

# AO08: Estimating surface reflectance using a neural network

Nikita Ostrovsky  
Somerville College, Oxford  
April 2022

*Supervised by Dr. Adam Povey and Prof. Don Grainger*

## Abstract

The Oxford-Rutherford Appleton Laboratory Retrieval of Aerosol and Cloud (ORAC) algorithm uses top-of-atmosphere (TOA) radiance data from a primary satellite-based radiometer and the Moderate Resolution Imaging Spectrometer (MODIS) to constrain land surface signal and estimate atmospheric aerosol properties. However, the spectral responses of the radiometers differ; it is shown that this leads to a difference in measured surface albedo of  $\sim 1\%$  between the instruments, which contributes to uncertainty in the aerosol retrieval. Two methods for reconstructing the true land surface contribution to the primary spectrometer's measurement are trialled – one involving singular value decomposition (SVD) analysis, the other using a multi-layer perceptron (MLP) neural network. It is found that the MLP reconstruction error is reduced by 61% when trained using a Limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) algorithm instead of the canonical Adam optimiser, but that reconstruction using SVD analysis outperforms the best-case neural network results by 53%. Despite this, impact of the SVD reconstruction on ORAC retrieval appears to be negligible.

## 1 Introduction

Atmospheric aerosols (suspended particles) play an important role in regulating the Earth's weather and climate: directly, by interacting with solar radiation and indirectly, by causing or altering processes (notably, clouds) that directly interact with solar radiation and the climate system as a whole. They are among the most significant and most uncertain drivers of anthropogenic climate change after greenhouse gas emissions [1]. Consequently, the measurement and modelling of aerosol properties and distribution is vital in understanding atmospheric dynamics and accurately predicting the weather and climate [2].

Aerosol properties can be accurately determined using land-based measurements, however these are limited in scope: global aerosol measurement from land-based instruments is highly impractical. Remote sensing enables more versatile, but less accurate, aerosol observation. Satellites measure top-of-atmosphere (TOA) reflected radiance in discrete spectral channels – typically at wavelengths where the atmosphere is relatively transparent, such that measurements of the Earth's surface are possible. The measured radiance in the  $k^{th}$  instrumental chan-

nel is

$$R_k = \int I(\lambda) \psi_k(\lambda) d\lambda \quad (1)$$

where  $I$  is the intensity incident on the spectrometer as a function of wavelength  $\lambda$  and  $\int d\lambda \psi_k(\lambda) = 1$  is the normalised spectral response function (SRF) of the  $k^{th}$  channel. These measurements can be used to infer atmospheric properties including aerosol optical depth (AOD),

$$\tau(\lambda) = \ln \frac{\Phi_i}{\Phi_t} \quad (2)$$

a proxy for the physical quantity of aerosols present in the atmosphere, where  $\Phi_i$  and  $\Phi_t$  describe the radiant flux incident on, and transmitted by, the atmospheric column; it is wavelength dependent through the spectral variation of flux scattering and absorption.

Oxford-Rutherford Appleton Laboratory Retrieval of Aerosol and Cloud (ORAC) is an optimal estimation retrieval algorithm using satellite measurements and *a priori* information on the state of the surface and atmosphere to determine the most probable characteristics of both. For retrievals over land, ORAC uses the Moderate Resolution Imaging Spectrometer (MODIS) MCD43C1 land data product as an *a priori* for the state of the land surface,

and measurements from the primary satellite – either the Advanced Along-Track Satellite Radiometer (AATSR) or Sea and Land Surface Temperature Radiometer (SLSTR) – are then used to determine aerosol characteristics.

The spectral response of MODIS differs from either of the primary radiometers, leading to a discrepancy between the MODIS land surface *a priori* and the land signal in the primary radiometer’s TOA measurement, which introduces uncertainty into the aerosol retrieval. Aerosol retrievals over land are generally less accurate than those over sea [3] due to the spectral inhomogeneity of land surfaces. Currently, no adjustment for this discrepancy in spectral responses is implemented in ORAC, thus a heuristic for the project aim is: given a land surface measurement by MODIS, how would the same land surface appear if measured by the primary radiometer?

