qquad (3$$

The second feature of a continuum is its shape. In the thermal part of the spectrum, the shape gives a measure of the temperature of the object being examined. It may be a ground temperature or a cloud temperature, but if it has the shape of a thermal spectrum, that itself, together with the temperature is important knowledge.

For a reflection spectrum the shape is a measure of both the illuminating source and the reflectivity versus wavelength of the surface. For this reason it is recommended that a device looking for the reflection spectrum use the same spectrograph to observe the star as well as the planet. It is then possible to obtain the reflectivity of the surface without needing to know about the wavelength sensitivity of the spectrograph. This could be particularly useful for looking for

the Chappius bands of O_3 in the visible range against a surface of clouds. For a long wavelength spectrograph it is probably useful also to obtain sensitivity versus wavelength by observing the star continuum. In both cases a calibrated near-neutral density filter would need to be used.

Continuum shape and spectral features. Very broad and shallow spectral features may be seen in a reflection spectrum against a neutral reflecting background such as clouds. For a pigmented reflector or for an emission spectrum shallow features can only be discovered by measuring the continuum shape well on either side, and by demonstrating that the observed intervening continuum deviates from the extrapolation through the continuum points. Low spectral resolution is unable to distinguish a broad shallow feature from a narrow deep feature or a set of such features. Any team proposing to observe a shallow feature should explicitly explain its proposed technique and validate it with calculations.

Another difficult determination is of a broad spectral feature whose center is off the edge of the observed portion of the spectrum. In these cases it is necessary to define the continuum sufficiently well that it can be projected through the region of absorption. Again there is a potential for systematic error from unpredicted spectral curvature. If there is some way to *predict* the continuum shape within an error range, that is the best way of extrapolating a continuum level.

The determination of a continuum spectral shape is a process that can be carried out at various levels of sophistication. A minimal shape determination has two points, while a more sophisticated one can use many more. A two-point observation determines the best straight line fit, and the errors in the amplitude and/or slope of the line can be used for simple analyses.

In order to be free of the influence of spectral bands, a number of points should be observed. The fewer the points selected as "continuum," the higher should be the signal-to-noise ratio of the measure. If only one or two points are selected, the sensitivity calculation should include the uncertainty in the continuum. A greater number of points is helpful in measuring spectral lines, because the continuum measure needs to have less noise than the measures within the line to minimize the required observing time.

Spectral Line or Band Observations
It helps to discuss two limiting cases:

1) Observations are limited by photon noise. This is independent of whether it is photon noise of planet, star-leak or zodiacal dust.
2) Observations are limited by detector noise.

Then we can discuss the likely case:

3) Both photon and detector noise are significant

There are p photons/second/unit of bandwidth in the observed continuum. In all cases it is assumed that there is a spectrograph whose efficiency does not change with the spectral resolution. We will perform a simple analysis for a rectangular feature of width W. The planet

radiation is a fraction f of the total radiation detected, and the average fractional depth of the spectral band in the planet signal alone is d. d = fractional depth of feature (shallow ≈0 = min, deep ≈1 = max). Note that d is the average absorption, which is about half the peak absorption for typical bands. For those who wish to use the equivalent width of a feature E, Wd = E.

Case 1. The continuum is assumed defined infinitely precisely at the wavelength of the feature because there could be many more continuum points than line points. If the continuum is not better defined than the line, then the observing time must be doubled.

Case (1) is photon noise limited. Let us do an analysis for the line absorption falling on a single pixel of spectral width KW, where K>1, that is, the line is totally unresolved. Over the width, in time t, pKWt/f +/- sqrt(pKWt/f) events are expected. Here p is the number of incident continuum photons per unit time and we have defined the measured quantity as corresponding to the sum of object, sky etc. It is assumed that an additional measurement or modulation technique separates the object from its background. Then the line deficiency Wdpt must be determined to 5 sigma accuracy.

