Sources and physical processes responsible for OH/H$_2$O in the lunar soil as revealed by the Moon Mineralogy Mapper (M$^3$)

T. B. McCord,$^1$ L. A. Taylor,$^2$ J.-P. Combe,$^1$ G. Kramer,$^1$ C. M. Pieters,$^3$ J. M. Sunshine,$^4$ and R. N. Clark$^5$

Received 11 August 2010; revised 4 January 2011; accepted 1 February 2011; published 14 April 2011.

Analysis of two absorption features near 3 $\mu$m in the lunar reflectance spectrum, observed by the orbiting M$^3$ spectrometer and interpreted as being due to OH and H$_2$O, is presented, and the results are used to discuss the processes producing these molecules. This analysis focuses on the dependence of the absorptions on lunar physical properties, including composition, illumination, latitude, and temperature. Solar wind proton-induced hydroxylation is proposed as the creation process, and its products could be a source for other reported types of hydrogen-rich material and water. The irregular and damaged fine-grained lunar soil seems especially adapted for trapping solar wind protons and forming OH owing to abundant dangling oxygen bonds. The M$^3$ data reveal that the strengths of the two absorptions are correlated and widespread, and both are correlated with lunar composition but in different ways. Feldspatic material seems richer in OH. These results seem to rule out water from the lunar interior and cometory infall as major sources. There appear to be correlations of apparent band strengths with time of day and lighting conditions. However, thermal emission from the Moon reduces the apparent strengths of the M$^3$ absorptions, and its removal is not yet completely successful. Further, many of the lunar physical properties are themselves intercorrelated, and so separating these dependencies on the absorptions is difficult, due to the incomplete M$^3$ data set. This process should also operate on other airless silicate surfaces, such as Mercury and Vesta, which will be visited by the Dawn spacecraft in mid-2011.


1. Background

The Chandrayaan-1 spacecraft [Goswami and Annadurai, 2008] was inserted into a polar orbit around the Moon on 8 November 2008 and continued to make measurements for more than nine months, until 16 August 2009. One instrument onboard was the Moon Mineralogy Mapper (M$^3$), a spectrometer, which measured the flux reflected and emitted from the lunar surface in the spectral region from about 0.406 to 2.991 $\mu$m (R. O. Green et al., The Moon Mineralogy Mapper (M$^3$) imaging spectrometer for lunar science: Instrument, calibration, and on-orbit measurement performance, submitted to Journal of Geophysical Research, 2011). One important discovery reported from the M$^3$ measurements is evidence of widespread OH and H$_2$O in the lunar surface layer [Pieters et al., 2009]. In this present article we discuss the possible explanations for hydroxyl/water in the surface and explore the M$^3$ observational evidence for the source. We focus here on solar wind-induced hydroxylation as at least one likely explanation for the global imprint and latitude dependence [McCord et al., 2009, 2010a, 2010b].

Pieters et al. [2009] initially described the M$^3$ observational evidence to that date, which consisted of absorption features in the Moon reflectance spectrum in the 2.8 to 3 $\mu$m spectral region, to the long-wavelength limit of the M$^3$ coverage. This absorption was interpreted as indicating OH and H$_2$O, and similar spectral features are commonly seen in spectra of outer Solar System objects. It was noted that this feature and the associated OH were not expected in the case of the Moon, as the surface is hot during the day and was determined to be very dry based upon study of lunar samples returned by the U.S. Apollo and USSR Lunar programs.

The absorption was reported from the initial M$^3$ measurements to be widely distributed and appeared strongest at cooler high latitudes and at several fresh feldspathic craters. However, thermal emission contaminates the reflected solar
radiation in this spectral region for much of the Moon observed and complete removal of this effect is difficult or impossible with the M sub 3 data. Some of these effects are discussed below. It also was only weakly correlated with the Lunar Prospector neutron spectrometer hydrogen abundances [Feldman et al., 1998, 2000, 2001; Lawrence et al., 2006; Pieters et al., 2009], suggesting probably a different source mechanism and possibly a surficial process that would not have been detected by the Lunar Prospector’s observations. Further, it was suggested that this hydroxyl/water production process could feed to polar cold traps and thus might form a source of useful water.

[5] This finding was so unexpected that there was a strong but unsuccessful attempt by the M sub 3 team to disprove it. Two sources of independent observations fortunately were available. The Cassini spacecraft carried the Visual and Infrared Mapping Spectrometer (VIMS) past the Moon on its way to Saturn in 1999. These earlier data were investigated and also show the OH/H sub 2O absorptions [Clark, 2009]. The VIMS covered a spectral range further into the infrared, to about 5 \( \mu m \) (unsaturated range to 3.3 \( \mu m \) in this case), so that the full extent of the absorptions became more apparent and better defined, with OH/H sub 2O amounts suggested to be in the range 10–1000 ppm. Furthermore, EPOXI, NASA’s extended mission for the Deep Impact spacecraft, had been making numerous close approaches to the Earth-Moon system and was observing the Moon as a calibration source, particularly with its 1.05–4.8 \( \mu m \) spectrometer (useful to 4.5 \( \mu m \) in this case). Prompted by the M sub 3 discovery, EPOXI observed the Moon on its June 2009 encounter. Again, the 3 \( \mu m \) lunar spectral signature was observed and confirmed [Sunshine et al., 2009]. These EPOXI observations were of the highest signal-to-noise ratio of the three observations and showed detailed spectral structure within the general absorption. They reported H sub 2O concentrations of <0.5 wt % with more toward the north polar region. They suggested that hydration varied with temperature, rather than total dose of solar radiation, and they reported a lack of correlation with regolith composition. Interestingly, they also reported diurnal changes in hydration that were greater for mare basalts (~70%) than for highlands (~50%) and occurred entirely between local morning and evening, an effect that they suggested required a ready daytime source of water-group ions consistent with a solar wind origin. However, Clark [2009] offered an alternative explanation: that the change in absorption strength of the water and hydroxyl is a multiple reflection effect due to topography. At lower solar incidence angles, mountains and crater walls facing the sun reflect light onto more horizontal surfaces causing a “reflectance squared” effect that increases apparent absorption depths.

[6] These two additional observations strongly confirmed and extended the M sub 3 data, so that there is little question that the spectral feature exists. The interpretation of OH mentioned above seems the obvious one and has not to date been seriously questioned. However, the presence of H sub 2O is a little less certain. The 3 \( \mu m \) spectral feature in general is due to O–H stretch of hydroxyl groups and can exist without H sub 2O. Spectral features diagnostic of H sub 2O also exist in the 3 \( \mu m \) region (e.g., near 3.1 \( \mu m \) due to H–O–H bend mode [cf. Dyar et al., 2010, Table 2]) but they are combinations and overtones with strengths 30–100 times weaker, and can be identified only at very high signal-to-noise ratios, a challenge for the existing observations. Less well determined or understood are the dependences of the spectral features on lunar properties and the nature of the physical process(es) creating the molecular species.

[7] Prior to the M sub 3 discovery, the Lunar Prospector’s neutron spectrometer observations were interpreted to show that excess hydrogen is present in the lunar polar regions within about 0.5 m of the surface [Feldman et al., 1998, 2000, 2001]. This result was suggested as evidence that water might be cold trapped and stored for long time periods in permanently shadowed regions to several meters in depth and gardened to nearer the surface by impacts. Such a scenario has been suggested by several scientists over the history of lunar exploration [e.g., Urey, 1952; Watson et al., 1961]. Feldman et al. [1998, 2000, 2001], however, could not ascertain that the hydrogen was in the form of water or even hydroxyl, and therefore, the nature and source were not evident. The Lunar Reconnaissance Orbiter mission also carried a neutron detector, the Lunar Explorer Neutron Detector (LEND) [Mitrofanov et al., 2010a, 2010b]. Maps from LEND of the south polar region reveal an elongated depression of the epithermal neutron rates that indicates a maximum concentration of hydrogen [McClanahan et al., 2010]. This depression was already reported by the Lunar Prospector Neutron Spectrometer, and it was validated by a direct comparison with LEND results that show a good correlation that increases toward the pole where the confidence is higher (higher density of measurements). The Lunar Crater Observation and Sensing Satellite (LCROSS) experiment impacted the Centaur upper stage rocket into a shadowed part of a lunar polar crater in October 2009; Evidence of water ice and other volatile materials [Colaprete et al., 2010] were reported. If confirmed, this water ice may be at least part of the source for the excess hydrogen reported by Lunar Prospector.

[8] In this article, we investigate the behavior of the M sub 3 3 \( \mu m \) region absorptions and discuss what these indicate about the sources for these OH and HOH observations, elaborating on previous discussions by McCord et al. [2009, 2010a, 2010b].

2. Potential Sources of Water on/in the Moon

[9] Water, OH and other volatiles, both on the surface and within the Moon, can come from three sources: (1) indigenous from the original material forming the Moon; (2) inflow of meteorite and cometary material containing water or hydroxyl; and (3) from processes associated with the solar wind irradiating the surface material that can actually produce such compounds. (See Lucey [2009] and Lucey et al. [2006] for background and reviews.)

