Teaching Atmospheric Correction Using a Spreadsheet

E. J. Milton

Abstract
There are several possible educational roles for spreadsheets in remote sensing. They may be used in cost-benefit analyses, in data processing, or in simulation models designed to demonstrate simple physical processes. This paper describes a spreadsheet implementation of the Chavez (1988) "Modified Dark-Object Subtraction" procedure for the first-order correction of path radiance effects on remotely sensed images. The model has been found to be useful in teaching atmospheric correction procedures and also in stimulating discussion among remote sensing students from different backgrounds.

Introduction
Most remote sensing educators are familiar with the challenge presented by a mixed class comprising students primarily interested in physics, mathematics, and engineering and those interested in image processing or environmental applications. Nowhere is this contrast in approach and discipline-related bias more apparent than in attitudes towards atmospheric correction procedures. The first group of students relish fine-tuning and tinkering with radiative transfer models while the second group often either ignore the problem or are uncritical in their adoption of whatever correction procedure happens to be offered by the image processing system they are using.

It is my belief that better communication between the two groups is best achieved through student-centered small group activities which involve members from each discipline area working together. This note describes a computer-based learning tool which may be used in small groups and which I have found to be useful in stimulating discussion and motivating students on the subject of atmospheric correction of remotely sensed images.

The paper describes a spreadsheet implementation of the modified dark-object subtraction method described by Chavez (1988; 1989). This is an image-based procedure designed to provide a first-order correction for the presence of haze in a scene. The method is an improvement on earlier dark-object subtraction procedures because it includes a realistic model of the way in which atmospheric scattering varies with wavelength. The Chavez method is clearly described with examples in the papers referred to above, and that exercise will not be repeated here. Instead, attention will focus on the use of a spreadsheet as a delivery system for the procedure.

The primary advantage of the spreadsheet implementation described here is that this approach is interactive. Students can modify any of the input parameters of the model and immediately see the effect of these changes on the intermediate calculations and the final corrected digital number (DN) values. They can question the assumptions of the model, refine it, and test it by applying the results to image data. Other advantages of a spreadsheet implementation include the excellent graphical output available from many modern programs, the good facilities for "error-trapping" of input data, and the convenience offered by a graphical user interface.

The Spreadsheet Model
A spreadsheet model was constructed in the form of a self-documenting worksheet using the program Quattro Pro. The spreadsheet comprises a series of screens, each dealing with one aspect of the subject of atmospheric correction using the Chavez procedure. The first screen is largely text, and introduces the subject, referring to associated practical exercises and recommended reading. The next five screens take the student through the logical steps in the correction procedure itself. The first provides boxes for the student to enter initial values of "starting haze DN (SHV)," "SHV band," and the "relative scattering exponent." Note that these terms are fully explained in Chavez (1988).

The third screen is headed "Sensor Model" and invites the student to enter calibration data for the sensor used. The fourth and fifth screens present the atmospheric model and reveal the results of intermediate calculations, before leading on to screen six which gives the results of the procedure as DN-equivalent haze values for each band of the sensor used. The various steps in the procedure are described in text windows distributed throughout the worksheet. Screen six also prompts the user to press a function key to reveal a graphical display of the results, an example of which is shown in Figure 1. The final screen provides suggestions for further reading on the subject of atmospheric correction.

In functional terms, the spreadsheet can be thought of as two linked sub-models (Figure 2): an atmosphere model and a sensor model. Each sub-model is kept as simple as possible but also left open to elaboration to allow students to investigate them further.

Atmosphere Model
The atmosphere model used in the Chavez procedure is grossly simplified as the main purpose of the procedure is to effect a physically based cosmetic improvement to images rather than accurately determine surface reflectance, although this may be possible under some circumstances (Chavez, 1989). Atmospheric absorption is ignored and scat-

Department of Geography, University of Southampton, Southampton SO9 5NH, United Kingdom.

Photogrammetric Engineering & Remote Sensing,
Vol. 60, No. 6, June 1994, pp. 751–754.

0099-1112/94/6006–751$03.00/0

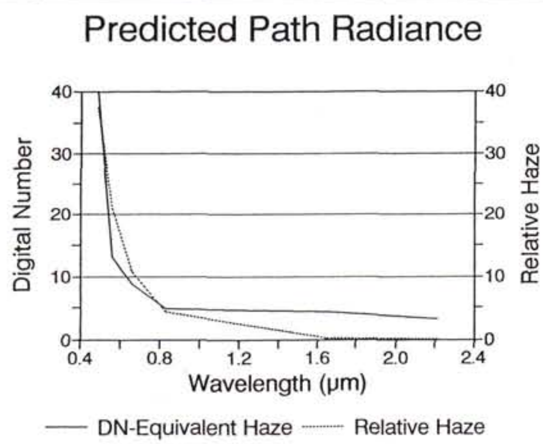

Figure 1. An example of graphical output from the spreadsheet for the case where the Starting Haze Value (SHV) was 40 in Landsat thematic mapper band 1 and the exponent was set at −4.0 (see text and Chavez (1988) for definition of terms).

