See next page for full paper.

Seeing through heavily polluted satellite imagery using QUAC

Steven Adler-Golden[*], Lawrence Bernstein, Benjamin St. Peter and Bridget Tannian

Spectral Sciences, Inc., Burlington, MA 01803

ABSTRACT

Extremely thick haze caused by air pollution is observed in many satellite images of the earth, and in particular over eastern China. Standard image display software typically provides satisfactory visualization of the ground through automated or user-driven scaling to enhance contrast; however, it does not perform well with these highly polluted scenes, where the haze is spatially non-uniform. Furthermore, estimation of surface reflectance using standard atmospheric correction software is highly problematic under these conditions due to very low visible transmission of the haze coupled with lack of knowledge of its optical properties, which may not conform to the haze or aerosol models in the software. In this paper we show that a version of the empirical Quick Atmospheric Correction (QUAC) algorithm, adapted for spatially dependent scattering, produces visually satisfying imagery of the entire ground in multispectral satellite scenes containing thick haze, and that the output reflectance spectra appear to be realistic enough for performing basic surface classification. The QUAC algorithm is applicable to multispectral and hyperspectral imagery with any number of wavelength bands, including true color (RGB) imagery, and does not require radiometrically calibrated data.

Keywords: atmospheric correction, multispectral, image, haze, air pollution, Planet Labs, QUAC

1. INTRODUCTION

Atmospheric correction algorithms are used for processing spectral imagery to remove absorption and scattering effects of the atmosphere, yielding imagery in units of spectral reflectance. The processing provides a number of benefits, including removing the effects of haze and solar angle variation, providing uniform image values for assembling large scene mosaics, and enabling identification of materials and terrain via comparisons with library spectra.

When atmospheric effects are reasonably uniform within an image, the relationship between the original measurements and the reflectance is close to linear. This is the basis of the Empirical Line Method (ELM)[1] for atmospheric correction using known surfaces in the scene. In this situation, atmospheric correction provides little or no benefit for image visualization with auto-scaling viewing software or for terrain classification based on in-scene information. However, under very hazy conditions, there is often enough patchiness in the haze to warrant a non-uniform atmospheric correction for most types of data exploitation, including visual inspection. In addition, the sensor's angular subtense can introduce apparent haze variation due to the variation in path length to the ground.

Spatially non-uniform atmospheric correction is relatively uncommon. In addition, quantitative atmospheric correction with very thick haze, whether patchy or uniform, is difficult to perform with first-principles methods due to uncertainty in the optical properties of the haze. This paper describes an empirical approach to atmospheric correction of multispectral imagery with thick and variable haze based on the Quick Atmospheric Correction (QUAC) algorithm. The basic QUAC algorithm[2] derives gain and offset spectra that are used to linearly transform the original measurements to reflectance units. The measurements need not be radiometrically calibrated, and nothing needs to be known about the illumination, the atmospheric conditions, the haze optical properties, or materials on the ground. The standard QUAC method is capable of obtaining good results with very hazy imagery when the haze is fairly uniform. Here we describe an improved QUAC-based method that handles situations where the haze also varies spatially. The new method derives an offset spectrum that varies smoothly across the image, and, in this initial version, combines this spectrum with a uniform gain spectrum.

*adlergolden@spectral.com; phone 1 781 273-4770; fax 1 781-270-1161; spectral.com

The new atmospheric correction method is illustrated using very hazy color (RGB) imagery of eastern China taken from Planet Dove satellites. The Planet constellation (Planet Labs, Inc., www.planet.com) is a set of small polar-orbiting satellites that are designed to collect daily multispectral four-band images over the whole globe. As test cases for our atmospheric correction software, Planet generously provided us with a series of three-band (RGB) images at ~4 m resolution taken over southeastern China under conditions of extreme air pollution. The data lack an absolute radiometric calibration, but it is not required for our current analysis.

2. EXAMPLE DATA

Figure 1 shows auto-scaled ("2% Linear") ENVI software displays of an example RGB image with very thick haze. The scaling in the display at left includes the zero background values beyond the edges of the data, and thus depicts the absolute signal level. The haze is so thick that the bulk of the signal is from the haze; ground features are barely visible. The Figure 1 display at right is auto-scaled to the signal range in the red box, where the image is a little less hazy than elsewhere. Here the ground is clearly visible, but many areas outside the box, especially towards the edges of the scene, remain hazy and obscure the ground features.

The small size of the ground signals relative to the haze is confirmed in Figure 2, which shows a horizontal slice of the signals in the red box. Here the largest ground signal, corresponding to a reflectance estimated as ~0.5, is only around one-quarter to one-fifth as large as the underlying haze signals. The slightly higher haze signal in the blue band compared with the others probably reflects Rayleigh scattering from air molecules rather than the color of the particulate Mie scattering.