2.1. Indigenous Volatiles in Material Forming the Moon

[10] It was generally thought in the lunar science community that the lunar surface was very dry and without water [Lunar Science Preliminary Examination Team, 1970]. Conditions on the Moon’s surface are hostile to the presence of water, with temperatures up to +125°C and a vacuum of 10 sub -12 to 10 sub -15 atmospheres. However, the search for evidence for water and other volatiles in lunar rocks, minerals, and soils has been pursued since the first return of Apollo
samples [e.g., Anders, 1970] (see also articles in The Moon Issue of Science (167, 1970) and Proceedings of the Apollo 11 Lunar Science Conference, Geochimica Cosmochimica Acta (1–3, 1970)]. Water and numerous other volatiles released from lunar soils were detected during early stepwise heating experiments [e.g., Gibson and Johnson, 1971]. However, this water was attributed to terrestrial contamination and to reactions with solar wind-implanted protons in the soils [Epstein and Taylor, 1971]. The “rust” [FeO(OH, Cl)] discovered in >40 Apollo 16 breccias has been attributed to the contamination by terrestrial moisture. This caused the oxyhydration of meteoritic lawrencite [FeCl₂], present in the lunar soil and rocks, to produce akaganeite [FeO(OH,Cl)], upon exposure of the samples to the terrestrial atmosphere [Taylor et al., 1973, 1974]. However, this hypothesis has not been universally accepted, with a cometary origin for the source of the water having been proposed [El Goresy et al., 1973].

[11] The early adopted paradigm for the formation of ubiquitous nanophase metallic Fe (np-Fe⁰) observed in agglutinitic glass in soil grains is that solar wind implants hydrogen in the lunar soil, which in turn reduces the FeO component in the soil melt to metallic iron during local micrometeorite impact melting of the soil [Housley et al., 1973]. If this reaction occurs, it should be accompanied by the production of considerable water, some of which should remain in the quenched melt (agglutinitic glass). However, an early infrared spectroscopic examination of some of this agglutinitic glass did not detect any presence of this hypothesized water [Taylor et al., 1995]. Because of the paucity of any substantiated lunar water, historically, the Moon has been believed to be “bone dry.”

[12] Recently, Saal et al. [2008] detected 20–45 ppm water in Apollo 15 green and Apollo 17 orange volcanic glasses using highly sensitive secondary ion mass spectrometry (SIMS). These samples are quench glasses believed to be the most primitive materials from the lunar mantle in our sample collection, and thus represent our best estimate of the composition of the lunar interior. Based upon diffusion gradients of hydrogen, they modeled that the melt, before degassing, may have originally contained as much as 260 to 750 ppm water. This important discovery has stimulated considerable research, and there are now several new searches for and “discoveries” of lunar water in Apollo samples underway using modern laboratory equipment.

[13] Reasoning that the best place to search for water in lunar minerals is in the very last portions of the crystallizing melt, the apatite in the mesostasis of basalts has become the target of study. Further evidence for water in the lunar interior, albeit in far greater quantities than reported by Saal et al. [2008] were presented recently at the 2010 Lunar and Planetary Science Conference [Greenwood et al., 2010; McCubbin et al., 2010; Liu et al., 2010]. They reported analyses of up to 7000 ppm of water in this late stage apatite – Ca₃(PO₄)₃(O,F,Cl,OH). In addition, Greenwood et al. [2010] reported D/H values for this water that are indicative of cometary input, probably incorporated into the ancient Moon, shortly after the Giant Impact forming the Earth-Moon system. This logically leads to the conclusion that the residue of the lunar magma ocean (termed KREEP, for its unusual composition) should also contain abundant water-bearing minerals. However, this indigenous water probably is not present in sufficient abundances at the optical surface to produce the spectral signal detected by the M³. The materials that might contain indigenous lunar water compose only a small fraction of the lunar surface sampled by M³ and other remote sensing investigations. The enhanced quantities of KREEP in the Procellarum KREEP Terrain [Jolliff et al., 2000] suggest that it would be the obvious region of the Moon that would contain the highest abundances of lunar indigenous water.

2.2. Volatiles From Infall of Meteorite and Cometary Material

[14] There was evidence of meteoritic material in the lunar soil from the first analyses [e.g., Keays et al., 1970]. It is well known also that certain chondrite meteorites, namely carbonaceous chondrites, contain matrices with abundant carbon- and water-bearing minerals and compounds. Much of this is the result of the alteration of primitive olivine and pyroxene and consists of various phyllosilicates such as clays and serpentine [Mg₃Si₂O₅(OH)₄]. It is to be expected that of the innumerable meteorites and comets that have struck both the Earth and Moon over time, such chondrites could effectively disaggregate and partially disintegrate upon hitting these bodies, thereby spreading these hydroxyl- and water-bearing minerals over the countryside. But there is little or no evidence of this material on the Moon. Because these phyllosilicates are so hydrogen-rich, their spectral signatures would be conspicuous. However, the impact velocity of these materials on the Moon is such that small impactors are totally vaporized, destroying the hydrogen-rich molecules. The peak in the meteoroid mass flux hitting the Moon is ~200 µm and the velocities can range from >11 km/s to as high as ~70 km/sec (Comet Halley would be ~76 km/s). The stability of these phases to subsequent space weathering also is a determining factor to their survivability.

2.3. Formation of OH From Interaction of Surface Material With the Solar Wind

[15] A possible source of the hydrogen that went into the production of OH and H₂O on the lunar surface is the solar wind. The solar wind consists mostly of a proton (H⁺) and electron (e⁻) plasma that strikes the Moon’s surface facing the Sun, with some disruption by the Earth’s magnetic field in parts of the Moon’s orbit. These protons, traveling at 300–800 m/s [Smith, 1967] can penetrate into the surface soil grains, which are ~45 wt % oxygen, at least 5–10 molecular levels (5–10 nm). Some of these protons could bond with the oxygen to form OH and even H₂O. This perhaps simplistic concept has been discussed nearly since the discovery of the solar wind. For example, Zeller et al. [1966] proposed the process and presented experimental evidence for it, and Starukhina [2001] discussed it and modeled the process.

[16] Each of the above three potential sources of OH and H₂O remains possible, given current knowledge. Indigenous water-rich material from the lunar interior seems insufficient in amount and surface area extent and concentration to completely explain the M³ observations. In addition, it is unlikely that infalling OH and H₂O-rich material could produce the widespread spectral feature with the intricate patterns associated with lunar geology observed by M³ [Pieters et al., 2009], in addition, little convincing evidence...
of hydrated/hydroxylated infall material has been found in the lunar samples. Thus, in this article we consider the process of proton-induced hydroxyl formation in the lunar surface material caused by the solar wind as the most likely explanation for most of the M\(^2\) observations of OH and H\(_2\)O in the optical surface layer of the Moon. Also, this is a process that has implications for other airless Solar System object surfaces. As part of our analysis, we explore the behavior of the M\(^2\) absorption in the 3 µm region and their correlations with lunar physical properties and search for physical evidence for the source(s).

3. The Solar Wind

[17] A flow of particles and energy away from the Sun into the Solar System was postulated in the midnineteenth century. Observations of aurora in the upper Earth’s atmosphere and terrestrial magnetic storms and their correlations with solar flares and of comet’s tails always pointing away from the Sun led early scientists to postulate and model what later became known as the solar wind [Parker, 1958] (see Schröder [2008] for a detailed discussion of the long and complicated history of solar wind investigations). The wind was first observed directly by the Soviet satellite Luna 1 in 1959 and verified by measurements from Luna 2, Luna 3 and Venera 1 (see Harvey [2007] for discussion of the history.). Solar wind was first observed by a U.S. spacecraft, Mariner 2, in 1962 [Neugebauer and Snyder, 1962]. It is composed mostly of protons and electrons, with about 4% helium and smaller amounts of heavier element ions. The flux at the Earth is about 4 × 10\(^8\) cm\(^−2\) s\(^−1\), with average energy \(\sim 0.5\) keV per amu and a flux energy half width between about 300 and 1500 keV as measured recently by the Sub-keV Atom Reflecting Analyzer (SARA) experiment [Wieser et al., 2009] on the Chandrayaan-1 mission. For this energy range, protons (H\(^+\)) are reported to have a penetration depth in the surface grains to 5–10 nm [Starukhina and Shkuratov, 2000; Johnson, 1990]. It has until recently been widely accepted that the solar wind plasma is almost completely absorbed by the Moon’s illuminated surface [e.g., Crider and Vondrak, 2002; Schmitt et al., 2000; Feldman et al., 2000; Behrisch and Wittmaack, 1991]. Recently, Wieser et al. [2009] using SARA measured reflected solar wind protons as neutral hydrogen atoms coming off the Moon. They reported that up to 20% of the impinging solar wind protons are reflected from the lunar surface back to space as neutral hydrogen atoms.

4. Nature of the Lunar Soil

[18] The solar wind impacts the lunar soil surface. The resulting interaction between the solar wind particles and the regolith depends largely on the nature of the lunar soil surface at the molecular level. Thus, we first discuss the nature of the lunar surface soil.