Designing such a reconstruction algorithm is difficult, largely because the required dataset (a large number of identical surfaces measured by the instruments of interest) is almost impossible to obtain. The rapid variability of atmospheric and land conditions means that two satellites never measure an identical land surface. Sayer et al. [4] instead simulated the measurement of land surfaces from different spectral instruments by numerically evaluating Eq. 1 for a variety of high-resolution land surface spectra. They used SVD analysis to identify dominant correlations in the different instruments’ simulated measurements, and, using these correlations, reconstructed the primary instrument’s true land measurement from the MODIS measurement, reducing the error in aerosol retrievals.

Simulated measurements (henceforth: ‘measurements’) of land surface albedo by the primary satellite (AATSR or SLSTR) differ from MODIS measurements on the order of 0.01, Fig. 1. Albedo has a maximum of 1, so differences in instrumental spectral response account for an error of several percent in land surface measurements. The magnitude of the error varies depending on primary instrument and measurement channel. For instance, in the 0.86  $\mu\text{m}$  channel, measurements differ less between MODIS and AATSR than MODIS and SLSTR, as seen in the steep gradient in Fig. 1. This is explained by examining Fig. 2: the SLSTR SRF does not fully overlap with the MODIS SRF in the 0.86  $\mu\text{m}$  band, leading to differences in the spectral features captured. The converse is true at 1.6  $\mu\text{m}$ , so AATSR albedo measurements differ more from MODIS in

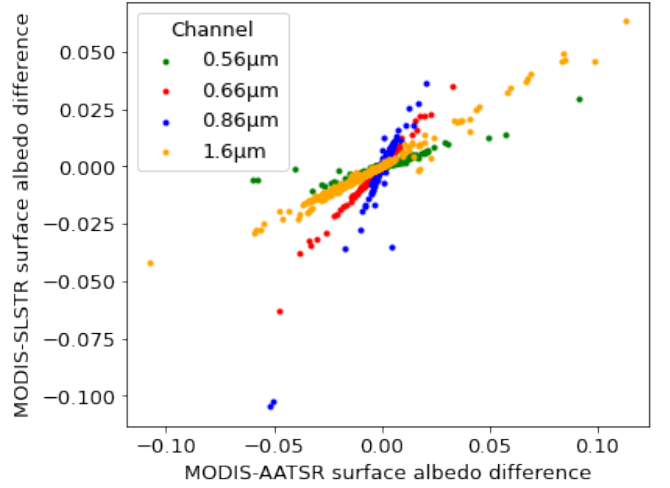


Fig. 1: Surface albedo measurements by primary satellite differ from MODIS measurements

this spectral region.

As Fig. 1 shows, simulated measurements (henceforth: ‘measurements’) of land surface albedo by MODIS differ from measurements by the primary satellite – AATSR and SLSTR – on the order of 0.01. Albedo, a measurement of diffuse reflectance of solar radiation, has a maximum of 1, so differences in instrumental spectral response account for an error of several percent in land surface measurements. The magnitude of the error varies depending on primary instrument and measurement channel. For instance, in the 0.86  $\mu\text{m}$  channel, AATSR measurements differ less from MODIS than SLSTR measurements, as seen in the steep gradient in Fig. 1. This is explained by examining Fig. 2: the SLSTR SRF does not fully overlap with the MODIS SRF in the 0.86  $\mu\text{m}$  band, leading to differences in the spectral features captured. The converse is true at 1.6  $\mu\text{m}$ , and correspondingly the AATSR albedo measurements differ more from MODIS in this spectral region.

This work seeks to build on Sayer et al., using updated and enlarged spectral databases and novel analytical methods. An outstanding concern for the authors of the 2012 paper was the presence of gaps in the surface spectral data used (Dr. A. Sayer, personal communication, October 20, 2021); this is explicitly addressed in Section 2.2.