Thus

$$Wdpt= 5\ sqrt(pKWt/f)$$

so that

$$t=25K/fpWd^2 \qquad (4$$

It can be seen that the observing time in this case will get longer as the spectral resolution is lowered, K is made larger, and f is smaller.

Case 2. For observations limited by detector noise, the sensitivity depends on NOT resolving the line, so that the entire region of depressed continuum falls on the least number of pixels. *If it falls on more pixels, the observing time increases proportionally with the number of pixels used.*

Let the detector rms noise on a single pixel be equivalent to the detection of N photons per second. (Internally it corresponds to the equivalent of N^2 being detected), and M pixels are used. Then the noise will be +/- Nsqrt(Mt) and the signal to be detected is Wdpt.

So for a 5 sigma detection,

$$5Nsqrt(Mt) = Wdpt$$

$$t = 25M(N/Wdp)^2 \qquad (5$$

The result above defines the detector noise limited case.

In contrast we now consider a high resolution spectrograph. For this we are looking for a total photon loss of WDpt in a continuum giving pt/f photons in the line width. So without weighting of the deeper absorption regions (impossible for a discovery process), the observing time comes from

$$Wdpt = 5\ sqrt(Wpt/f)$$

$$t = 25/fpWd^2$$

This result is independent of how high the resolution is, and is the same as the result for a wide band if K=1. However, as the resolution is raised, one moves away from being photon noise limited towards detector noise limited. The preferred condition is the photon noise limited condition, and optimum design pushes the spectral resolution to the point where detector noise is noticeable but smaller than photon noise.

Case 3. Both photon noise and detector noise are significant. If for high resolution, we put K=1 (to minimize detector noise), and we include both photon and detector noise,

$$Wdpt = 5sqrt(pW/f + MN^2)$$

Again, M is the number of pixels used for the line detection. Typically it will be ≈10, but will depend on the device design.

$$t = 25(pW/f + MN^2)/W^2d^2p^2 \qquad (6$$

It is necessary to incorporate all efficiency calculations before applying the above analysis.

Line detection is of course not the same as line interpretation! Presumably some higher value than signal/noise value of 5 will be needed for line interpretation. The true signal/noise value of the line observation is (continuum signal/noise*d). Previous TPF analyses have asked for a signal/noise value of 20 for a value of d ≈0.25 which corresponds fairly closely to signal/noise =5. This is a bare minimum for interpretation.

Line Identification

If there is only a depression over a single resolution element, the identification by peak wavelength is quite uncertain. One can say that the feature lies within the wavelength range of the pixel. In fact if the feature has appreciable strength one can place the average wavelength of the feature towards the pixel center. When there are multiple pixels showing absorption, it is possible to determine the likely peak wavelength of the line by graphing the absorption against wavelength, and determining the centroid by visual inspection. Three pixels are the minimum number required to do this, and it is unlikely that the band center will be determined to better than a precision ≈0.2 pixel.

If there are more pixels, it is possible to help identify the feature by spectral shape. Does absorption degrade to longer or shorter wavelengths? Does the absorption feature break into two parts? Identification of spectral features by a single absorption is particularly difficult. Do the circumstances suggested by an identification of a few bands seem to suggest an environment that is plausible for the object? Identification of broad features is only possible if the spectral shape is used as part of the identification process.

Comparison of Planet Detection and Line Detection

If we wish to use a spectral detection instrument to both detect a continuum and to detect a spectral feature in the presence of a dominating background noise, as is the case for most instruments now envisioned, then the relative time to do this is scaled from the continuum detection time.

$$t_f = t_{c1}(S/N)^2(R/R_0)(cf)[\eta_c/\eta_l]/d^2$$

where

t_f	=	time to measure the total strength of the spectral feature
t_{c1}	=	time to detect the continuum alone, at signal/noise = unity
S/N	=	desired signal to noise ratio of feature measurement to object measurement
cf	=	continuum factor, on the order of 1 to 2 and equal to 1 for a perfectly known continuum.
d	=	average line depth in the passband (0=shallow, 1=deep)
R	=	resolution needed for feature (wavelength / bandpass)
R_0	=	resolution assumed to detect the continuum in time t_{c1}
$[\eta_c/\eta_l]$	=	is a factor that is only used if a different device is used to observe the continuum and the line. It is the ratio of the transmissive efficiencies of the continuum detection device and the line detection device.