[19] The lunar regolith blankets the surface with a layer of crushed rock, minerals, and glass [McKay et al., 1991]. On the Moon, regolith formation contrasts drastically with that on Earth in that it is a direct function of “space weathering.” Space weathering is a combination of all the physical and chemical factors that occur on airless bodies, like the Moon. The agents of space weathering include a wide range of types and sizes of impactors, such as meteorites, micrometeorites (<1 mm), solar wind particles, solar-ultraviolet and galactic/cosmic rays. The micrometeorites process the lunar surface by crushing and grinding from rocks into soil (<1 cm regolith) to dust (<20 µm), and to nano-sized grains. This is accompanied by impact melting of the silicate soil and vaporization and redeposition as an oxygen-depleted silica-rich glass. On the other hand, solar wind particle-induced spallation of soil grain surfaces produce an amorphous glass layer enriched in oxygen (i.e., superstoichiometric) [Keller and McKay, 1997].

[20] It is a combination of these agents of weathering that are responsible for the formation and evolution of the lunar regolith, its soil and dust. These processes occurring in a vacuum also affect and enhance the chemical reactivity of the soil grains and produce some chemical changes. In addition, this impact-induced comminution, with its instantaneous crushing and fracturing, produces myriads of unsatisfied bonds, most of which are those of oxygen.

4.1. Agglutination

[21] Although the craters formed by meteorites are obvious on the lunar landscape, it is the extreme flux of micrometeorites that perform the major weathering of the regolith/soil, causing comminution, agglutination, and vaporization (Figure 1). These small meteoritic grains (<1 mm), although with minute mass, have tremendous impact velocities (as high as \(\sim 70\) km/s), and hence high kinetic energies. The crushing and shattering of grains is the most obvious effect of micrometeorite impacts, and this comminution is the major force decreasing grain size. However, it is this same micrometeorite impact force that has sufficient energy to partially to completely melt the soil, with the melt quenching to glass. And this melt/glass weds the soil particles into aggregates called agglutinates. This glass is brittle and makes the agglutinates easily comminuted [Taylor et al., 2001]. The major melting of the lunar soil is of the finest fraction of the soil, which already consists of 60–80% glass and chips off the rims of grains. Taylor et al. [2001] demonstrated that the amount of agglutinitic glass increases with a decrease in grain size, suggesting maturation. The micrometeorites also effectively garden the regolith and soil, with the occasional larger impacts churning and overturning the grains (impact gardening). This stochastic impact processing creates a chaotic layering to the lunar soil such that a lower layer may be younger than an upper one, and layers do not necessarily correlate from one place to another only meters away [McKay et al., 1991].

4.2. Vaporization and Deposition

[22] Soil grains melted by impact may vaporize as a result of the extreme temperatures (>2000°C) reached by some melts, giving off disassociated molecules, ions and atoms into the vapor [Keller and McKay, 1997]. Depending upon the temperature, these molecules and ions will dissociate further into charged ions of elements, as depicted in Figure 2. This process places a portion of the silicate melt into vapor, which permeates and condenses on the rims of all soil grains, largely as an amorphous glass rinds from 50 to 100 nm thick. This process also provides disassociated O\(^2−\), which can bond with available H\(^+\). The texture and major element compositions of these thin, amorphous rinds on the surface of lunar
soil grains is a testament to the process of vaporization and subsequent deposition of these silica-rich patinas with their myriads of np-Fe\(^{0}\).

[23] In general, oxygen is not vaporized or sputtered as easily as many other elements. The vapor-deposited glass as rinds with np-Fe\(^{0}\) has a deficiency of oxygen, whereas the sputter-produced glass has excess oxygen as a residue from cation loss [Keller and McKay, 1997].

### 4.3. Solar wind Particle Sputtering and Implantation

[24] Sputtering is a process whereby solar wind particles bombard the exposed surfaces of lunar soil grains, producing an amorphous layer and effecting atoms, ions or molecules to be ejected from a lattice site in the target material [Stern, 1999; Behrisch and Eckstein, 2007]. Such ejected particles from the source material can be redeposited as a thin film on a surface in the ballistic path of the ejected atoms. The sputtered particle can be charged, but is most often neutral [Crider and Vondrak, 2002]. The sputtering energy is inherent in the solar wind velocity and particle mass. Protons (H\(^{+}\)), which comprises ~95% of solar wind particles, have an incident energy range of ~300–1500 eV and average energy of ~500 eV under normal solar wind velocities of 300–800 km/s [Smith, 1967]. Heavier ions (especially He at ~4% of solar wind particles) and other heavy species, with greater incident energies, also play a significant role in the sputtering process [Behrisch, 1983; Brunetto and Strazzulla, 2005; Strazzulla et al., 2005].

[25] Solar wind produced amorphous rims and their crystalline hosts have similar yet different major element chemistry, although in different proportions [Keller and McKay, 1997]. Specifically, the silicate rims are depleted in some cations (especially Mg), which is thought to result from preferential sputtering of cations with the lowest crystalline binding energy [Bradley, 1994; Bradley et al., 1996]. It is common for such amorphous, silicate matrices to contain some chemical constituents not indigenous to the host grain. These are particularly present on ilmenite and other opaque, nonsilicate phases. The most common of these inclusions is nanometer-sized grains of Fe metal, termed nanophase iron (np-Fe\(^{0}\)).

### 4.4. Solar UV Radiation

[26] Photoelectrons are emitted from the lunar dayside surface by solar UV radiation, leaving the surface charged several volts positive and leaving many dangling positive ions. The released electrons move to the unlighted side of the Moon and into the solar wind plasma wake, formed by the Moon absorbing most of the solar wind [Manka, 1972; Freeman and Ibrahim, 1975; Benson, 1977; Farrell et al., 2008]. Some of these effects were observed by the Electron Reflectometer onboard Lunar Prospector [Halekas et al., 2002] and have stimulated study of the electrical charging of objects on the Moon, such as astronauts [Farrell et al., 2008]. The photoejection of electrons is yet another effect that enhances the chemical activity of the lunar soils.

### 4.5. Cosmic and Galactic Ray Spallation and Implantation

[27] The bombardment flux of cosmic/galactic rays on the Moon are very energetic (on the order of ~1 GeV), so energetic that they penetrate several meters into the regolith as they fragment, producing less energetic products, but potentially dangerous neutrons. They can damage crystalline grains by forming damage tracks. At a high density, these accumulated tracks can cause amorphization.

### 4.6. Temperature Variations

[28] Diurnal temperatures on the Moon (125°C to −150°C) near the equator are extreme by Earth standards, but with the absence of fluid water, these extreme temperature variations do not have an appreciable effect on mechanical weathering of lunar rocks/soil. However, the extremely cold temperatures (30–40 K) in permanently shadowed craters at the poles [Paige et al., 2010] should affect the brittleness and change the overall crushability of soil grains, and therefore the melting and development of agglutinates, and potential chemical reactions. Low-temperature comminution of the lunar soil in permanently shadowed craters may have developed a finer grain size than present in “normal” soil.

---

**Figure 1.** Schematic depiction of the several processes that operate to form and space weather the lunar soil.

**Figure 2.** An illustration of the some of the chemistry the solar wind ions and micrometeoroids enable in the lunar surface grains.
One result might be higher concentrations of solar wind particles, which are surface area correlated with smaller grains that have higher surface areas.

5. Solar Proton Implantation and OH Formation

5.1. The Concept

Solar wind impacting protons (H\(^+\)) strike the lunar soil grains with enough energy to penetrate to depths of 5–10 nm [Starukhina and Shkuratov, 2000; Johnson, 1990]. If there were no chemical processes, the protons would quickly diffuse out of the surface and would be lost to space, an effect that is mineral specific [e.g., Nichols et al., 1989]. However, the soil grains are extremely space-weathered and thus have very irregular surfaces from comminution, and the fractured minerals have random surface cuts through the crystal structures. This weathering in a vacuum results in unsatisfied bonds at the grain surfaces and at many defects and dislocations within the outer portions of the grain. These surface and internal sites are locations where protons could be trapped and bonded with the abundant oxygen, forming individual OH and HOH groups to depths of 5–10 nm. In a distorted crystal lattice the bond lengths and strengths will be irregular, and can be more readily broken and will present distorted and broadened spectral signatures.

5.2. Surface Sites

The surface sites for OH and H\(_2\)O formation are essential for the solar wind protons to react. The lunar soil grain surfaces are irregular and have huge surface areas. This is due to micrometeorite impact and partial melting (Figure 3) and to sputtering by particle radiation (Figure 4), all in deep vacuum (10\(^{-12}\) torr). The lunar soil grains have been shown to be very chemically reactive, even under terrestrial ambient conditions. For example, Wallace et al. [2009] studied the chemical activity of freshly ground (in the laboratory under ambient conditions) fine lunar soil grains (Apollo mare and highland soils) and compared these with the activity of freshly ground lunar soil simulant (JSC-1A-vf) and quartz samples. They found that the simple process of grinding with mortar and pestle under laboratory ambient conditions increased the surface area of the already fine grains by factors of 2 (simulant) to 5 (lunar soil) [Wallace et al., 2009, Table 2]). In these ambient conditions, they reported that the grains tended to clump, so their surface area increase must be only a fraction of what it would have been in vacuum without much contact with other materials. They determined the chemical reactivity of the ground samples by measuring the sample’s capability of producing hydroxyl radicals. It was determined that it was the presence of np-Fe\(^0\) in the soil glasses that greatly increased the chemical reactivity and also correlated this with the degree of maturity of the soil: mature soils = more np-Fe\(^0\) = greater reactivity (Figure 5).