Section 2 describes the instrumental and surface-spectral data used, and the interpolation used to prepare data for analysis; Section 3 details the SVD and neural network methods used in reconstructing the land measurements; Section 4 examines the accuracy of the reconstructions and their impact on

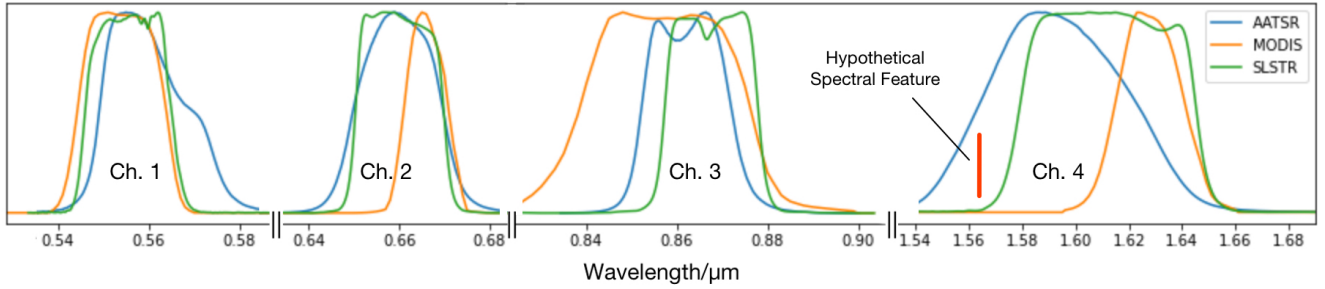


Fig. 2: Instrumental spectral response functions; a spectral feature could be measured by AATSR ch. 4, but missed by MODIS

ORAC retrieval; and an assessment of the work and suggestions for further development are made in Section 5.

## 2 Datasets

### 2.1 Spectral Response Functions

In order to simulate the measurement of land surfaces, spectral response functions for each of the relevant instruments were required. Where there were multiple instances of an instrument (MODIS AQUA and TERRA, SLSTR-A and -B) the data from the more stable version instrument was used. MODIS TERRA and SLSTR-A, having historically experienced hardware issues [5], were foregone in favour of AQUA and SLSTR-B, respectively.

Spectral radiometers measure radiance in a large number of spectral bands. Aerosols have diameters  $\lesssim 1 \mu\text{m}$ , so bands centered at  $0.56 \mu\text{m}$ ,  $0.66 \mu\text{m}$ ,  $0.86 \mu\text{m}$  and  $1.6 \mu\text{m}$ , in which aerosols are visible, are used in ORAC retrieval. They are designated channels 1, 2, 3 and 4, respectively (cf. Fig. 2).

As seen in Fig. 2, there is significant variation between instrumental spectral responses, meaning that a land surface with a spectral feature at  $1.56 \mu\text{m}$ , for instance, would be measured as having a different brightness in channel 4 of AATSR compared to MODIS. Information about the spectral responses of the instruments alone is therefore insufficient to determine how surface measurements vary between instruments: it is necessary to examine correlations between surface measurements in different channels, the assumption being that there are correlations within land surface spectra themselves, which would enable prediction of measurements between channels.

### 2.2 Surface Spectra

Land surface spectra were required to represent the land surfaces signal typically present in aerosol retrieval. ADAM, ECOSTRESS and USGS [6] datasets were considered. The latter was selected for its breadth (comprising over 1,800 entries), high resolution and variety, with surface spectra in seven categories. This database is significantly expanded compared to the spectral databases used Sayer et al. [6].

From this database, three categories (minerals, coatings, compounds) were immediately excluded as they were considered poor representatives of the surfaces encountered in satellite measurements. The remaining four categories (artificial, liquids, vegetation and soils) comprised 766 spectra. The USGS data used were captured in field conditions, meaning that the surface spectra are modulated by the solar spectrum. However, given that all land surfaces measured by satellite radiometers are subject to the same solar modulation, this was assumed to be unproblematic.

The large dataset contained multiple variants of extremely similar surfaces (e.g. 11 spectra of lodgepole pine.) The performance of SVD analysis and neural networks generally depends sensitively on the datasets used to train them, so this prompted concerns about designing a reconstruction algorithm that did not adequately represent the Earth's surface. A 125-spectrum subset of the large dataset was therefore curated as a more accurate representation of the Earth's surface. Where there were multiple spectra of a single relevant surface, one was chosen at random under the assumption that variation between spectra was negligible, thereby eliminating redundancy. Surfaces were selected to represent the dominant types of Earth's land cover [7]: 30% of the selected spectra were tree-covered regions, 22%