The use of this formula assumes that the spectral feature is just resolved (and uses the same number of pixels to detect the line as are used to detect the planet image). A virtue of this formulation is that the purely spectroscopic factors (cf and R and d), and the desired signal quality (S/N), are cleanly separated from the planet-detection factors (t_l and R_0). Note that this formulation uses the width and depth independently, not just their product (the equivalent width).

Summary and Conclusions

This note discusses the detection criteria for planets and spectral lines for TPF. The detection criterion for a planet is given by Equation 3, and for spectral lines by Equation 6. The formulae are part of a suite of optical and mechanical items required to assess whether a certain TPF concept does indeed satisfy TPF design constraints.

It is recommended that, for consistency, Equation 4 be used for all devices. Note that all signal values are used as they would be after all inefficiencies have been allowed for. The photon noise is calculated for the sum of planet and star leaked signal and local and external system zodiacal emission.

If parts of the detection process have reduced efficiency below that calculated for photon noise, that inefficiency must be allowed for also. The efficiency of the process of separating the planet signal from the background must also be accommodated.

Because of the complexity in the scenario for use of the TPF variants, it is not possible at this time to define all the processes that will enter into the use. It is recommended that the scenarios of operation for the different TPF variants be scrutinized to understand the appropriate data reduction process for each, and that this be compared with the formula that is used to calculate the signal-to-noise ratio.

B. Effects of Planetary Rings, Dust Wakes, Moons and Binary Planets: Interpretation of Visible and IR Spectra of Terrestrial Planets.

Rings:

For a full ring with external radius $\mathbf{R_R}$, an inclination $\mathbf{i_R}$ and the same albedo as the planet, the reflected flux (comparable to the planet flux) is $\mathbf{F_r = A\pi\phi(t)\cos(i_R)R_R^{\ 2}}$ (neglecting the part of the ring hidden by the planet) when the observer sees the illuminated face of the rings. But, because the rings are planar, along half of the orbit the observer sees them from the back with respect to the illuminating stellar source, that is, their dark side and the ring reflected flux is then $\mathbf{F_r = 0}$. The total ring plus planet flux is, then, at least for the half portion of the orbit where the observer sees the illuminated side of the rings: $\mathbf{F_r = A\pi\phi(t)(\cos(i_R)R_R^{\ 2} - R_{pl}^{\ 2})}$. The flux variation along the orbital revolution is therefore no longer sinusoidal, and thus indicates a nonspherical shape (Fig 14).

Depending on the relative observer-ring orientation-orbital position configuration, the situation may be complex; depending on the relative planet/ring size, the dark side of the rings may even completely hide the planet for some orbital positions (see the accompanying Figure 14).

"Binary planet"

Because both components of a pair of planets are spherical, one just adds the fluxes of each individual component and there is no particular orbital phase effect similar to the ring case. No "counter-measure" is foreseen before the visual separation of the two components in eventual, very high, angular-resolution imaging. In about 10% of cases, there will nevertheless be an exception: a mutual eclipse of the two components of the binary planet will happen every half-orbital period. For instance, for a radius ratio 0.5 for the two components and similar albedo, the relative planet flux drop during the two mutual eclipses is 25% (the detection of that effect requires a sufficient photometric precision and continuous time coverage in planet imaging).