Additional recent work relevant to the reactivity of lunar soil was reported in a study of the chemical reactivity (formation of peroxy and hydroxyl groups) for ground pure minerals compared to ground quartz [Hurowitz et al. 2007]. They found basaltic silicates much more reactive than quartz, implying a compositional dependence for hydroxyl formation. Further, they reported that heating of the samples (>400°C) increased the reactivity. This suggests a possible temperature-dependent enhancement of the lunar process, at least associated with the microimpact process, if not also the diurnal cycle.

These studies addressed above suggest that the finely ground lunar soil grains under vacuum conditions could be...
bonding on a grain surface in the incident solar
E00G05 is likely at least Bustin and Gibson [1992] reported heating [1973, 1974] also described the concept of tetrahedron, the basic trapped in defects –
Gibson and Moore E00G05 and after activation (grinding). The emission spectrum that is compared for the dust before terephthalate (10 mM; at a wavelength of 324 nm). This is was measured by the fluorescence of synthetic 2 aqueous solution. The OH contents of the dust (3.8 mg/mL) was monitored the ability of the dust to produce hydroxyl radicals in aqueous solution. The OH contents of the dust (3.8 mg/mL) was measured by the fluorescence of synthetic 2-hydroxyterephthalate (10 mM; at a wavelength of 324 nm). This is the emission spectrum that is compared for the dust before and after activation (grinding).

very receptive to chemical reaction, whereby freshly broken surfaces cannot be pacified by an effective atmosphere and its contained constituents (water, oxygen, etc.). These dangling bonds that have not yet been annealed might readily form OH.

[33] Of course, the actions of space weathering and impact gardening are capable of both activating and neutralizing grain surfaces. Micrometeorite and larger impacts can activate mineral surfaces through mechanical forces. Crushing and shattering of minerals leaves their surfaces with partially unsatisfied chemical bonds. Generally, an impact event provides a chemically reactive surface through the physical breaking of bonds and gardening of lunar soil, which buries more mature, coated grains and exposes fresh uncoated as well as fractured grains. These surfaces tend to have a net negative charge resulting from the coordination of ligands around their cation host. Micrometeoroid impacts and solar irradiation will continuously break bonds and chemically activate surfaces, but these processes are also responsible for the creation and deposition of neutral atomic species and glass coatings, which works to deactivate surfaces. The rate of production of active surfaces and their neutralization will reach some equilibrium, with some broken bonds always remaining.

[34] Continual irradiation by solar wind protons could pacify broken bonds through the formation of OH and H$_2$O. However, these are likely weak bonds (hydrogen or Van Der Waal forces) and could be easily broken at a critical temperature in the vacuum of space. There could be a balance between H$^+$ bonding on a grain surface in the incident solar wind, and the revolatilization of H$^+$ at some temperature, depending on the bond strengths. There might be a temperature dependence that is observable.

[35] As a side comment, it is interesting to consider the observations by the Lunar Prospector and LRO-LEND neutron detectors, which found elevated hydrogen abundances in many permanently shadowed craters. Since the reactions of the surface with solar wind H$^+$ is likely at least partially a grain surface effect, it is possible that the finer grain sizes would be more reactive in the permanently shadowed polar regions due to the low temperatures and probably more brittle nature of the soil grains. This would enhance the available grain surface area and result in greater hydrogen retention in these colder regions. This might explain many of the LEND observations without resorting to deposits of ice.

[36] Evidence exists of enhanced solar wind species trapped in the lunar soil from the earliest analyses of the Apollo lunar samples. For example, Gibson and Moore [1972] and Bustin and Gibson [1992] reported heating lunar soil samples in vacuum in the laboratory and measured released trapped solar wind materials, including 10–60 ppm by weight H. They found strong correlation of the amounts of trapped gases with soil maturity and with decreasing grain size. This result agrees with other measurements of enrichment of noble gases and carbon from the solar wind.

5.3. Interior Sites

[37] We now focus on interior sites within individual lunar soil grains for OH and H$_2$O formation. Note that this is different from mobility of H within regolith, i.e., grain-to-grain diffusion (see Grieeves et al. [2010] and Hibbitts et al. [2010] for discussions of the mobility of H). The constant mechanical reworking of the surface grains must induce microcracks and stress fractures resulting in crystal lattice dislocations. Further, processes associated with partial melting would distort the crystal structure. A high density of defects in the crystal structure would result. These would be fertile sites for bonding and trapping of ions passing through the material, especially highly mobile species such as H$^+$. Note that these sites are not within a regular crystal lattice structure, but rather are at breaks or defects in the crystal structure. Since oxygen is so abundant and is preferentially located on the outside of the Si-O$_2$ tetrahedron, the basic building block of silicates, abundant sites for OH formation are expected at these defects at least until annealed and pacified by some other mechanism. For H$^+$ trapped in defects and dislocations within the soil grains, the bond formation should be more stable than for surface sites, and with a continued flux of protons, the reactions might more readily be driven to the formation of H$_2$O.

5.4. Not a New Concept! The History

[38] The concept of OH and H$_2$O formation by proton bombardment on the lunar surface is not new. Shortly after direct measurement of solar wind, Zeller et al. [1966] proposed proton-induced hydroxyl formation on the Moon and presented experimental evidence of the process of OH formation, although for much higher energies (MeV protons) than the solar wind and for simulated lunar material (chemically similar glass). He reported a conversion efficiency (OH produced per incoming proton) of 5 to 100%. Housley et al. [1973, 1974] also described the concept of...
solar wind induced chemistry in the lunar surface soil grains. The solar wind implants supposedly created an autoreduction of the impact melt. The focus was on explaining the submicron Fe metal in the soil grains but the reduction reaction would also produce OH and H$_2$O as well as the reduced Fe. They also pointed out that the formation of vesicles in the glassy lunar soil grains is direct evidence of the grains having been saturated with solar wind gas (H$^+$) and its volatile products (e.g., OH, H$_2$O); and this supposed np-Fe$^0$ production was associated with agglutinitic glass formation.

More recently, Starukhina [2001] proposed that implanted protons in oxygen-bearing regolith grains form isolated disordered OH groups that belong neither to water molecules nor to hydroxylate compounds. She presented calculations supporting this hypothesis. She also pointed out that the H$^+$ trapped on the surfaces of grains is more weakly held than those trapped internally in defects, and also suggested that the highly irregular OH sites in regoliths would produce a range of energies for the H-O bonds. This would result in a broadened spectral signature, such as in the 3 μm region where M$^2$, VIMS and EPOXI detected the features for the Moon and for telescopic asteroid band detections.

It remains to be seen whether such extreme broadening can be achieved otherwise, so as to eliminate the need for H$_2$O to explain the observed spectral signature. Nevertheless, it brings into question whether the recent spectral observations prove the existence of H$_2$O specifically.

Recently, Zent et al. [2010] reported producing hydroxyl in Apollo lunar soil samples by proton irradiation at solar wind energies. They heated and dehydrated lunar soil samples and exposed them to proton fluxes of solar wind energies. IR spectroscopy was used to demonstrate OH formation in the irradiated sample compared with unirradiated and unheated samples. They also irradiated with deuterium (D$^+$) and produced the expected OD absorption, demonstrating that the effect was due to irradiation. Further experimentation into this OH production mechanism is under way.

On the other hand, some experimentalists have reported their inability to produce hydroxyl by proton irradiation [Burke et al., 2010; G. Strazzulla, personal communication, 2010]. These studies mostly used samples with polished crystalline surfaces, which do not have the similar reactivity to the highly irregular lunar soil grains and they used particle energies far higher than that of solar wind protons e.g., ~100 keV versus ~1 keV, resulting in much greater penetration depths. In general, these reactions may be highly dependent on composition, crystallinity of the material and other physical conditions. Thus, extensive laboratory work will be needed to constrain to what extent solar wind proton bombardment of lunar soil is a viable mechanism for producing OH.

5.5. Nanophase Iron Formation

The formation of nanophase metallic Fe due to chemical reactions occurring in/on lunar soil grains due to space weathering is well known. These grains are produced through the reduction of Fe$^{3+}$, present naturally in the lunar soil and particularly in mare basalts, to a metallic state (see section 4). These grains have a strong effect on the optical spectra character of the Moon’s surface (see Lucey et al. [2006] for a review), and have a significant effect on spectroscopic measurements of surface maturity [e.g., Lucey et al., 2000; Noble et al., 2007; J. W. Nettles et al., A new lunar optical maturity index based on Moon Mineralogy Mapper data, submitted to Journal of Geophysical Research, 2010].