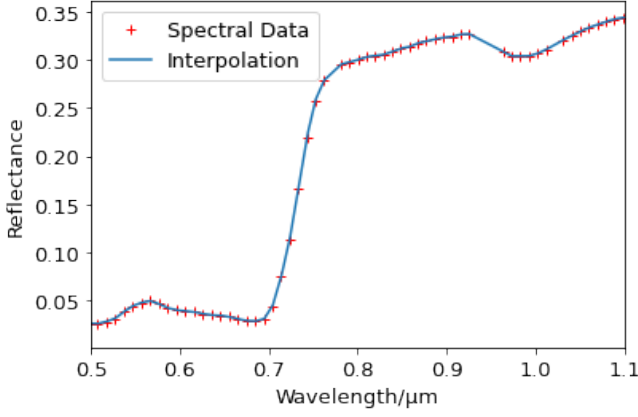


Fig. 3: Interpolation of Aspen spectrum

grassland, 17% terrestrial barren land, 16% shrubs, 11% herbaceous crops, 3% inland water, 1% artificial surfaces – as determined from the metadata accompanying the surface spectra.

The large and curated datasets were both used independently in the reconstruction process, to assess the relative importance of size and diversity within the training dataset. During algorithm training (i.e. neural network training and SVD analysis), one third of the selected dataset was reserved for validation.

### 2.3 Data Preparation

To facilitate the numerical convolution of the SRFs with surface spectra (Eq. 1), all datasets were interpolated onto a grid-spacing matching the highest resolution dataset (the SLSTR SRF). A simple linear interpolation procedure was deemed sufficient as the surface spectra were smoothly varying, e.g Fig. 3.

However, some surface spectra were found to have gaps, prompting concerns about interpolation from insufficient data. This was resolved by examining the spectral position of the gaps relative to the instrumental bands: it was found that, in all cases, the instrument bands fell outside of gaps in surface spectral (Fig. 4), meaning that the absence of data did not affect numerical evaluation of Eq. 1, nor, by extension, conclusions about differences between instrumental measurements of land surfaces.

The reason for this fortuitous coincidence is likely that the gaps in spectral data were caused by atmospheric absorption: gaps in the data appeared only for surfaces that were measured from the air [6], e.g. manzanita – a type of tree. Consequently, data was recorded only in those spectral regions where the

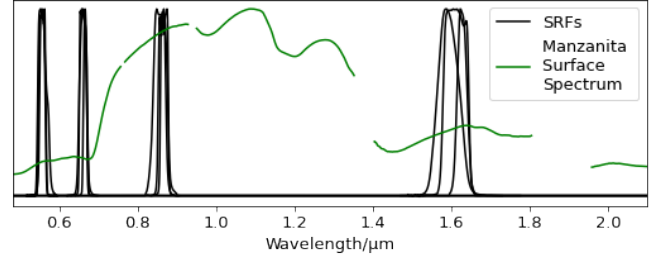


Fig. 4: Gaps in surface spectra fall outside spectral responses of instruments

atmosphere was significantly transparent, which, as noted above, matches the wavelength bands used by satellite-based spectrometers.

With all data interpolated onto a homogeneous grid spacing, Eq. 1 could be evaluated numerically, yielding a  $12 \times N$  array containing 12 simulated radiance measurements (3 instruments with 4 channels each) of  $N$  surfaces.

For neural network and SVD analysis, the array was randomly segmented into a training and testing portion: the former was used to train the neural network and generate the transformation matrix using SVD analysis, the latter to validate the reconstruction method.

## 3 Methods

### 3.1 SVD Analysis

Singular value decomposition (SVD) is a computationally-light method for extracting the dominant correlations in a set of data, which is useful for our purposes in characterising instrumental responses to land surfaces.