The QUAC atmospheric correction process subtracts an offset equal to the minimum value of the data in each band. Thus, its visual output is equivalent to Figure 1 at right. Similar results would be obtained using other atmospheric correction methods that do not account for haze variation within the image, including the ELM and well-known first-principles algorithms[3,4,5]. (An exception is ATCOR4, which can correct for variable haze and thin clouds[6].)

In the following section we review the standard QUAC method, and in Section 4 we describe an upgrade for spatially varying haze that has proven very successful with all of our hazy Planet imagery, including the Figure 1 data.
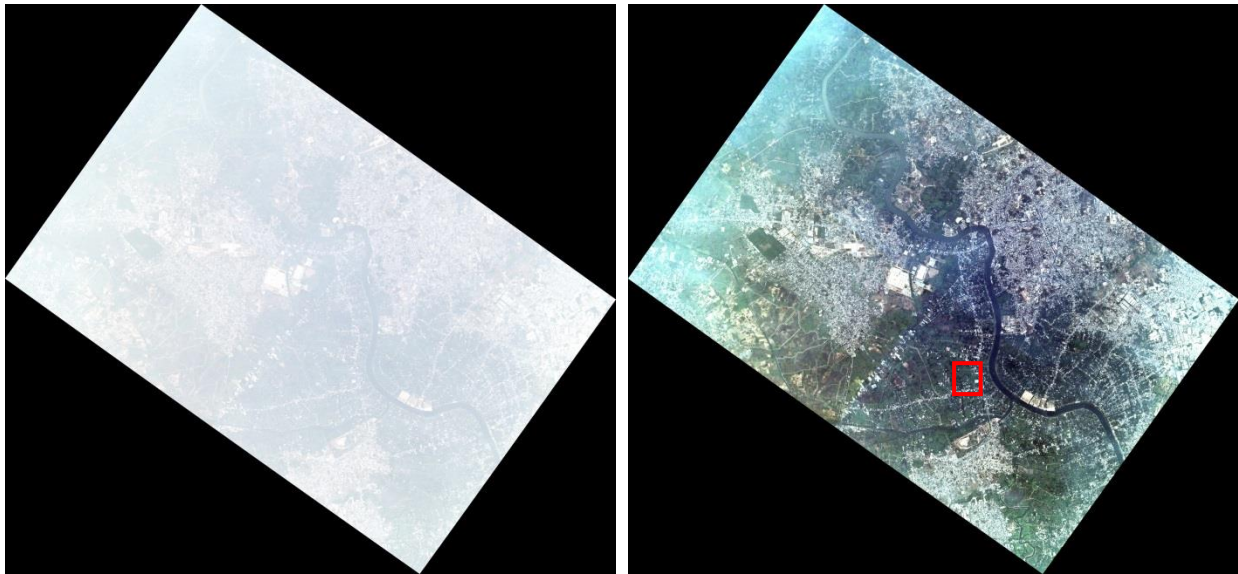

Figure 1. Hazy Planet image of southeastern China taken on March 31, 2016. At left, scaled to the full data range, including zero; at right, scaled to the data range within the red box.
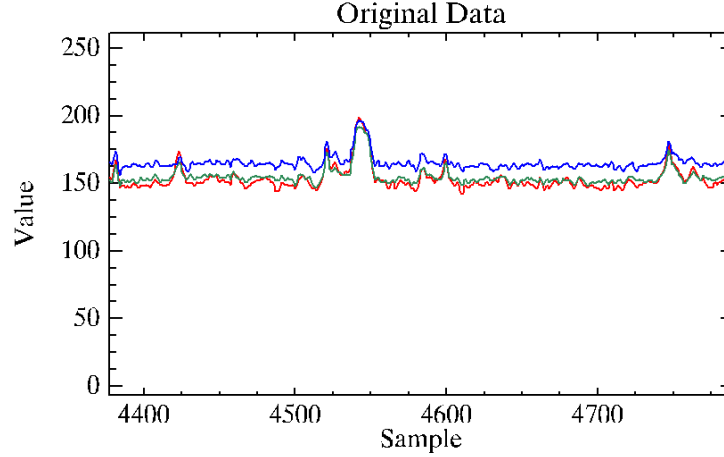
Figure 2. A horizontal slice of the Figure 1 image within the red box.