Housley et al. [1973, 1974] proposed that the solar wind protons present in lunar soil formed the agent for reduction of the FeO in the soil melt to np-Fe$^0$, with the release of water, and they demonstrated that the np-Fe$^0$ grains are associated with the agglutinates and regolith breccias. Largely because of its simplicity, this became the paradigm for the production of most of the np-Fe$^0$ in lunar soil. However, the proof of the impact-induced vaporization and deposition by Keller and McKay [1997] as minute (50–100 nm) layers of oxygen-deficient glass with myriads of np-Fe$^0$ has changed our considerations of these unique minute metallic Fe grains. The fact that the np-Fe$^0$ in the rims of virtually all soil grains in a mature lunar soil occurs as equally dispersed grains in the glass demonstrates that homogeneous nucleation of the Fe occurred, not something that most spallation or autoreduction by protons would produce.

The extent to which implanted solar wind protons (H$^+$) facilitate iron reduction remains a point of debate. Some workers [e.g., Sasaki et al., 2001] contend that vapor deposition alone is sufficient on the Moon and that the presence of solar wind H$^+$ as a reducing agent is not necessary. Sputtering by solar wind ions has also been shown to create np-Fe$^0$ [Hapke 1973; Hapke, 1974, 1977; Cassidy and Hapke, 1975]. More recent work supports this [e.g., Loeffler et al., 2009]. The lack of water, a byproduct of the reduction of FeO to np-Fe$^0$ in lunar soil samples argues against solar wind implanted H$^+$ as a significant contributor to space weathering [Taylor et al., 1995; Hapke, 2001]. However, space weathering is still not well understood, and new observations suggest other processes could be responsible for the apparent lack of water in returned samples [Taylor et al., 1995; McCord et al., 2009, 2010a, 2010b].

Different space weathering processes tend to create a particular range in np-Fe$^0$ sizes. Grain sizes of np-Fe$^0$ deposited on soil grain surfaces by vapor deposition/sputtering range between 1 and 15 nm with a median of 3 nm [Keller and Clemett, 2001], where the median represents an equilibrium state between creation/deposition and revaporization. The np-Fe$^0$ grain sizes created by solar wind protons might have a wide range. Incident H$^+$ has the energy to directly reduce Fe$^{3+}$ at mineral surfaces, in which case the resultant np-Fe$^0$ grains would be small, on the same order of sizes created by sputtering. The np-Fe$^0$ grains in agglutinates are larger and have a wide range of sizes, up to several hundred nanometers [Housley et al., 1973; Keller and Clemett, 2001]. This range is caused largely by the heating of the glass and np-Fe$^0$ grains causing random diffusion of the metallic Fe, which subsequently ripens into coarser grain sizes.

Oxygen loss and creation of reduced phases is a function of the binding energy of the atoms/ions at the grain surfaces [Kelly, 1987; Dukes et al., 1999]. The Fe-O bond is weaker than the Si-O bond [Reed, 1971; Lieb, 1985]; therefore, oxygen is more easily sputtered from a grain surface if bound to Fe than if bound to Si. This explains why
reduced Fe and, to a lesser extent Si, are observed in lunar samples. Dukes et al. [1999] showed that, for olivine under irradiation, all the near-surface Fe-O bonds are destroyed leaving an aggregation of metallic iron.

The formation of np-Fe⁰ indicates that ferrous iron is being reduced in the lunar soil, perhaps due mainly to chemical reduction by solar wind protons. An obvious question is what happens to the OH or H₂O that should form as a byproduct of at least some of the other formation processes. Since no detectable OH was observed in the initial study of the lunar samples, this byproduct was suggested to simply be lost to diffusion and heating or was simply ignored. Another explanation is that some oxygen atoms are sputter, depleting the material of oxygen and leaving only reduced Fe. With the M⁰ discovery, some OH and H₂O are now known to exist in the lunar surface soil.

6. OH Behavior on the Moon From M³ Observations

The nature of the processes that might form OH and H₂O on/in the soil grains of the Moon suggests that there might be some dependence of the amount of these molecular groups on various lunar physical properties. These properties include temperature, soil composition, time of lunar day, latitude and lighting. Therefore, we have conducted analyses of the M³ data, by far the most complete spectral image data set obtained for the Moon to date, to search for any relationship [McCord et al., 2009, 2010a, 2010b]. We present here examples of the results so far.

There are difficulties in conducting this analysis. Most of the physical parameters available to explore are themselves intercorrelated for the Moon. Lighting and temperature are correlated and both are related to latitude. Composition is also correlated with latitude in that most maria are in the lunar frontside, near-equatorial regions, and lighting and temperature are correlated with surface topography. Additionally, the M³ investigation was not able to complete its measurement plan, due to loss of communications with the Chandrayaan-1 spacecraft. Thus, its data set is incomplete. Nevertheless, a meaningful analysis was possible. Additional values of this analysis are that it should help define the measurement plan for subsequent observational efforts and that it should provide some guidance for future analyses of M³ data.

6.1. Calibration of M³ Data

We worked with the entire M³ data set projected into a global mosaic for monitoring trends at global scale and/or with individual image strips for closer examination of surface details. We used calibration version “K” (as defined by the M³ team) for all the analyses discussed here. We converted radiance data into reflectance factor I/F (where I is the surface irradiance and pi*F is the solar flux) with a basic photometric correction that assumes the lunar surface is an isotropic scatterer (Lambertian surface): I/F = Radiance × pi/(solar radiance × cos(incidence)). A thermal correction has also been applied [Clark et al., 2011], unless specified otherwise. This is important because thermal emission may have a significant contribution in M³ spectra beyond 2.178 μm, especially when the surface of the Moon is hot (above 250 K), and that may affect the estimate of the absorption band depth. Calibration residual errors that remain in the data as systematic, high-frequency features in the spectra are erased using a single corrective spectrum

Figure 6. Examples of the narrower 2.8 μm (dotted line) and broader 3 μm (dashed line) absorption features in the overall M³ Moon reflectance spectrum. The third spectrum (solid line) is for a mature highlands soil from a region where these absorptions are weak.

Figure 7. Correlation plot of the strength of the 3 μm versus the 2.8 μm absorption features. Band strength measurement technique and the M³ data set used are described in the text. The two absorption strengths are highly correlated. There are several data clouds evident, suggesting several different phenomena affecting the band strengths. Most of the pixels are concentrated in the lower 3 μm band strength region, including many with zero strength and lying on the horizontal axis.
spectraClark\textsuperscript{m} Clark spectra, and it is extrapolated as a straight\textsuperscript{m} wavelength channels in order to increase\textsuperscript{m} data: a broad band centered\textsuperscript{m} average I/F at 2.876\textsuperscript{m} \textsuperscript{Sunshine et al.}\textsuperscript{spectra using two\textsuperscript{m} \textsuperscript{-\textsuperscript{m}\textsuperscript{E00G05}} The strength of (left) the 3\textsuperscript{m} I/F\textsuperscript{52} the band\textsuperscript{5} shorter wavelengths, since the long\textsuperscript{wavelength shoulder of the band. Instead, a\textsuperscript{2}\textsuperscript{m} \textsuperscript{-\textsuperscript{m}\textsuperscript{Pieters et al.}} Sinusoid dependence on latitude.\textsuperscript{45}\textsuperscript{o}, with many pixels at or near zero\textsuperscript{m} strength and plotted on the horizontal axis. The 2.8 \textsuperscript{m} feature is present nearly everywhere with a distinct sinusoid dependence on latitude.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8.png}
\caption{The strength of (left) the 3 \textmu m versus (right) the 2.8 \textmu m absorption features versus lunar latitude. Clearly, the behavior is different between the two absorptions. The 3 \textmu m feature is concentrated at high latitudes and is mostly absent (as we measure it) below about 45\textdegree, with many pixels at or near zero strength and plotted on the horizontal axis. The 2.8 \textmu m feature is present nearly everywhere with a distinct sinusoid dependence on latitude.}
\end{figure}

(\textsuperscript{polisher}) multiplied to all M\textsuperscript{3} spectra [\textsuperscript{Clark et al., 2011}]. Further, the global mosaic data set has had an estimate of any thermal emission radiation contribution removed.

\subsection{6.2. Calculation of the Relative Spectral Feature Depths at 2.8 and 3.0 \textmu m}

\textsuperscript{[51]} There are at least two types of absorptions in the 3 \textmu m region apparent in the M\textsuperscript{3} data: a broad band centered beyond 3 \textmu m, but with the band edge extending shortward to about 2.6 \textmu m, and a narrower absorption centered near 2.8 \textmu m (Figure 6) [\textsuperscript{Pieters et al., 2009; Clark, 2009; Clark et al., 2011; Sunshine et al., 2009}]. Sometimes both features appear superimposed and sometimes only the 2.8 \mu m feature is present.

\textsuperscript{[52]} Estimates of the spectral absorption feature’s relative depth at 2.8 \mu m uses a straight-line continuum removal between 2.696 and 2.976 \mu m. The depth at 2.8 \mu m = 1 – I/F at 2.8 \mu m/continuum at 2.8 \mu m. The absorption band depth at 3 \mu m is estimated using an extrapolated continuum from shorter wavelengths, since the long-wavelength wing of the band [\textsuperscript{Clark et al., 2011}] is not available in the data, and therefore a straight-line continuum cannot be anchored on the long-wavelength shoulder of the band. Instead, a straight-line continuum is fit on top of M\textsuperscript{3} spectra using two flexible anchor points in the range 1.419–1.519 \mu m and 2.477–2.756 \mu m, respectively. This continuum is designed to not go below M\textsuperscript{3} spectra, and it is extrapolated as a straight line toward the longer wavelengths. Then, the average relative band depth is calculated in the range 2.876–2.956 \mu m that uses three M\textsuperscript{3} wavelength channels in order to increase the signal-to-noise ratio: Average depth at 2.876–2.956 \mu m = 1 – average I/F at 2.876–2.956 \mu m/average continuum at 2.876–2.956 \mu m.