A real  $m$ -by- $n$  matrix  $\mathbf{M}$  can be factorised as

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (3)$$

where  $\mathbf{\Sigma}$  is a real  $m$ -by- $n$  diagonal matrix, and  $\mathbf{U}$  and  $\mathbf{V}$  are orthogonal  $m$ -by- $m$  and  $n$ -by- $n$  matrices, respectively.  $\mathbf{U}$  is the eigendecomposition of the  $m$ -by- $m$  matrix

$$\begin{aligned} \mathbf{M}\mathbf{M}^T &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T \\ &= \mathbf{U}(\mathbf{\Sigma}\mathbf{\Sigma}^T)\mathbf{U}^T \end{aligned} \quad (4)$$

The elements of  $\mathbf{M}\mathbf{M}^T$  are the dot-products of the rows of  $\mathbf{M}$ , hence our  $12$ -by- $12$  correlation matrix  $\mathbf{M}\mathbf{M}^T$  describes correlations between the 12 instrumental channels across all of the surfaces measured. The eigenvectors  $\mathbf{u}_i$  of  $\mathbf{M}\mathbf{M}^T$ , which form the columns of  $\mathbf{U}$ , provide a basis which describes the

correlations between the rows of  $\mathbf{M}$ . Each eigenvalue  $\sigma_i$ , contained in the diagonal matrix  $\mathbf{\Sigma}\mathbf{\Sigma}^T$ , describes the proportion of the variance in  $\mathbf{M}$  captured by  $\mathbf{u}_i$ . Hence the eigenvector with the largest eigenvalue captures most of the variance in  $\mathbf{M}$ . In this way, the simultaneous decomposition of measurements in the instrumental channels enables characterisation of their responses to land surfaces.

The Python library EOFs was used to extract the singular vectors for this matrix. In general, all of the left singular vectors (contained in  $\mathbf{U}$ ) are needed to provide a complete basis for the matrix  $\mathbf{M}$ . However, since the input space is four-dimensional (spanning the four MODIS channels), the first four singular vectors form a complete basis for the space.

Projecting the MODIS measurement onto the first four components of the singular vector, the final eight components uniquely specify the expected AATSR/SLSTR measurement. The measurement 12-vector  $\mathbf{m}$ , describing a land surface measurement by each of the three satellites, may be expressed in the dominant correlation basis  $\mathbf{u}_i$

$$\mathbf{m} = \sum_{i=0}^4 c_i \mathbf{u}_i \quad (5)$$

where  $c_i$  are basis transformation coefficients to be found. These four unknowns are specified by the four MODIS elements of the measurement vector,  $\mathbf{m}_{(m)}$

$$\mathbf{m}_{(m)} = \sum_{i=0}^4 c_i \mathbf{u}_{(m),i} \quad (6)$$

more conveniently expressed in matrix notation

$$\mathbf{m}_{(m)} = \mathbf{U}_{(m)} \mathbf{c} \quad (7)$$

where  $\mathbf{U}_{(m)}$  are the MODIS components of the left singular vectors. This yields the coefficients

$$\mathbf{c} = \mathbf{U}_{(m)}^{-1} \mathbf{m} \quad (8)$$

which immediately specify the predicted AATSR and SLSTR measurements,  $\mathbf{m}_{(a,s)}$  as

$$\begin{aligned} \mathbf{m}_{(a,s)} &= \mathbf{U}_{(a,s)} \mathbf{c} \\ &= \mathbf{U}_{(a,s)} \mathbf{U}_{(m)}^{-1} \mathbf{m}_{(m)} \\ &= \mathbf{T} \mathbf{m}_{(m)} \end{aligned} \quad (9)$$

where  $\mathbf{T} = \mathbf{U}_{a,s} \mathbf{U}_m^{-1}$  encodes the transformation from MODIS-space to AATSR/SLSTR-space: multiplying the MODIS measurement  $\mathbf{m}_{(m)}$  by  $\mathbf{T}$  yields the expected results when the primary satellite spectrometer measures the same surface. This matrix was incorporated into the ORAC algorithm by Dr.

Povey, enabling assessment of its impact on retrieval, detailed in Section 4.4.

### 3.2 Neural Networks

A multilayer perceptron (MLP) – a simple type of neural network – was used as an alternative method for reconstructing primary instrument surface signals from MODIS measurements.

The MLP was built with a single hidden layer, Fig. 5, using Scikit-learn. Each node in the network takes at least one input and returns an output determined by the activation function. The network's reconstruction of a surface measurement in the  $k^{th}$  primary-radiometer channel is defined by the function

$$F_k(\mathbf{x}) = \sum_i^N \phi(\mathbf{w}_{i,k} \cdot \mathbf{x} + b_{i,k}) \quad (10)$$

where the summation is over the  $N$  hidden nodes,  $\mathbf{x}$  is the MODIS measurement four-vector,  $\phi(x)$  is the activation function and  $\{\mathbf{w}_{i,k}, b_{i,k}\}$  are parameters to be found, such that the predictions of the network are optimised.