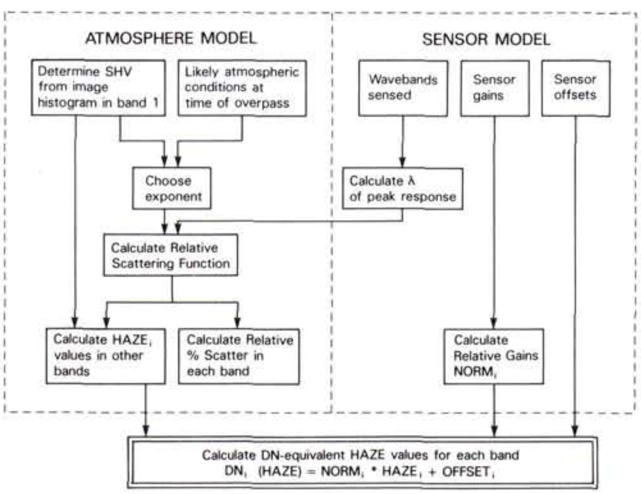

Figure 2. Flow diagram showing the logical structure of the spreadsheet implementation of the haze correction procedure described by Chavez (1988).

tering is represented by a function which varies according to the wavelength raised to a power between −0.5 and −4.0. It is left to the user to select a value for the exponent based on (a) the DN of the darkest pixels in the scene in a short-wavelength band and (b) knowledge of the likely atmospheric clarity at the time of the overpass. Chavez suggests the values listed in Table 1.

The Relative Scattering Model represents the interplay of Rayleigh and Mie scattering, with Mie scatter becoming more important in hazy atmospheres. Once again, the interactive nature of the spreadsheet is important, as it allows students

TABLE 1. SUGGESTED EXPONENT VALUES

Atmospheric Conditions	Relative Scattering Model
Very Clear	$\lambda^{-4.0}$
Clear	$\lambda^{-2.0}$
Moderate	$\lambda^{-1.0}$
Hazy	$\lambda^{-0.7}$
Very hazy	$\lambda^{-0.5}$

to experiment with different values of the exponent and to see the results graphically. Interest and discussion naturally focuses on the effect of changing the scattering model upon the corrected DN and not on the mathematical manipulation or programming steps needed. This leads students to wish to explore more elaborate atmospheric correction procedures and thence to discussion of the trade-offs involved between the greater realism of radiative transfer approaches and the minimal input data requirements of image-based methods.

Sensor Model

The sensor model performs a relative calibration of each band of the sensor based on gain and offset values entered by the user. Extracting these values from the published literature and dealing with conversions between the radiometric units involved is the first challenge facing the student. Relative gains are calculated, normalized to the band chosen by the user for the starting haze value. The sensor model also identifies the wavelength of peak sensitivity for each band, which in the present implementation is assumed to be the center wavelength.

Several discussion issues are raised by this simple sensor model, for example, the methods by which the gain and offset values are determined, the relative accuracy of pre-flight and in-flight calibration, the effect of sensor non-linearity, the importance of wavelength calibration, and the effect of changing the radiometric resolution. The spreadsheet allows the student to assess the significance of these factors through the manner in which they affect the corrected DN, and he/she is thereby introduced to the subject of sensitivity analysis. The suggestions for further reading lead the student into the burgeoning literature on absolute and relative calibration of remote sensing systems.

The combined output from the two sub-models is a list of estimated dark-object DNs for each band of the sensor. These must then be compared with the dark-object DNs which would have been predicted from use of the image histograms alone. The main premise of the Chavez procedure is that better estimates of true dark-object radiances are produced by use of a physically-based relative scattering model than relying on minima derived from image histograms alone. In practice, there is sometimes overcorrection of the image data, and this invites careful thought on the nature of dark-object subtraction procedures, and, in particular, on whether any radiance is likely to emanate from pixels in deep shadow. The iterative nature of the Chavez procedure makes it highly suitable as a starting point for discussing the nature, importance, and effectiveness of image-based atmospheric correction procedures in general.

Discussion

At one level, the material presented here might be regarded as trivial. The use of a spreadsheet does not improve the physical basis or applicability of the Chavez procedure by itself, nor does this paper demonstrate its application to a new or novel data set. Rather, this paper is an attempt to make more accessible an interesting and practical procedure which combines physically based scene/sensor understanding and image pre-processing as a precursor to qualitative image interpretation. The rationale for doing this is two-fold. First, a belief in the importance of dialogue between physically based scientists and practitioners of applied remote sensing at all levels. Second, a feeling that this dialogue is best encouraged by shared problem-solving activities such as those which a skilled educator may formulate around a device such as the spreadsheet described here.