\subsection{6.3. Correlation With Surface Illumination}

\textsuperscript{[53]} The occurrence and strengths of the two spectral features are correlated with each other, as shown in Figure 7. There seems to be at least three different clusters in this scatter diagram. This suggests that there are several factors affecting the band strengths and that these affect each of the bands differently. Most of the pixels are located in the lower core of the cluster (low or zero 3 \mu m band strength), which maps mostly to mare regions. The pixels with the strongest band strengths map mostly to higher latitudes, where the 3 \mu m band already was reported to be strongest [\textsuperscript{Pieters et al., 2009; Clark, 2009; Sunshine et al., 2009}]. A plot of band strengths with lunar latitude (Figure 8) clearly shows this effect. The behavior shown in Figure 8 for each of the two absorptions (Figure 6) that are analyzed in Figure 7 is different. The 3 \mu m feature strength is concentrated at high latitudes and is mostly absent (as we measured it) below about 45\textdegree, with many pixels in this region at or near zero strength and plotted on the horizontal axis (not visible in Figure 8). The 2.8 \mu m feature is present nearly everywhere (as we measured it) with a distinct sinusoidal dependence on latitude, quite different from the shape for the 3 \mu m band latitude dependence.

\textsuperscript{[54]} In an attempt to avoid the confusion of the correlation of mare units with low latitude on the lunar frontside, we studied a pole-to-pole limited longitude strip of data, which mostly covers highlands regions but does include Mare Orientale (Figure 9a). Figure 9b shows that the same general behavior as depicted in Figure 8 for the global data set but with a double cluster (higher and lower) for the 3 \mu m feature at lower latitudes. The upper of these two clusters contains mostly highland region pixels. The lower cluster on the right edge corresponds to pixels in mare regions (mostly in
Mare Orientale). The remainder of the pixels in the lower cluster corresponds to highland regions distinguished by having very steep, well-illuminated (and presumably high-temperature) slopes.

[55] It is interesting that there are two topographic slope types distinguished for the highland pixels (higher and lower band depth for lower-latitude clusters in Figure 9b for the 3 μm band depth plot): (1) steep and well illuminated with the lowest band strength and (2) all others. This suggests that there is something special about the very steep slopes and/or how we treat their analysis using M3 data or that the distribution of slopes for the lunar highlands is not random and there are few slopes except for the very steep ones. Another possibility may be related to the steep, well-illuminated slopes being much higher temperatures than other areas. This could imply a physical temperature effect or some artifact of the standard thermal emission removal algorithm, since this algorithm would operate only on surfaces with a temperature higher than about 250 K. At lower temperatures, thermal emission near 3 μm is not detected by the thermal emission removal algorithm.

6.4. Correlation With Composition

[56] From the latitude dependence analysis shown in Figure 9a it is clear that the mare regions (lower, right cluster for the 3 μm band) have lower band strength than the highland areas. This begs the question of the association in general of the absorptions with composition. To study this, we used the 1 and 2 μm pyroxene absorptions as a measure of composition, and Figure 10 shows how these are correlated with the 2.8 and 3 μm absorptions. The 3 μm and the 2.8 μm absorptions show little or no correlation with the 1 μm absorption. But, note that the 3 μm absorption is strongest where the 1 μm feature is very weak or nonexistent (as we measured it), which is mostly for highland-type material. Note also that the 2.8 and 3 μm bands are weakest where the 1 and 2 μm absorptions are strongest, which is in mare regions. There are local exception of some small fresh craters and some central peaks whose spectra display strong 1 μm absorptions and contain high OH [Pieters et al., 2009; Clark et al., 2011]. Figure 10 again shows that the OH/H2O-related spectral features occur preferentially for highland-type material and are weakest for mare-like
Figures 10c and 10d for the 2 μm pyroxene absorption show more of a linear anticorrelation and strongest 3 and 2.8 μm absorptions for weakest 2 μm absorption on a global average. This is the same global trend seen for the 1 μm absorption. In this case, we see again that the 2.8 μm absorption exists (as we measured it) everywhere, whereas the 3 μm absorption (as we measured it) is mostly absent where the 1 μm feature exists, as is the case for the 2 μm feature.

[57] We hasten to point out that to some extent the nature of Figure 10 is dependent on how these band strengths are measured. In our case, for the 1 and 2 μm bands, this is done by fitting a straight line, tied to the continuum on either side of the absorption, across the top of the absorptions, as described earlier. This is only an estimate, and is likely to have offsets in the absolute values, as is illustrated by the 1 and 2 μm values plotted having values below zero. Nevertheless, the overall trends should be correct.

[58] One compositional effect pointed out by Pieters et al. [2009, supporting online material Figure S1] is that small, fresh feldspathic craters in one subscene they studied showed stronger absorptions (both the 2.8 μm and the broader 3 μm bands) compared to a mature soil nearby. We have shown here that the 2.8 μm absorption is present for both mare and highlands, whereas the 3 μm absorption is mostly absent or is very weak (depending partly on how one measures band depth) for maria regions (Figures 8, 9, and 10). To explore this effect, in Figure 11 we show spectra for a selection (15) of small fresh craters in highland regions taken from the data set used for Figure 9. Although located in general highland regions, these craters display 1 and 2 μm absorptions that indicate a range of composition, feldspathic as well as olivine- and pyroxene-rich. The spectra are organized so that they range from the brightest to the least bright craters from top to bottom. This arrangement of the spectra also correlates with composition. All these spectra
show the 2.8 μm OH absorption but none show the 3 μm absorption, usually evident as a broad turndown in the spectrum toward 3 μm (Figure 6). All the bright-crater spectra show a stronger 2.8 μm absorption than seen in the mature highlands soils (Figure 6) [Pieters et al., 2009, Figure S2]. Thus, there is something about the fresh highlands crater material that encourages OH formation. Perhaps this property is the fresher exposed grain microsurfaces.

Figure 11. Examples of fresh soil spectra from small fresh craters with various compositions (feldspathic, olivine-rich, and pyroxene-rich) that show absorption feature mostly at 2.8 μm. All these fresh soils are in highland regions, but the compositions vary from normal highlands felspathic to more mare-like mafic (e.g., note 1 μm region absorptions). (a) These regions differ in brightness, ranging from brightest to darkest, top to bottom. (b) However, the spectra have been scaled to have the same reflectance at 2.656 μm, thus normalizing the absorption strengths. No thermal emission has been removed, and no visual evidence of the presence of thermal radiation is apparent.

Figure 12. Plots of the strength of the (a) 3 μm and (b) 2.8 μm absorption features versus lunar surface temperature. The surface temperature calculation is described by Clark et al. [2011].
that are more crystalline and more chemically reactive than the mature, more glass-rimmed soil grains. The explanation for the lack of the 3 \( \mu \)m absorption is unclear. Perhaps it is being filled in by thermal emission, although there is no visual evidence of thermal emission in any of the Figure 11 spectra, and the \( \text{M}^3 \) team thermal emission removal algorithm would not detect any in these spectra, and it would be just the right amount so as to balance a 3 \( \mu \)m absorption.

So, most likely the 3 \( \mu \)m is very weak or absent for these craters at this time of observation.

The 2.8 \( \mu \)m absorption seems to decrease in strength with a decrease in brightness (Figure 11a). However, when the spectra are all scaled to the same reflectance at 2.656 \( \mu \)m (Figure 11b), the 2.8 \( \mu \)m absorption appears about the same for all craters. Thus, the strength of the absorption is approximately the same proportion of the scattered light in all cases. This illustrates another complication in attempting to investigate the dependence of these absorptions on physical properties. Here, the multiscattering nature of the lunar highlands fresh soils creates an anticorrelation of apparent absorption band strength on the amount of scattering (apparent brightness) within the surface material.

### 6.5. Correlation With Surface Temperature

The dependence of the two absorptions on surface temperature was also examined in detail. Figure 12 shows a correlation plot for the global data set. A problem in this part of the analysis is that we cannot determine the surface temperature below about 250 K from the \( \text{M}^3 \) data, due to the lack of spectral coverage beyond 3 \( \mu \)m that is needed to be sensitive to colder temperature thermal emission. The method of calculating the effective surface temperature and removing the thermal emission from the \( \text{M}^3 \) data is described by Clark et al. [2011]. It basically involves fitting a straight line to the lunar spectrum continuum shortward of about 2.5 \( \mu \)m, assuming that the lunar reflectance spectrum continues with the same slope through the 3 \( \mu \)m region, and attributing to thermal emission any contribution that causes the spectrum to rise above that extrapolated continuum from 2.5 to 3.0 \( \mu \)m. A thermal emission radiation model is used to associate a temperature with the removed signal. For the...
The thermal emission removal algorithm has been applied. Plots of how the measurement of absorption band strength would be affected by incomplete or inaccurate removal of thermal emission contribution to the measured flux for the 3 μm absorptions. (a) The model calculations produce the curve, showing the band apparent strength (as we measure it) decreasing as the surface temperature rises if no thermal removal is attempted. (b) The 3 μm absorption band depth from M^3 spectra as a function of actual temperature measurements by the Diviner investigation. M^3 data come from a subset of observation M3G20090212T162141 of Mare Orientale and adjacent highlands taken at 807 local lunar time. Diviner observations have been acquired at 750 local lunar time.