The behaviour of the neural network is determined by the weights and biases on each input,  $\{\mathbf{w}_{i,k}, b_{i,k}\}$ . During training, a cost function quantifying the error in the neural network's predictions for the entire training dataset is minimised with respect to the node weights: in our case, an 800-parameter optimisation problem. This is done iteratively, with each evaluation of the cost function and adjustment of the network weights termed an 'epoch', until the cost function converges at a local minimum of the weight-space.

Due to the non-linearity introduced by the activation function, the action of the MLP is more general than that of the linear SVD-based reconstruction. In fact, the universal approximation theorem [8] states that Eq. 10 can approximate any continuous function arbitrarily well, which implied that a single-layer MLP was sufficient. It was hoped that the MLP could provide a higher-precision reconstruction than the SVD method while using an identical dataset.

The challenge, of course, is that the function being approximated – which maps MODIS surface measurements to primary instrument surface signal – is not known, so the predictive power of the network is highly dependent on the size of the training dataset available. Moreover, the universal approximation theorem places no constraints on the size of

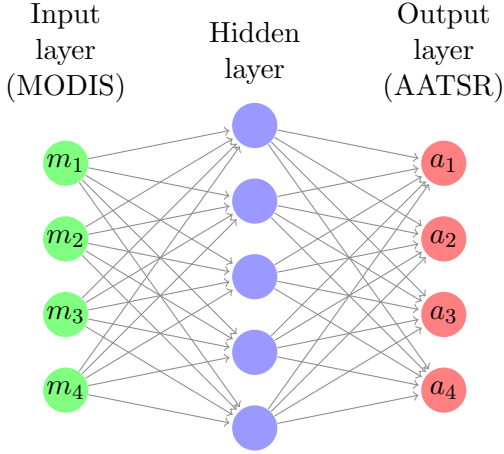


Fig. 5: Neural network taking measurements from four MODIS channels, returning predictions for the four AATSR channels

the hidden layer, which could be very large and thus computationally impractical.

Two methods for optimising the weights were trialled here: Adam, a stochastic gradient-based optimiser and the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) algorithm, a quasi-Newtonian iterative optimisation method. Adam is commonly used for weight optimisation as it is computationally efficient and scales well to problems with large volumes of data. However, given that the datasets in this project are relatively small (<1000 entries), LBFGS was trialled as an alternative.

## 4 Results

### 4.1 SVD Analysis

The matrix transforming MODIS measurements to primary spectrometer measurements (Eq. 9) was approximately diagonal, Fig. 6. This was to be expected: there is significant overlap between the SRFs of different instruments in corresponding channels (Fig. 2), meaning that a measurement of a surface in channel  $i$  is the main predictor of the outcome of a measurement of that surface in channel  $i$  of a different instrument.

$$\begin{bmatrix} 0.97 & 0.03 & -0.0006 & 0.01 \\ 0.03 & 0.98 & -0.01 & 0.02 \\ -0.002 & -0.005 & 1.01 & -0.03 \\ 0.005 & -0.005 & 0.004 & 1.0 \end{bmatrix}$$

Fig. 6: Matrix describing SVD-based transformation from MODIS measurement to SLSTR

The most significant off-diagonal terms are adjacent to the diagonal terms – the MODIS channel 1 measurement is predictive of the measurement in SLSTR channel 2. This suggests that spectral features are spectrally localised: the presence of a spectral feature in instrumental channel  $i$  is correlated with its presence in channel  $i \pm 1$ .

Reconstruction error in channel  $k$  was defined as

$$\epsilon_k = R_k - R_{k,m} \quad (11)$$

where  $R_k$  is the true radiance measurement in channel  $k$  by the primary radiometer and  $R_{k,m}$  is the prediction of that measurement outcome from the MODIS measurements; dividing this by  $R_k$  gives the fractional error. This metric was used in assessing the performance of both the SVD and MLP reconstructions.