Infrared Issues

An Earth-like planet seen from a distance of several parsec, with instruments that cannot separately resolve a Moon-like satellite circling the planet, or a planet-trailing dust cloud similar to the one observed near Earth, would contribute an infrared flux of the order of 20 to 30% that of the planet. However, the albedo of the barren satellite and the dust would likely be extremely low, the surface of the Moon having been darkened by the solar wind. The low albedo would raise the temperature of the satellite by several tens of degrees relative to the planet's black sphere temperature, so that some care would need to be taken in interpreting the data.

The planet trailing cloud data considered are discussed by Beichman et al. *[Beichman et al., 1999]* Chapter 5. There it is suggested that the solar system dust ring will be at a similar temperature to Earth and emit a flux about 10% of Earth. However the strength of the dust glow will depend on the amount of dust, and planetary systems probably exist with a range of dust wake glows from a few percent of Earth to many tens or hundreds of times that of the Earth.

The effect of such a dust glow would be to cause us to interpret the observations as indicating a larger planet, a 10% glow being interpreted as a planet 5% larger. Also, spectral

features will be affected. Radiation pressure would remove all small grains rapidly and these are ones that would otherwise show spectral emission features. The Poynting-Robertson effect will remove the small particles, so the remaining grains of ≈40 μm diameter are expected to show a smooth featureless continuum. The overall effect would be to wash out spectral features due to the planet. Therefore, for the wake discussed above, spectral features would become 10% weaker.

The effect of a satellite lacking an atmosphere can be calculated. Because it rotates slowly, our Moon has a hot side facing the Sun and a cold side away from the Sun. If the axis of the planetary system is substantially inclined away from the line of sight, we will see a strong seasonal effect. The planet's contribution will be very significant when we have the planet on the opposite side of its system and very weak when the planet is on our side of its system. If instead the pole is pointed at us, then the flux from the Moon would get integrated with that of the Earth.

A few calculations have been made of the effect of a Moon-like satellite on the infrared spectrum of the joint Earth-Moon system as compared with an Earth alone. We assume that the sunlit side of the Moon has an average temperature of 375 K, and the dark side 200 K. For Earth we assume 285 K surface temperature. Two sizes of Moon have been modeled: ½ Earth diameter and ¼ Earth diameter. The calculations are for an orbit that is almost edge-on, so that in one season we see the full Moon, and six months later we observe the new Moon. There are two effects, namely the modification of the apparent temperature as measured by the 9/12-μm ratio and the washing out of the O_3 band at 9.7 μm.

The results are as follows:

1) At the season where we see the dark side of the Moon, the Earth spectrum is essentially unchanged.

2) At the season where we see the bright side of the Moon, the Earth temperature appears higher—297 K for a moon of ¼ Earth diameter, 315 K for a moon of ½ Earth diameter.

3) The O_3 band appears at reduced strength, 84% of full strength for ¼ Earth diameter, 57% for ½ Earth diameter.

The flux change with season would also be noticeable, particularly for the larger moon, but we might misinterpret that as the effect of a planet with a thin atmosphere. We would need to use the seasonal change in band strength to eliminate that possibility. Therefore, for a planetary system with its ecliptic pole near normal to the line of sight, and for a moon appreciably larger than our own, the "Earth" study would be virtually unaffected. However, an appreciably larger moon would be evident, and we could even interpret some of its properties. If, instead, the ecliptic pole is along the line of sight, and the moon is large, we will see a somewhat washed-out spectrum. Even here it is possible that a study of the core of the 15-μm CO_2 band will allow us to separate the planet emission from that of wakes and satellites.

If, instead of color temperature, we consider a temperature defined by total radiation, then for an angle α between the star's line of sight and the star to planet radial direction ($\alpha = 0$ when we see the moon's dark side, $\alpha = \pi$ when we see the sunlit side), the apparent Earth temperature is

$$T=285(1+(R_M/R_E)^2[(200/285)^4(1+\cos\alpha)/2+(375/285)^4(1-\cos\alpha)/2])^{(1/4)}$$

There thus is a temperature modulation along the orbit. Such a modulation would also result from two other configurations: eccentric orbit and inclination of the exo-earth rotation axis with respect to the normal of its orbital plane (planet seasons). According to the simulations by J. Laskar on the chaotic behavior of the planet rotation axis, such a situation is not exceptional.

