DEEP LEARNING OF RADIATIVE ATMOSPHERIC TRANSFER WITH AN AUTOENCODER

A. Abigail Basener
Virginia Military Institute,
Student in Applied Mathematics
Lexington, VA
basenerag24@mail.vmi.edu

B. Bill Basener
University of Virginia,
School of Data Science
Charlottesville, VA
wb8by@virginia.edu

ABSTRACT
As electro-optical energy from the sun propagates through the atmosphere it is affected by radiative transfer effects including absorption, emission, and scattering. Modeling these affects is essential for scientific remote sensing measurements of the earth and atmosphere. For example, hyperspectral imagery is a form of digital imagery collected with many, often hundreds, of wavelengths of light in pixel. The amount of light measured at the sensor is the result of emitted sunlight, atmospheric radiative transfer, and the reflectance off the materials on the ground, all of which vary per wavelength resulting from multiple physical phenomena. Therefore measurements of the ground spectra or atmospheric constituents require separating these different contributions per wavelength. In this paper, we create an autoencoder similar to denoising autoencoders treating the atmospheric effects as 'noise' and ground reflectance as truth per spectrum. We generate hundreds of thousands of training samples by taking random samples of spectra from laboratory measurements and adding atmospheric affects using physics-based modelling via MODTRAN (http://modtran.spectral.com/modtran_home) by varying atmospheric inputs. This process ideally could create an autoencoder that would separate atmospheric effects and ground reflectance in hyperspectral imagery, a process known as atmospheric compensation or atmospheric correction. Our method uses an autoencoder [1] similar to a denoising autoencoder, treating the atmosphere as noise and ground reflectance as the image.

An autoencoder is a deep learning neural network that passes data through a series of layers that decrease in size leading to a dimensionally smaller representation of the data (encoding), and then passing the data through a symmetric series of layers back to the original data shape (decoding). When trained well, the process learns a representation of the data in the reduced dimensional space so that the encoding stage removes noise and the decoding stage recovers the data in the original space [2, 3, 4, 1]. A diagram of the architecture of our autoencoder is shown in Figure 1.

This method to convert a hyperspectral image from at-sensor radiance (the amount of light per wavelength measured at the sensor) to ground reflectance (the percent reflectance per wavelength), a process known as atmospheric compensation or atmospheric correction. While the accuracy of our method is not as good as other methods in the field, this an important first step in applying the growing field of deep learning of physical principles to atmospheric compensation in hyperspectral imagery and remote sensing.

Index Terms— Machine Learning, Deep Learning, Autoencoder, Radiative Transfer, Atmospheric Compensation

1. INTRODUCTION

In this paper we present a method for deep learning of atmospheric radiative transfer from a few observed spectra from a ground-based or overhead spectra. This method could be use to measure atmospheric constituents, but we focus on using
While our deep learning autoencoder method does not outperform a naive regression method based on the universal mean principle in QUACC [5], it is an important first step in applying the growing field of deep learning of physical principles to atmospheric compensation in hyperspectral imagery and remote sensing. We expect, based on the trajectories of other efforts in deep learning of physical phenomena, that better inclusion of physical principles in the architecture of the autoencoder would substantially improve the quality of output. For example, the inclusion of skip connections in the ResNet Network enables layers that learn functions closer to the identity, leading to the ability to train much larger, more complex, networks [6] with a smoother loss function that has fewer local minima [7]. All of our data and methods are provided open access\(^1\).

A hyperspectral image is a digital image in which each pixel has more than the usual three color (red, green, blue) bands, but often hundreds of bands across wavelengths sufficient to get spectral information about the materials in each pixel. We focus on hyperspectral images that have bands wavelengths from about 400nm to 2500nm - for comparison the visual colors occur around 450 (blue), 550nm (green) and 650nm (red) - and our spectra have 452 wavelengths. For a hyperspectral images collected at these wavelengths, the measured light at the sensor is from sunlight, having passed through the atmosphere, reflected off materials, and passed again through some amount of atmosphere (which may be small for a ground-based space or significant if the sensor is on board an aircraft or satellite.) Each band is typically 2nm to 10nm wide, and the bands are contiguous across the wavelength range. The reflectance for a material is important because it is the result of the interaction of photons at different wavelengths and the resonant frequencies of molecular bonds (for the wavelengths above the visible range) and the interaction of photons and electrons moving between quantum states. Specifically, important information about the constituents and bonds present in a material can be computed from reflectance spectra, for example distinguishing between different polymers, or distinguishing talcum powder from powdered sugar from dangerous white powdered substances.