Mean fractional reconstruction error was no greater than 2% in any band, and measurements in corresponding bands for the primary instruments were reconstructed with similar accuracy (Fig. 7, hence further analysis of the SVD-reconstruction is restricted to SLSTR as this is representative of AATSR.

The intention of the reconstruction was to reduce the discrepancy between the MODIS albedo measurement and the albedo measurement performed by the primary satellite. This was successful for SVD reconstructions using both the curated and large datasets: the majority of measurements fall below the  $y = x$  line – the locus of measurements for which reconstruction made no difference. In the curated case, the mean albedo difference across all SLSTR channels was reduced by 99.9%; for the large dataset, by 52%. This is indicative of the importance of dataset curation, however a larger valida-

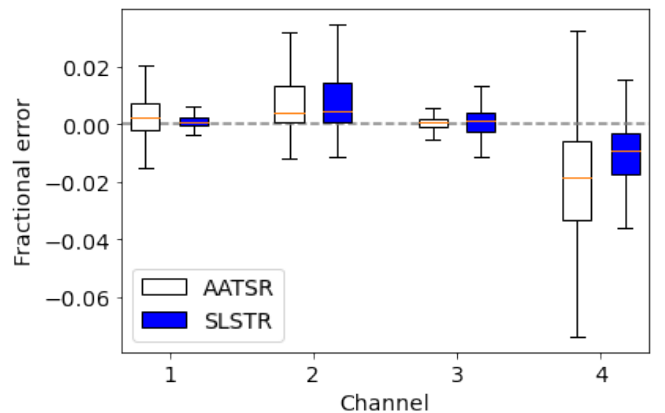
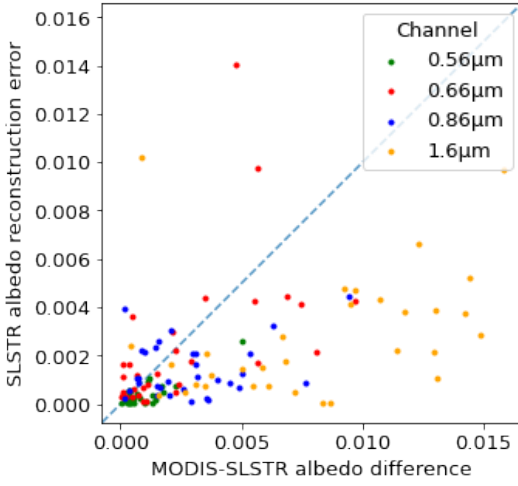
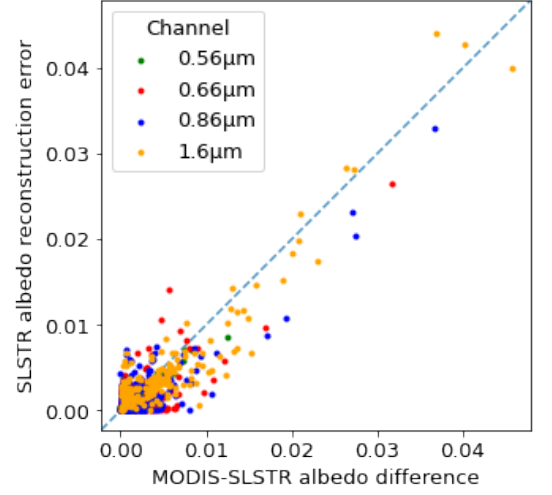


Fig. 7: Comparison of SVD reconstruction error between instrumental channels; outliers omitted





(a) Curated SVD validated on 32 spectra



(b) Large SVD validated on 192 spectra

Fig. 8: Comparison of difference between original MODIS and SLSTR albedo measurements, and SLSTR reconstruction error

tion dataset is needed before robust conclusions can be drawn.

The greatest adjustment in albedo occurred in those channels where overlap between the SRF of the primary instrument and MODIS was small, Fig. 9. The magnitude of the alteration was largest in channels 2, 3 and 4, where there are significant differences between the MODIS and SLSTR SRFs, and negligible in channel 1, where they overlap.

The SVD reconstruction algorithm developed here using the large dataset was tested against the existing SVD reconstruction algorithm developed by Sayer et al. [4] by validating both on the 192-surface spectrum dataset. The algorithm presented here was found to reduce the magnitude of the fractional reconstruction error across all channels by 46% compared to the existing one, Fig. 10.