3. QUAC OVERVIEW

We briefly review the key concepts underlying QUAC for a scene with uniform atmospheric properties. A more detailed description is available in Bernstein[2]. The observed spectral radiance, L_{obs}, for a pixel with surface reflectance, ρ_{sur}, is the sum over the three optical paths in Figure 3,

$$L_{obs} = \left(A + C\rho_{ave}\right) + B\rho_{sur}.$$

The components in $(A+C\rho_{ave})$ are grouped together because they tend to be approximately constant over an image, and thus can be considered as an offset common to the image pixels. This simple linear relationship can be re-arranged to express the retrieved surface reflectance in terms of the observed signal and derived atmospheric parameters, such that

$$\rho_{sur} = Gain(L_{obs} - Offset),$$

where $gain=1/B$, and $offset=(A+C\rho_{ave})$. For a physics-based approach, A, B, and C are retrieved by comparison of certain spectral features to those predicted by radiation transport calculations. In QUAC, we determine these parameters directly from the in-scene spectral data and a key underlying assumption. The offset is given by the minimum signal in each band, after noise and data dropouts have been filtered out. The gain is given by

$$Gain = \frac{< \rho_{end} >_{lib}}{< (L_{obs} - C\rho_{ave})_{end} >},$$

where $<\rho_{end}>_{lib}$ is the average of the endmember spectra representing a reference library of material reflectance spectra, and $<(L_{obs}-C\rho_{ave})_{end}>$ is the average of a collection of endmembers retrieved from the observed in-scene pixel spectra. An endmember represents an "extreme" spectrum from a collection of spectra. In most cases, sparse positive linear combinations of tens of endmember spectra can accurately represent the much larger number of spectra associated with an image or a comprehensive spectral library. In QUAC, we use the SMACC (Sequential Maximum Angle Convex Cone) algorithm to find endmembers[7]. With multispectral data, SMACC is tasked to find roughly an order of magnitude more endmembers than the number of bands.

The gain equation above embodies the key, empirical QUAC assumption, which holds reasonably well for most scenes, that the average of diverse endmember reflectance spectra, excluding highly structured materials (i.e., vegetation, shallow water, mud), is always the same. More specifically, as long as an image contains at least ~10 or more spectrally diverse materials, their average reflectance spectrum can be taken as a "universal" reference. The materials may include both natural and manmade materials, such as bare soil, deep water, bare rock, vehicles, roofs, roads, etc. The universal reference is calculated by running SMACC on a large library of non-vegetated material reflectances and applying a small final correction to the average.
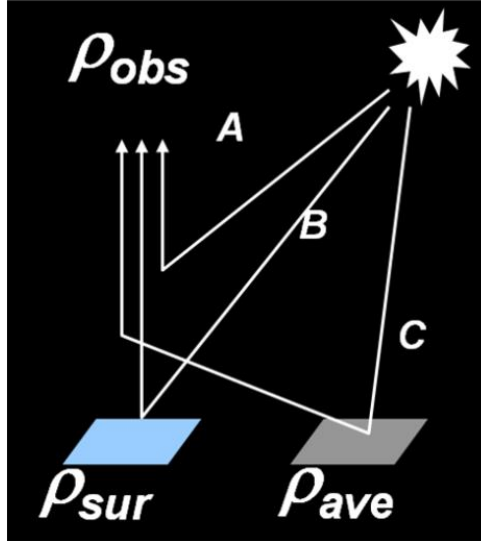
Figure 3. Radiance contributions to the apparent reflectance ρ_{obs} of a pixel.

4. MULTISPECTRAL QUAC UPGRADE FOR NON-UNIFORM HAZE

To handle spatially varying haze, we employ QUAC with a locally varying offset spectrum derived from sub-regions of the image, taken as square tiles. Ideally, the gain spectrum would also vary locally; however, much larger areas of the image are often required to obtain enough diverse endmembers for the gain calculation. Here we calculate the gain in the usual QUAC manner from the entire image once the offset is calculated and subtracted.

Calculating locally varying offsets proved to be less straightforward than one might expect. The tiles need to be small enough to account for the haze spatial variation, but large enough to ensure that the vast majority of them contain dark pixels. The tile offsets need to be smoothly blended across the image so that edge effects are not apparent. Tiles at the edges of the image may contain insufficient data for a good offset determination, and interior tiles may also produce erroneous offsets due to data dropouts or lack of dark pixels. To deal with these challenges, we developed the procedure described below:

1. The image is divided into ~150x150-pixel tiles, and 3x3 median filtering is applied. Tiles are declared valid if at least ¼ of their pixels contain valid data. In addition to excluding the zero-valued pixels beyond the edge of the image, we exclude a 35-pixel border of the image where there are dimmer pixels, presumably due to camera vignetting. For each valid median-filtered tile, a provisional tile offset spectrum is defined as the minimum signal in each wavelength band.