The percentage of light that passes through the atmosphere (as opposed to being absorbed) is called transmittance, and varies depending on the wavelengths for each band. A plot of transmittance for two different atmospheric models across our wavelength range with a spectral resolution (band width) of 5nm is shown in Figure 2. Part of our goal is to determine this transmittance amount from spectra on the ground measured from a sensor even when the reflectance of the ground material is unknown.

The plot in Figure 3 shows more components involved in the radiative transfer model The Direct Solar (100km) gives the amount of sunlight measured at the top of the atmosphere.

The main shape of this curve is from the blackbody radiation given the temperature of the sun. The Direct Solar (0km) is the amount of sunlight reaching the ground, which is the Direct Solar (100km) times the transmittance shown in Figure 2. The Downward Diffuse (0km) is the amount of light per wavelength that reaches the ground after scattering in the atmosphere; this is the indirect illumination on an object that is in a shadow from the sun. The upward Diffuse is the amount of upward light, which at 100mk (top of atmosphere) is from atmospheric scattering and at 0km is from the blackbody radiation of the ground (which is insignificant given the wavelength range and ground temperature assumed in this model).

The at-sensor radiance is the upward Diffuse (at the elevation of the sensor) plus the total illumination (Direct Solar + Downward Diffuse) times the percent reflectance per band of the material on the ground per wavelength \(\lambda\),

\[
UD_\lambda + (DS_\lambda + DD_\lambda) Ref_\lambda = Rad_\lambda.
\]

There are also nonlinear effects as well. Light that has passed through a leaf and reflected off the ground would have the leaf transmission times ground reflectance in place of \(Ref_\lambda\) in this equation. In the lower wavelengths, especially blue and below, photons will take multiple bounces/scattering (collectively, 'haze' in the image) in the atmosphere (which you can observe in the higher values for Upward/Downward Diffuse in Figure 3), but the amount of haze is highly dependent on constituents and length of pass through the atmosphere; moreover has causes photons reflecting off material at one

\(^1\)https://www.kaggle.com/code/billbasener/autoencoder-atmospheric-compensation/notebook
location to scatter in the atmosphere and enter the sensor at locations/directions for other pixels, causing in "Upward Diffuse" that varies across the image comprised of a nonlinear mixture of nearby ground material spectra and atmospherics, rather than the ideal Upward Diffuse from the atmosphere alone shown in figure 3.

All of these values change with respect to atmospheric constituents, water vapor, CO2, Ozone, CH4, aerosols, sun angle, fraction of sun and sky visible to each pixel (shadow from objects, terrain, clouds, etc.), sensor angle, angle and roughness of the ground material, and other factors. The MODTRAN software (http://modtran.spectral.com/modtran_home) can simulate these effects if they are known, and provide a modeled at-sensor radiance for reflectance spectra, or a ground reflectance spectrum for a measured at sensor radiance.

The purpose of hyperspectral imaging is to perform spectroscopy writ large; that is, to be able to determine the materials in each pixel from the reflectance spectra for those pixels. As such, good atmospheric compensation is an essential step. The most accurate methods are usually either based on physics based modelling with MODTRAN (MODerate resolution atmospheric TRANsmission, http://modtran.spectral.com/modtran_home) such as FLAASH [8] or using materials of known reflectance in the image, for example the empirical line method [9, 10], in which case it is usually preferable to have materials that are spectrally flat, for example a set of five panels which are 5%, 30% 50% 80% and 95% reflectance across all wavelengths, which can be used to estimate a best fit regression line per wavelength to convert from radiance units to reflectance. However, all of these methods have some approximations and attempt to measure physical parameters. For example, a good ELM will estimate the Upward diffuse as the intercept and the Direct Solar and Downward Diffuse as the slope, but assume these are consistent values for every pixel in the image. FLAASH attempts to estimate the physical parameters from the image, even estimating water vapor content per pixel, and use a physics based model to compute the grouped reflectance from each at-sensor radiance measurement.