An orbit with eccentricity **e** gives a relative temperature orbital modulation with a peak to peak amplitude $\Delta T/T=e$. But then the star to planet distance varies by an amount $2ea(\sin\alpha)=0.1a$ in case of an eccentricity of 0.1 and $\alpha=$ 30 deg. between the two positions of extreme temperature variation.

Barycenter and Photocenter

In the case of an exo-earth similar to our Earth, the summer to winter temperature difference for a given hemisphere is ≈20 K, similar to the 315 to 285=30 K difference due to an 0.5 earth diameter moon. But then, contrary to the lunar case, the thermal photo-center is identical to the planet center and has a Kepler motion. For a Moon-Earth system, the thermal photo-center makes a monthly orbit around its barycenter with a peak to peak amplitude

$$(\Delta a_M)_{phot} =2a_M(M_M T^4_E R^2_E -M_E T^4_M R^2_M)/[(M_M +M_E)(T^4_M R^2_M +T^4_E R^2_E)]$$

For an 0.5 earth diameter moon, the photo-center amplitude is $3.3\ 10^{-3}$ UA = 0.3 mas at 10 pc. (It would be measurable with long exposures, in particular if there is a background reference star in the field of view). If this monthly photo-centric modulation is too small to be measured, the ambiguity with a pure seasonal effect would remain unremoved. If it is measurable, one can infer the planet mass range *[Schneider et al., in preparation]*

VII. References

Beichman, C.A., N.J. Woolf, and C.A. Lindensmith, The Terrestrial Planet Finder: A NASA Origins Program to Search for Habitable Planets, JPL Publication, 99-3 1999.

Caroff, L.I., and D.J. Des Marais, Eds., Pale Blue Dot 2 Workshop: Habitable and Inhabited Worlds Beyond Our Solar System, NASA Publication, NASCP 2000-209595, pp. 320, 2000.

Goode, P.R., J. Qui, V. Yurchyshyn, J. Hickey, M.-C. Chu, E. Kolbe, C.T. Brown, and S.E. Koonin, Earthshine Observations of the Earth's Reflectance, in, *Geophysical Research Letters*, pp. 1671–1674, 2001.

Kasting, J.F., Runaway and moist greenhouse atmospheres and the evolution of Earth and Venus, *Icarus*, *74*, 472–494, 1988.

Kasting, J.F., Habitable zones around low mass stars and the search for extraterrestial life, in, *Origins of Life*, pp. 291–307, 1997.

Kasting, J.F., and L.L. Brown, Setting the stage: the early atmosphere as a source of biogenic compounds, in, *The Molecular Origins of Life: Assembling the Pieces of the Puzzle*, edited by A. Brack, pp. 35–56, Cambridge Univ. Press, New York, 1998.

Kasting, J.F., J.B. Pollack, and T.P. Ackerman, Response of Earth's atmosphere to increases in solar flux and implications for loss of water from Venus, *Icarus*, *57*, 335–355, 1984.

Kasting, J.F., D.P. Whitmire, and R.T. Reynolds, Habitable zones around main sequence stars, *Icarus*, *101*, 108, 1993.

Kieffer, H.H., T.Z. Martin, S.R. Peterfreund, B.M. Jakosky, E.D. Miner, and F.D. Palluconi, Thermal and albedo mapping of Mars during the Viking primary mission, *J. Geophys. Res.*, *82*, 4249–4291, 1977.

Leger, A., M. Ollivier, K. Altwegg, and N.J. Woolf, Is the presence of H_20 and O_3 in an exoplanet a reliable signature of a biological activity, *Astronomy and Astrophysics*, *341* (1), 304–311, 1999.