2. The 2-D array of provisional tile offsets, denoted **T**, is then filtered to remove anomalous values, which may arise from occasional data dropouts or lack of sufficiently dark terrain in the tile. In this step, tile offsets are declared invalid if they are greater than three standard deviations from the mean, and logical values, either 1 (valid) or 0 (invalid), are placed in a mask array **M**.

3. The remaining, valid tile offsets will be used to derive an offset data cube—that is, a unique offset spectrum for each pixel—by resampling **T** to the full-size image using linear interpolation. Since the interpolation is performed only between tile centers, to handle the image edges we enlarge the **T** array with a one-tile border, thus increasing its dimensions by two in each direction.

4. To replace missing or invalid tile offsets with valid offsets and smooth the **M** array, we use 3x3-element averaging over the valid offsets, given by **TM**/**M**, where the array product in the numerator denotes element-by-element multiplication and the underline denotes averaging. Any remaining unassigned offsets (i.e., where **M** equals zero) are replaced with the mean offset, and a 3x3 median filter is applied to yield the final **T** array.

5. The **T** array is resampled to the image pixel locations using linear interpolation.

The resulting offset data cube is subtracted from the original data cube, and the result is processed with standard QUAC.

5. RESULTS AND CONCLUSIONS

Figure 4 shows the result of processing the image of Figure 1 with the new variable haze method. Here the excess haze outside the red box area is removed, and the entire scene has a good visual appearance. Figure 5 shows the data slice corresponding to Figure 2. The underlying haze has been completely removed, and the data are now in reflectance units. Since we do not have ground truth measurements for this scene, the absolute accuracy of the reflectances is unknown, but it is likely within the typical QUAC accuracy of ~25%.

Figure 4. Image displayed as in Figure 1 (left), but after processing with the variable-haze QUAC atmospheric correction.
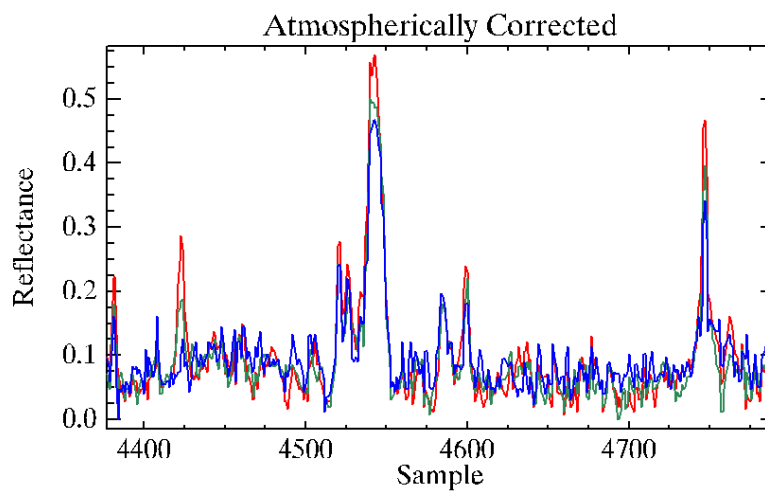

Figure 5. As in Figure 2, but after variable-haze QUAC atmospheric correction.

In summary, we have demonstrated spatially non-uniform atmospheric correction on extremely hazy RGB imagery from the Planet constellation using an empirical, QUAC-based method, producing reflectance imagery with good visual quality and superior to those from methods that assume a fixed amount of haze in the scene. The method is applicable to spectral imagery with any number of wavelength bands, and does not require an absolute radiometric calibration or any information other than the image itself.

The empirical nature of the QUAC atmospheric correction has the distinct advantage of avoiding uncertainties, inherent in first-principles methods, associated with lack of knowledge of the haze optical properties, i.e., the aerosol model. In particular, first-principles methods use the atmospherically scattered radiance, which includes adjacency scattering, to infer aerosol optical depth (AOD), defined as $-\ln(\tau)$ where τ is the vertical direct transmission from ground to space. The size of the scattered radiance is dependent on the single-scattering albedo and phase function of the aerosol, so that uncertainty in these parameters leads to uncertainty in the transmission and hence in the magnitude of the inferred ground reflectance. If dark urban aerosol models are considered, this uncertainty can be very large. This is illustrated in Figure 6, which shows MODTRAN calculations of atmospherically scattered radiance at 550 nm for a surface albedo of 0.2, which is an estimated average for the above scene, and two different MODTRAN aerosol models, rural and urban, which we take as limiting cases. The AOD range of 0.5 to 2.3 corresponds to a visibility range from 20 km to 2.5 km and a vertical transmission range from 0.6 to 0.10. With the urban model, the scattered radiance signal is insensitive to AOD, making AOD retrieval uncertain. Moreover, depending on whether either of the two aerosol models, or an intermediate one, is selected, a scattered radiance value of 6×10^{-3} W/cm^2/sr/μm might correspond to any AOD in this range and beyond. The corresponding uncertainty in vertical transmission would be at least a factor of $0.6/0.10 = 6$, leading to an even larger uncertainty in the retrieved reflectance, considering that the illumination at the ground is also varying.