A heuristic and approximate but surprisingly effective method for atmospheric correction is to make the assumption that the mean spectrum of a significantly large library of reflectance spectra will be consistent, and use this assumption to compute a single gain and offset that is applied per wavelength across the image. A method based on this assumption is QUACC (QUick Atmospheric Correction Code), in which a sample of 50 different spectra (called endmembers) are selected from the radiance image, usually iteratively so each new endmember is optimally different than the previous ones, and the mean of these 50 endmembers is determined. The least measured radiance value across the image per wavelength is assumed to be the upwelling radiance, and is thus subtracted from the image. Then the ratio between the ideal mean and mean of the endmembers (after subtracting estimated upwelling radiance) is computed and used as a ‘gain’ and is multiplied by every upwelling-subtracted radiance value to convert to reflectance. This is often implemented with additional heuristic improvements such as removing spectra of mud or vegetation from the spectra, is provided in QUACC [5]. There are other semi-heuristic methods for example SHARC [11]. The QUAC method is faster than physics-based methods require no manual input. In tests, is provides reflectance spectra that are ±15% in comparison to FLAASH generated reflectance spectra, and perhaps more importantly QUACC tends to generate reflectance spectra that retain the features of the true spectra, which is the most important factor for spectroscopy. A small to moderate variation in the magnitude of a spectrum is often not important. Materials are identified more from variation in reflectance at different bands which indicates which bands are absorbing vs. reflecting, resulting from the constituents and molecular bonds present. The total reflectance (i.e. total albedo) can vary with illumination amount, angle to sun, and surface roughness, inconsistent calibration, none of which are dependent on the molecular bonds and constituents of the material. The primary questions in atmospheric compensation methods is whether they retain the features of the true reflectance and whether they avoid creating new features not present in the material.

For this paper, we started with a set of about 5,000 reflectance spectra of known materials, each of which passed some basic quality check. We then took random samples of size 39, randomly selected a set of parameters for MODTRAN (solar zenith angle from 0-85 in increments of 5, random selection from the 6 possible atmospheric models, random selection

Fig. 3. Two plots of transmittance for two different atmospheric models created using MODTRAN.
from the 12 possible aerosol models) and created an associated set of 39 radiance spectra. For each set, we also computed the mean spectrum and added this as a 40th spectra. This creates arrays that have even dimension (number of spectra and number of bands) which are preferable for some Deep Learning methods. So our input data is a set of 40 radiance spectra each with 452 wavelengths, and our output data is a set of 40 reflectance spectra. In Figure 4 we show a plot of the 40 radiance spectra (top) and the associated reflectance spectra (bottom). We created a set of 10,000 such samples of 40 spectra.

![Figure 4](image)

**Fig. 4.** At-sensor radiance for 39 spectra and the mean plotted in black (top plot) along with the true reflectance for these materials and the mean of their reflectance in black (bottom plot). Below are the data arrays for these plots (40 spectra by 452 bands each) are shown below as images.

2. BASELINE REGRESSION METHOD AND RESULTS

We create a baseline regression method for atmospheric correction based on the assumption that the mean spectrum of a significantly large library of reflectance spectra will be consistent, called the reference spectrum. The is the same assumption used in the QUACC [5] method, although we are using the synthetically generated data which has different distributions for the types of spectra present. Our baseline is not an approximation to QUACC, but a baseline for comparison to the Deep Learning methods that is simple and based on an accepted heuristic approximation.

The mean spectra for a each of a random selection of ten of our sets of 40 spectra are shown plotted in Figure 5. This assumption is very approximately true, but the approximation seems more consistent using 120 per sample, shown below in this figure. A physics-based support for the general shape of the reference reflectance spectrum is provided in [5] as follows:

The general shape of the reference reflectance spectrum has a simple physical origin. The decrease toward the longwavelength edge arises because the molecular constituents of materials have relatively strong NIR vibrational absorption features that increase in strength with increasing wavelength. The decrease toward the short-wavelength edge arises because the molecular constituents have strong electronic absorption features that increase in strength with decreasing wavelength. Although we normalize the peak of this curve to unity for reasons discussed later, it is important to note that the peak average reflectance is $\sim 0.4$...