Lunine, J.I., The occurrence of Jovian planets and the habitability of planetary systems, in *Proceedings National Academy of Science*, *98*, pp. 809–814, National Academy of Science, 2001.

Reynolds, R.T., C.P. McKay, and J.F. Kasting, Europa, tidally heated oceans, and habitable zones around giant planets, *Advances in Space Research,* 7, 125, 1987.

Sagan, C., W.R. Thompson, R. Carlson, D. Gurnett, and C. Hord, A search for life on Earth from the Galileo spacecraft, *Nature*, *365*, 715–721, 1993.

Schneider, J., Extrasolar planets: detection methods, first discoveries and future perspectives, pp. 621, C.R. Acad. Sci. Paris, Paris, 1999.

Schneider, J., A. Labeyrie, and F. Coliolo, A universal biosignature: generalized chlorophylls, in, *AGU 2000 Spring Meeting. EOS Supp*, 9 May, S30, 2000.

Tomasko, M.G., The thermal balance of Venus in light of the Pioneer Venus Mission, *J. Geophys, Res.*, *85*, 8187–8189, 1980.

Traub, W.A., and K.W. Jucks, *A possible aeronomy of planetary systems beyond the solar system, in*, AGU, 2001 (submitted).

Traub, W.A., and M.T. Stier, Theoretical atmospheric transmission in the mid- and far-infrared at four altitudes, *Applied Optics*, *15*, 364–377, 1976.

Turner, E., E. Ford, and S. Seager, Theoretical Light Curves of Extrasolar Terrestrial Planets, *BAAS*, *32* (4), 42.06, 2000.

Contacts

David J. Des Marais
Mail Stop 239-4
Ames Research Center
Moffett Field, CA 94035-1000
ddesmarais@mail.arc.nasa.gov

Douglas Lin
University of California
Kerr 461
Santa Cruz, CA 95064
lin@ucolick.org

Martin Harwit
Cornell University
511 H Street, SW
Washington, D. C. 20024-2725
harwit@bellatlantic.net

Sara Seager
Institute for Advanced Study
School of Natural Sciences
Princeton, NJ 08540
seager@ias.edu

Kenneth Jucks
Center for Astrophysics
Smithsonian Institution
Cambridge, MA 02138
kjucks@cfa.harvard.edu

Jean Schneider
Observatoire de Paris
92195 Meudon, France
jean.schneider@obspm.fr

James F. Kasting
Department of Geosciences
The Pennsylvania State University
State College, PA 16802
kasting@essc.psu.edu

Wesley Traub
Center for Astrophysics
Smithsonian Institution
Cambridge, MA 02138
wtraub@cfa.harvard.edu

Jonathan I. Lunine
University of Arizona
Department of Planetary Sciences
Tucson, AZ 85721
jlunine@lpl.arizona.edu

Neville Woolf
University of Arizona
Steward Observatory
Tucson, AZ 85721
nwoolf@as.arizona.edu

Additional Reading

Berkener L. V., and L.C. Marshall, The History of Oxygenic Concentration in the Earth's Atmosphere, *Faraday Society Disc*ussions, 37, pp. 122-141, 1964.

Lederberg J., Signs of Life Criterion-System of Exobiology, Nature, 207, pp. 9-13, 1965.

Lovelock J. E., A Physical Basis for Life Detection Experiments, *Natur*e, 207, pp. 568-570, 1965.

Lovelock J. E., Thermodynamics and the Recognition of Alien Biospheres*, Royal Society of London Proceedings B*, 189, pp., 167-181, 1975.

Owen T., 1980: The Search for Early Forms of Life in Other Planetary Systems—Future Possibilities Afforded by Spectroscopic Techniques, in "Strategies for the Search of Life in the Universe", Papagiannis ed., *Reidel, Dordrecht*, pp. 177-185,