Figure 14. Plots of how the measurement of absorption band strength would be affected by incomplete or inaccurate removal of thermal emission contribution to the measured flux for the 3 μm absorptions. (a) The model calculations produce the curve, showing the band apparent strength (as we measure it) decreasing as the surface temperature rises if no thermal removal is attempted. (b) The 3 μm absorption band depth from M^3 spectra as a function of actual temperature measurements by the Diviner investigation. M^3 data come from a subset of observation M3G20090212T162141 of Mare Orientale and adjacent highlands taken at 807 local lunar time. Diviner observations have been acquired at 750 local lunar time.

6.5.1. Limitations of the Thermal Removal Correction

An example of the confusion that may be caused by any remaining, unremoved thermal emission effects arises when we attempt to study lighting geometry effects on the 3 μm absorption. For example, Figure 13 shows an area where the 3 μm band depth seems to follow the lighting conditions, for example, for the crater near the upper middle part of the image but also for other regions with topography in the image. This data set has no thermal emission removed, and so thermal emission will appear to fill in any absorption longward of about 2.5 μm and make the band appear less strong. Of course, lighting conditions are strongly correlated with surface temperature. Thus, to separate these effects, the thermal emission must be removed.

A further concern is that the thermal emission removal algorithm used on the M^3 data and the lack of spectral information beyond 3 μm will limit the effectiveness of thermal emission removal and/or cause errors, and thus still will be affecting our calculation of absorption band strengths even for data treated by the M^3 thermal emission removal algorithm. The method used by the M^3 team [Clark et al., 2011] is simplistic, although about the best that can be done given the limited M^3 data long-wavelength spectral coverage. For example, the surface temperature experienced within a pixel is not just a single value: it is a complicated combination of several temperatures created by thermal gradients in the soil and by subpixel mixing of surface elements of different temperatures due to a number of effects including thermal inertia, albedo and topographic differences. These effects vary from pixel to pixel.

6.5.2. Uncertainties and Errors due to Remaining Effects of Thermal Emission

By modeling, we explored how an error in thermal removal might affect the band depth calculations. It appears from both models (Figure 14a) and observations (Figure 14b) that incomplete thermal radiation removal will tend to fill in the absorption bands and effect the band strength measurements by reducing the apparent strength. This affect must be kept in mind when attempting the sort of investigation we are describing here, i.e., of the relationship of the absorptions to other properties that may themselves be associated with surface temperature. A similar but more complicated affect occurs for the 2.8 μm absorption apparent strength as well. This is because the 2.8 μm absorption is affected by the shape of the 3 μm absorption, on which it is superimposed, as well as the broader spectrum shape. The effect of thermal emission contamination on the M^3 spectrum absorption strengths is being further investigated, but its effect is not fully understood at the time of this article.

6.5.3. Need for Accurate Surface Temperature Measurements

The effect that thermal emission has on the apparent strength of the two absorptions under investigation indicates that better surface temperature measurements are needed. Fortunately, the Diviner investigation on the Lunar Reconnaissance Orbiter spacecraft [Paige et al., 2010] is providing these needed measurements. We illustrate here how sensitive the absorption band strength measurement can be to thermal radiation contamination related to surface temperature. Figure 15 shows two strips of M^3 data for approximately the same area over Mare Orientale, one at 0807 and...
one at 1535 local lunar time. These data strips are color-coded to indicate 3 \(\mu\)m absorption band strength as measured by our method. Figures 15a and 15b have no thermal correction made. Figures 15c and 15d have thermal emission removed according to a model created by Diviner and using their temperature calculation. Diviner did not measure the same M\(^3\) area at exactly the same time, but rather at 0750 and 1502 lunar time. The Diviner surface temperature model under development was used to calculate from the Diviner measurements what the thermal emission and temperature would be at the exact time of the M\(^3\) measurements. As can be seen, the morning and afternoon band depth measurements/calculations still do not match, with the 3 \(\mu\)m absorption appearing stronger in the afternoon. This suggests that the bands are stronger for higher temperatures, counter to what the M\(^3\) data analysis suggests (Figure 12) or that the Diviner model still is not sufficiently correct, or possibly that multiple reflections as proposed by Clark [2009] are influencing the apparent absorption strengths. Thus, this analysis using the Diviner thermal model and temperature measurements should continue as well as study of indirect illumination from local topography.

6.6. Distinction Between Temperature Effects and Surface Composition Effects in M\(^3\) Spectra

It is beyond the scope of this article to develop methods to separate incomplete thermal emission removal effects from actual absorption band depth variations on the lunar surface or to model multiple reflections from lunar topography. But, we can present evidence that some variations in the band exist that are not due to thermal emission, in addition to making the reader and future M\(^3\) data users aware that interpretations of band depth variations can be complicated and misleading. One way of demonstrating
actual band depth changes is described in Figure 16. Here, the ratios of two spectra are presented, one for each of two different illuminations and temperatures (Figures 16a and 16b) for the same location. These ratios (Figure 16c) show the expected smooth spectral shape for a difference in thermal emission contributions. It appears that there is a difference in surface temperature and thermal emission component between the two spectra used in each of these ratios. Further, the amount of thermal emission implied by the two ratio spectra in Figure 16c suggests that there is less temperature difference for the morning-illuminated/afternoon-shaded locations than for the other case, probably because the surface warming-cooling relationship with time and illumination is asymmetric. This helps indicate that we are dealing with thermal emission in this case.

Figure 16e shows the ratio of the spectra (Figure 16d) for two similar areas within Ryder crater under different illuminations. There is a definite kink in this ratio spectrum at the wavelength where the 3 μm OH absorption is expected to occur (about 2.6 μm). This shape suggests that the OH/H₂O absorption is different between the two spectra used in the ratio and the different band depth is not (at least entirely) a thermal emission effect.

It is interesting that the band depth of the 0.9 μm pyroxene band also changed with illumination (Figure 16d). This may indicate an effect unrelated to thermal emission also is influencing the spectrum, possibly multiple reflections from surface topography. There are many other examples of this effect, and this suggests a method to explore the behavior of the 3 μm OH/H₂O absorption somewhat independent of correcting for the thermal emission effects.


An obvious answer to this limitation in thermal emission correction is to measure to longer wavelengths, and this is strongly recommended for any future observational investigations. The EPOXI spacecraft spectrometer did cover the spectral range to longer wavelengths and more

Figure 16. Examples of spectra of the same or similar areas (Ryder crater) under morning (M3G20090125T172601) and afternoon (M3G20090528T213152) lighting conditions and their ratios. This is to illustrate a method for separating differences due to changes in 3 μm absorption from differences in thermal emission contributions. Thermal emission has not been removed from the M₃ spectra that are treated here. (a) Spectra for the same area under morning/low lighting and afternoon/well illuminated. (b) Same as Figure 16a but opposite lighting conditions. (c) The ratio of the spectra in Figures 16a and 16b. (d) Spectra from similar areas of Ryder crater under different lighting conditions from the same data set. (e) ratios of the spectra from Figure 16d. The ratio spectra shapes in Figure 16c are similar to what is expected for differences in thermal emission, while the ratio spectra shapes in Figure 16e have kinks at the wavelength where the 3 μm absorption begins, suggesting an actual difference in the 3 μm band strength from place to place, perhaps due to physical effects of temperature or lighting on the process creating the OH/H₂O responsible for the 3 μm absorption.
fully revealed the 3 \( \mu \text{m} \) absorption than do the M\(^3\) measurements, although the M\(^3\) data have much greater spatial resolution and coverage [Sunshine et al., 2009]. Figure 17 shows examples of our analysis of EPOXI observations [Sunshine et al., 2009] for a highland region at different times of the lunar day. The 3 \( \mu \text{m} \) feature varies in strength in a way that suggests that OH/H\(_2\)O is least abundant at midday. The absorption appears to change in a symmetric way across the absorption, unlike what is expected if thermal emission is mainly responsible, which would be greater filling to longer wavelengths. This conclusion is also suggested by the apparent near symmetry of the plot of band strength with time of day (Figure 17b). This suggests that lighting angle is more important than surface temperature in controlling the strength of the 3 \( \mu \text{m} \) absorption and the amount of OH/H\(_2\)O, which along with similar changes in mineral bands like the 0.9 \( \mu \text{m} \) absorption are an indication of topographic lighting effects.