From the assumption that a sample of spectra will have a fixed mean reflectance $m$, we can simply take this mean divided the mean measured radiance spectra to obtain a ‘gain’ multiplication factor for each band. An approximate atmospheric compensation can be done by multiplying this gain times each measured radiance value, per band. The output from the baseline method is shown in Figure 6. Some of the
features in the true reflectance spectra column (third, right-hand, column) can be observed in the spectra predicted using the baseline method with the standard mean reflectance assumption (center column). There are regions in the predicted column where the spectra have values near zero - these are not errors, but regions where the radiance has near-zero values. Effectively, the water in the atmosphere has near-zero transmittance in these wavelength regions, and there is insufficient signal for prediction. These regions are removed from all hyperspectral images created with solar illumination. Wavelength regions around 0.9 and 1.2 microns are also often removed because some atmospheres will also have very low transmittance in these regions, and these regions can be observed particularly strong in the 4th row. Figure 7 shows these same spectra but with the water bands removed (replaced with a straight line in the plots) to aid in visually comparing spectra.

To measure the effectiveness of this method we randomly selected 500 sets of 39 spectra and computed the correlation between each true reflectance and predicted reflectance (after removing bands in $0 - 0.45, 1.3 - 1.5, 1.75 - 2.05$, and $2.4 - 3$ micron regions) and the mean correlation of these 19,500 spectra was 0.87 with a standard deviation of 0.29. Of these spectra, 69% of the predicted spectra were within 15% of the true spectrum. These results are not great, and not sufficient without addition modifications like those in QUACC, but this is sufficiently promising that a deep learning method may be viable.

3. AUTOENCODER METHOD AND RESULTS

Because the simple baseline regression method works reasonably, it can be expected with a proper architecture and proper training, a deep learning method could improve on accuracy. Perhaps the model could learn the baseline regression plus adjustments based on common types of variation in the atmosphere and illumination. Because the input and output have the same shape, we decided to try an autencoder network treating the atmosphere as noise (see [1]).

The input into our autencoder is the $40 \times 452$ array of 40 spectra in radiance units, each of which has 452 bands. The output is an array of the same shape, number of spectra, and bands, but in reflectance units. The architecture for the network is shown in Figure 1, constructed in Keras and trained for 50 epochs in batches of 256 each with an adam optimizer and binary_crossentropy loss. All activation functions are ReLU except the final decoder layer which is sigmoid, in which case the $0 - 1$ output of the sigmoid neurons match the $0 - 1$ range of values for reflectance.

The loss curve from optimizing the network is shown in Figure 8 (using 33% of the data for validation). Example output showing original radiance, predicted reflectance, and true reflectance are shown in Figure 9. The output predicted reflectance of the autoencoder along with true reflectance with the standard water bands removed are shown in Figure 10.
Fig. 8. The loss curve from training the autoencoder.

Table 1. Evaluation metrics for the baseline regression and autoencoder methods.

| METRIC               | REGRESSION RESULT | AUTOENCODER RESULT |
|----------------------|-------------------|--------------------|
| MEAN CORR            | 0.87              | 0.52               |
| STD CORR             | 0.29              | 0.44               |
| PERCENTILE IN ±15%   | 69%               | 34%                |

To measure the effectiveness of the autoencoder comparable to our baseline method, as before we randomly selected 500 sets of 39 spectra from the validation set and computed the correlation between each true reflectance and predicted reflectance (after removing bands in $0 - 0.45, 1.3 - 1.5, 1.75 - 2.05$, and $2.4 - 3$ micron regions) and the mean correlation of these 19,500 spectra was 0.52 with a standard deviation of 0.44. Of these spectra, 34% of the predicted spectra were within 15% of the true spectrum. These results are significantly worse than the baseline method.

4. CONCLUSIONS

It is clear that our autoencoder performed poorly in comparison to the baseline method, summarized in Table 1.