A spreadsheet implementation of the Chavez procedure is particularly useful in an educational context for the following reasons:

- It introduces the Rayleigh Law of molecular scattering and shows how this must be modified for real atmospheres to allow for the effect of aerosols. Students can easily change the exponent in the relative scattering function and/or the dark object DN value and observe the effect on the predicted haze levels. They may easily test the results of the model against data from the literature and against images from different platforms and sensors.
- It demonstrates the importance of sensor parameters such as gain and offset values, and it provides a link to discussion of sensor calibration problems. Typically, the exercise would first be done with sensor gain = 1.0 and offset = 0, in which case the relative scattering function will vary smoothly with wavelength and show most clearly the effect of changes to the exponent value. Next, real-world sensor parameters can be introduced, to show how the DN-Equivalent Haze Value may vary with wavelength. The ability to produce graphical output from the spreadsheet allows students to see this in a visual form.
- It forms a bridge between physical principles and image processing. Students of remote sensing often have difficulty reconciling the precise and elegant mathematics of the basic physical laws with the less formal and more subjective procedures of image interpretation. A simplified treatment of the atmospheric effect such as that presented here helps to deal with this educational schizophrenia. It demonstrates that the physics of the real-world is complicated and uncertain, and it shows that digital images are quantitative data sources, not simply colorful pictures.

Broader Issues

There are several possible roles for spreadsheets in remote sensing education. First, they may be used as financial modeling tools to assess the costs and, possibly, the benefits of remote sensing techniques. For example, students could be provided with a spreadsheet which related the cost of a remote sensing survey to variables such as platform/sensor combination, areal coverage, and data processing methods. This would allow questions common in the real-world, such as "What if the flight line was extended by so many kilometres?" or "What if SPOT HRV were substituted for Landsat TM?" to be addressed. In this case, the educational benefit comes partly from students thinking about the relative costs predicted by the model, but perhaps more importantly, it comes from them thinking about the basis of the model, challenging specific assumptions, and making changes to the model.

Second, spreadsheets may be used as data processing tools, to complement or even replace specially written software. For example, data from a multiband radiometer may be processed from raw voltages to percentage reflectance. In this case, the advantage of a spreadsheet is that it is quick to program, easy to debug, and often has sophisticated data editing and error trapping procedures built-in.

Third, spreadsheets may be used to simulate the outcome of processes under different input conditions, such as the implementation of the Chavez procedure described here. Once the model has been established and verified, the user may change one or more of the input variables and observe the effect on the output and on any intermediate stages of processing. This role offers perhaps the most exciting educational possibilities because it involves students in active learning and focuses attention upon the performance of a model under different conditions rather than on the programming of the initial model.

Of course, there are dangers, the greatest of which is students not understanding the assumptions or limitations of the model provided and drawing false conclusions because of this, but these can be overcome by using the computer to complement rather than replace more conventional teaching methods. In particular, students must be encouraged to test the outcome of the model against real-world data (Shepherd, 1985). In the present context, this may be achieved by using the Chavez procedure to calculate DN-equivalent haze values for a scene and then assessing the effectiveness of the atmospheric correction. This may be done in several ways: a physics student may prefer to compare the results with those derived from a radiative transfer approach, while a geology student may wish to study the effect of the correction upon the color balance and interpretability of image data for example.

Conclusion

Simple interactive learning tools such as the spreadsheet model described here have a role to play in remote sensing education at all levels. They can motivate students to think about the assumptions inherent in a particular procedure, to evaluate results produced, and to explore what happens when key parameters are changed. Spreadsheets are particularly useful delivery systems for such models because they are interactive, intuitive, and extendable. Furthermore, they are widely used in all quantitative disciplines, which means that students from varied backgrounds can easily relate to them and focus on the purpose and assumptions of the model rather than on the programming structure and language used. The particular implementation described here has been found to be useful in giving students both an increased understanding of atmospheric correction procedures and an increased understanding of the perspective of remote sensing scientists from different backgrounds. Finally, it also is a convenient and practical method of providing a first-order haze correction to remotely sensed images.

If you would like a copy of the spreadsheet model for use with the Quattro Pro spreadsheet (DOS version), please send a blank disk for the IBM PC (3½- or 5¼-inch) to the author at the address heading the article.

References

Chavez, P. S., 1989. Radiometric calibration of Landsat thematic mapper multispectral images. *Photogrammetric Engineering & Remote Sensing*, 55:1285–1294.

———, 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24:459–479.

Shepherd, I. D. H., 1985. Teaching geography with the computer. *Journal of Geography in Higher Education*, 9:3–23.

(Received 11 August 1992; accepted 27 January 1993; revised 2 March 1993)

Edward Milton
Edward Milton has a first degree in geography from the University of Exeter and a PhD in remote sensing from the University of Reading. He has held lectureships at the Universities of Reading and Durham, and at the University of Southampton where he has been since 1981. His research interests include field spectroscopy and the application of IT to education and training in remote sensing.