### 6.8. Other Effects That Might Confuse Analyses of the 2.8 and 3 \( \mu \text{m} \) Absorptions

[69] Changes in the strengths of absorptions in the lunar soil spectrum with time/illumination angle would be important for understanding the physical process(es) creating the molecules responsible. We have shown apparent evidence for these and then have pointed out possible complicating factors. One of these is the effect of reradiation, that is sunlight that is scattered from one region of the Moon surface, strikes a second region, and is again scattered before heading off to the M\(^3\) [Clark et al., 2011]. We show the results of one test for this effect in Figure 18. We used a shaded area within a crater that showed sufficient indirect illumination to produce an M\(^3\) signal and support calculation of a reflectance spectrum (top solid line). This scattered light also could be from instrument scattered light (which exists in M\(^3\) [Green et al., 2010], as in most all spectrometers). We estimated the instrument scattered light in this case by using the M\(^3\).
signal from a totally shadowed region (bottom solid line), and we used the mature highlands spectrum from Figure 6 as representative of the once scattered sunlight. Note that the instrument scattered light spectrum has no resemblance to the lunar surface spectrum and therefore probably contains little light scattered from another region, but rather is due to properties of the instrument optics. We then modeled the observed shaded area spectrum by a linear combination of once and twice scatter light and instrument scattering. The best match to the observed spectrum is a combination of instrument scatter light and once-scattered light (top dashed line). The model using only twice-scattered light and instrument scattered light (middle spectrum) could not be made to match well the observed spectrum (top solid line).

The direct illumination model has a ratio of 2.27% of direct light (from the mature highlands feldspathic soil spectrum shown in Figure 6) over instrument scattered light. This model is a linear combination of 0.047 × the illuminated soil spectrum and 2.089 × the deep shadow spectrum (assumed to be mostly instrument scattered light). The conclusion is that in this case most of the light from the observation is from once-scattered sunlight and scattering in the instrument. The significance is that twice-scattered light has the effect of doubling any absorption due to the surface material (assuming the material at the origin of the scatter light and at the observed site are similar), and this effect can mislead studies of specific regions [Clark et al., 2011].

7. Discussion

[70] The analysis of the behavior of the 3 μm region absorptions in the M^3 data described in section 6 helps to confirm that the 3 μm and the 2.8 μm absorptions exist. The search for dependences of the absorption on physical properties is more difficult, due to the limited coverage (space and time) of the M^3 data and to the mixing of thermal emission and reflected solar radiation in the 3 μm region. We look forward to new M^3-like observations of the Moon with longer-wavelength spectral coverage in the near future in order to help untangle the complicated behavior of the absorptions and better understand the physical processes creating the OH/H_2O molecular groups responsible for the spectral absorptions.

[71] Of the three possible sources of OH and water on/in the Moon, discussed early in this article, the widespread nature of the 2.8 and 3 μm region absorptions alone [Pieters et al., 2009], seems to us to favor solar wind-proton-induced hydroxylation as the source of OH and possibly H_2O. Water in the lunar interior, suggested by recent analysis of lunar sample material thought to come from the deep interior [e.g., Saad et al., 2008], seems not likely to exist in the widespread deposits observed by M^3. Hydroxylated and hydrated material from infall could be widespread, but it should not be correlated with lunar composition as the spectral absorptions seem to be, and, again, no evidence of this kind of infall material is found in the returned lunar samples. However, solar wind products are found in the lunar samples and are associated with smaller grains with larger surface areas.

[72] From the analysis above, we conclude that the behavior of the 2.8 μm and 3 μm absorptions is somewhat different, although their strengths are correlated, suggesting the same or related processes. These two absorptions might be due to somewhat different molecular groups and may be formed by somewhat different stages of the process(es).

[73] The 2.8 μm feature, usually assigned specifically to the OH molecular group, appears to have the greater presence (as we measured it). The broader 3 μm absorption also could be created only by OH but could also at least partly be due to H_2O, and if so, then the implication is that OH is more abundant. This effect seems consistent with the solar wind proton implantation process, as OH would be formed first and most often. H_2O is more likely to be formed less often and mostly only when the OH formation process becomes close to saturation. In the solar wind-induced hydroxylation process, H_2O is more a byproduct after production of OH. Nevertheless, this byproduct could be the source of deposits of H_2O, particularly after transport to cold traps in permanently shadowed regions or as the result of hydration of lunar materials.

[74] Several lines of evidence are presented suggesting that there is a compositional effect, with more absorption for highland material. This conclusion is also consistent with other studies [Pieters et al., 2009]. The compositional effect seen in M^3 data is possibly due to a higher abundance of volatiles in the highlands than in the mare, probably both because of the formation process and stability at high temperatures. Protons from the solar wind are more likely able to break molecular bonds in highland-type minerals such as anorthite than in mare-type minerals such as pyroxene and olivine because of crystalline binding energy. On the other hand, adsorbed water in mare materials is released at lower temperatures than adsorbed water in highland materials [Graham et al., 1979], which is consistent with stronger absorption bands at 2.8 μm and 3.0 μm on fully illuminated highland surfaces than on fully illuminated mare surfaces.

[75] The relatively stronger 2.8 μm and 3.0 μm absorption bands in highlands spectra than in mare spectra may also result from an observation bias. For the same concentration of OH at the surface, reflectance spectra show a stronger absorption band for highlands than for mare [Starukhina and Shkuratov, 2010].

[76] There appears to be a small dependence of the band strength with temperature (less absorption with higher temperature), which suggests that the amount of absorber at the optical surface is decreased and perhaps less is formed at higher temperatures. This is after the M^3 “standard” thermal removal process is applied.

[77] Complicating the interpretation of these apparent relationships is the obvious evidence that incomplete thermal emission removal can affect the calculation of the 2.8 and 3 μm band strengths (more thermal, less apparent absorption band strength). Since temperature is correlated with most of the physical parameters studied (composition, latitude, lighting), thermal emission removal is essential to improve on the study presented here.

[78] Other observational biases also can exist, such as the presence of rescattered light, which can enhance absorption band strengths and other color effects when the materials responsible for the first and subsequent scattering are similar. This must exist in some cases, but we did not find evidence for it in the test case we explored. Similarly, photometric function effects, controlled by surface topography and light
conditions in general, can enhance absorptions (more absorption as the sun illumination is farther from the surface normal). These effects require a detailed surface topography (digital terrain model) map to study effectively.

[79] An interesting question is “How much OH/H₂O?” We make one form of estimate here, following the calculations of Starukhina and Shkuratov [2000; Johnson, 1990]. O–H therefore might be created in amounts up to 1–5 x 10⁴⁵ cm⁻² of surface area. This is in agreement with laboratory studies of H trapped in OH groups per unit surface of oxides [Gruen et al., 1976; Siskind et al., 1977]. One liter of water = 1 kg = 55.55 mol of O = 3.3 x 10⁴⁵ atoms of O. The surface area needed to provide this amount is >6 x 10⁴⁵ cm² or 60 m² square area (or about one football field) if all the OH were driven to form H₂O. Gardening of the surface and migration of H [Grieses et al., 2010] would bury OH-saturated surfaces and expose fresher material, increasing the overall amount of OH per surface area.

[80] These observations and the processes involved likely have implications for other airless silicate bodies in the inner Solar System. The solar wind streams out from the sun and strikes not only the Moon, but also Mercury and Vesta, for example. The distance from the Sun affects the particle flux density but not so much the particle energy. Thus, similar processes may be operating on these other objects but perhaps with longer or shorter timescales. Vesta is of interest because the Dawn spacecraft [Russell et al., 2007] is approaching this object and should be able to make M³-like observations with less contamination by thermal emission. Thus, the results discussed here might be tested by Dawn. Observations of silicate asteroids also might show evidence of OH formation, such as the recent Rosetta observations of the asteroids Steins and Lutetia.

[81] Many of these results and conclusions presented here are at least somewhat dependent on the technique used to measure the absorption strengths. More research and more complete observations are needed, including digital terrain and temperature models. The results and interpretations presented here are intended partly as a beginning to the M³ analysis process, to give some knowledge basis for future research, and to help future users of M³ data to understand the power and the limitations of the data set.

Acknowledgments. We thank D. Dyar for helpful comments and a reading of the manuscript. The manuscript also benefited considerably from a careful review and editing by David Blewett and one other reviewer. This research was supported by the NASA M³ Project under contract from Brown University.

References

Green, R. O., C. M. Pieters, J. Boardman, M. Eastwood, P. Mouroulis, S. Lundeen, and M. White (2010a), Assessment of the complete Moon


---

R. N. Clark, U.S. Geological Survey, MS964, Box 25046, Federal Center, Denver, CO 80225, USA. (rclark@usgs.gov)

J.-P. Combe, G. Kramer, and T. B. McCord, Bear Fight Institute, 22 Fiddler’s Rd., Box 667, Winthrop, WA 98862, USA. (tmccord@bearfightinstitute.com)

C. M. Pieters, Department of Geological Sciences, Brown University, Box 1846, Providence, RI 02912, USA. (Carle_Pieters@Brown.edu)

J. M. Sunshine, Department of Astronomy, University of Maryland, College Park, MD 20742-2421, USA. (jess@astro.umd.edu)

L. A. Taylor, Planetary Geosciences Institute, Department of Earth and Planetary Sciences, University of Tennessee, 1412 Circle Dr., Knoxville, TN 37996-1410, USA. (lataylor@utk.edu)