It should be noted that recent upgrades to the Planet constellation are yielding higher quality imagery than what is shown in this paper. Prior to mid-2017, the constellation included satellites launched from the International Space Station, whose orbital plane crosses the equator at varying times of the day. Afterwards, Planet started operating its production flock at a polar sun-synchronous orbit with crossing times mid-morning at the equator, providing consistent illumination conditions across long periods with a daily revisit time. Results might vary and likely improve with a morning consistent crossing time. Additionally, unlike the data shown here, the newer data are radiometrically calibrated and contain NIR band information. A thorough analysis of this newer data has the potential to yield much better information on aerosol optical properties, with applications to environmental monitoring as well as atmospheric correction.
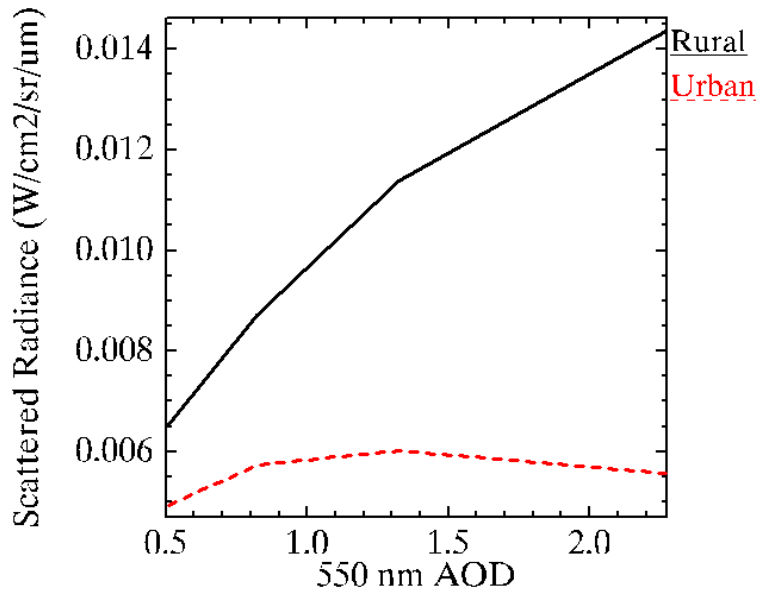

Figure 6. MODTRAN-predicted scattered radiance for two different aerosol models and a surface albedo of 0.2. The sun is overhead.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] Smith, G.M. and E.J. Milton, "The use of the empirical line method to calibrate remotely sensed data to reflectance," International Journal of Remote Sensing, 20(13), 2653-2662, DOI: 10.1080/014311699211994 (2010).

[2] Bernstein, L.S., X. Jin, B. Gregor and S.M. Adler-Golden, "Quick atmospheric correction code: algorithm description and recent upgrades," Opt. Eng. 51(11), 111719 (2012).

[3] Perkins, T., S.M. Adler-Golden, M.W. Matthew, A. Berk, L.S. Bernstein, J. Lee and M.J. Fox, "Speed and Accuracy Improvements in FLAASH Atmospheric Correction of Hyperspectral Imagery," SPIE Optical Engineering, 51(11), 111707 (2012).

[4] Green, R., "Atmospheric Correction Now (ACORN)," developed by ImSpec LLC, available from http://imspecco.ipower.com (2018).

[5] Montes, M.J., B.-C. Gao, and C. O. Davis, "NRL atmospheric correction algorithms for oceans: Tafkaa user's guide," Naval Research Laboratory, NRL/MR/7230--04-8760, Washington, DC (2004).

[6] Makarau, A., R. Richter, D. Schläpfer, and P. Reinartz, "Combined Haze and Cirrus Removal for Multispectral Imagery," IEEE Geoscience and Remote Sensing Letters, 13, No. 3, 379-383 (2016).

[7] Gruninger, J., J. Lee, and R. L. Sundberg, "The application of convex cone analysis to hyperspectral and multispectral scenes," Proc. SPIE 4888, 188–198 (2003).