It seems likely that the architecture of CNNs is making the network unable to use the wavelength information. Specifically, the architecture we used follows a standard framework for images. In images, a combination of pixels forming a nose/ear/wheel/etc. is meaningful at any location in the image, and CNNs are able to leverage this information. But in spectroscopy, the meaning of a combination of values is dependent on the location of the feature; although a network might learn the shape of features in general and use both the shape and location in final layers. We tried a significant number of modifications to the architecture, including replacing some or all Conv layers with fully connected dense layers, modifying the amount of dropout, using different optimization methods and loss functions (for example, rmsprop in place of adam and/or mean squared error in place of binary crossentropy). We also tried building the network to use just the mean spectrum (to learn an approximation to the baseline regression concept), or building convolution layers with different rectangular windows. However, none of these methods provided benefit above the autoencoder provided in this section in any manner worth reporting. It does seem the architecture should learn a function for removing atmosphere in a way that the same function is applied to every row (spectrum) in the data array.

Our approach was to use an autoencoder and treat the atmosphere as noise since this is a well-developed method of application. But perhaps since the baseline regression method is a reasonable approximation, perhaps a regression neural network would be better, or some combination that leverages the known heuristics and physics together with a deep learning approach. We believe the physical explanation of the data and reasonable effectiveness of the baseline regression method strongly suggest that a Deep Learning approach to atmospheric compensation, perhaps incorporated with some heuristics and physics, has the likelihood of being completely automated and highly effective, with the potential to outperform currently available methods such as QUACC and FLAASH.

5. REFERENCES

[1] Dor Bank, Noam Koenigstein, and Raja Giryes, “Autoencoders,” arXiv preprint arXiv:2003.05991, 2020.
Fig. 10. Predicted reflectance using the autoencoder (left-hand column), and true reflectance (right-hand column). The water absorption regions have been removed (replaced with a straight line in plots) to aid in visual comparison of features. Observe that, especially in the 0.45 to 1.25 micron region, the autoencoder is introducing features that are not present in the true reflectance spectra.

[2] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol, “Extracting and composing robust features with denoising autoencoders,” in Proceedings of the 25th international conference on Machine learning, 2008, pp. 1096–1103.

[3] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, Pierre-Antoine Manzagol, and Léon Bottou, “Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion,” Journal of machine learning research, vol. 11, no. 12, 2010.

[4] Cai Ruikai, “Research progress in image denoising algorithms based on deep learning,” in Journal of Physics: Conference Series. IOP Publishing, 2019, vol. 1345, p. 042055.

[5] Lawrence S. Bernstein, Xuemin Jin, Brian Gregor, and Steven M. Adler-Golden, “Quick atmospheric correction code: algorithm description and recent upgrades,” Optical Engineering, vol. 51, no. 11, pp. 1 – 12, 2012.

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[7] Hao Li, Zheng Xu, Gavin Taylor, Christoph Studer, and Tom Goldstein, “Visualizing the loss landscape of neural nets,” Advances in neural information processing systems, vol. 31, 2018.

[8] Timothy Perkins, Steven M. Adler-Golden, Michael W. Matthew, Alexander Berk, Lawrence S. Bernstein, Janine Lee, and Marsha Fox, “Speed and accuracy improvements in FLAASH atmospheric correction of hyperspectral imagery,” Optical Engineering, vol. 51, no. 11, pp. 1 – 8, 2012.

[9] Baabak G Mamaghani, Geoffrey V Sasaki, Ryan J Connal, Kevin Kha, Jackson S Knappen, Ryan A Hartzell, Evan D Marcellus, Timothy D Bauch, Nina G Raqueño, and Carl Salvaggio, “An initial exploration of vicarious and in-scene calibration techniques for small unmanned aircraft systems,” in Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping III. International Society for Optics and Photonics, 2018, vol. 10664, p. 1066406.

[10] Bo-Cai Gao, Marcos J Montes, Curtiss O Davis, and Alexander FH Goetz, “Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean,” Remote sensing of environment, vol. 113, pp. S17–S24, 2009.

[11] Leonid V. Katkovsky, Anton O. Martenov, Volha A. Siliuk, and Dimitry A. Ivanov, “SHARC method for fast atmospheric correction of hyperspectral data,” in Remote Sensing of Clouds and the Atmosphere XXIII, Adolfo Comeron, Evgeni I. Kassianov, Klaus Schäfer, Richard H. Picard, and Konradin Weber, Eds. International Society for Optics and Photonics, 2018, vol. 10786, pp. 53 – 65, SPIE.