The superior performance of the updated SVD algorithm may be explained by the use of a training spectral dataset over three times larger than the

one used by Sayer et al. The performance may have been helped by the fact that, although the validation dataset was ‘unseen’ in training, it originated from the same spectral library as the one used to train the updated algorithm. Given more time, it would be illuminating to validate the algorithms against an entirely independent spectral library, to ensure that the improved performance seen here indicates the algorithms’ ability to reconstruct radiometer measurements, rather than merely identify incidental characteristics of the spectral library used in training.

## 4.2 Neural Network

As discussed in Section 3.2, the MLP was trained on either the large or curated training dataset, using either Adam or LBFGS for weight optimisation.

A preliminary analysis showed that a 100-node hidden layer minimised mean reconstruction error (Eq. 11) across 192 surface spectra, Fig. 11. The choice of activation function was found to have a

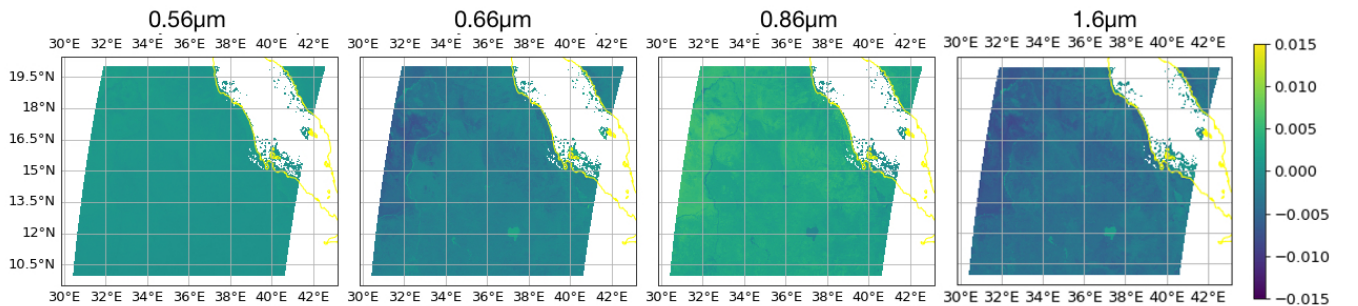


Fig. 9: MODIS albedo adjustment when reconstructing SLSTR measurement using curated SVD over Sudan on February 19<sup>th</sup>, 2022

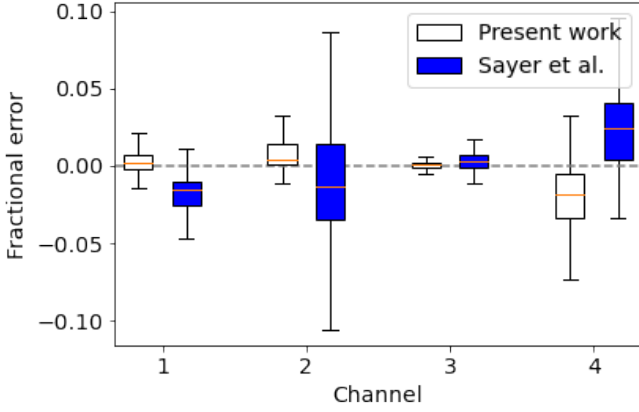


Fig. 10: Comparison of fractional error in AATSR reconstruction of 192 surface by the algorithm designed by Sayer et al. against the algorithm presented here

negligible impact on performance. The rectified linear unit activation function  $\phi(x) = \max(0, x)$  was preferred to  $\tanh$  and logistic functions for its computational efficiency.

As with the SVD analysis, performance varied little across instrumental channels, so discussion is restricted to a single SLSTR channel. Fig. 12 shows the outcome of validation against the testing batch of the selected dataset, where the curated reconstruction was validated against 32 spectra, the large dataset against 192.

The network optimised using Adam was visibly dependent on the size of the training dataset used; the LBFGS-optimised network saw a more modest reduction in mean reconstruction error of 27%. During training, network parameters converged in  $< 20$  epochs when optimised using LBFGS and  $< 60$  using Adam, suggesting that the former is better suited to smaller datasets. The mean fractional reconstruction error was 61% lower for the large-dataset LBFGS

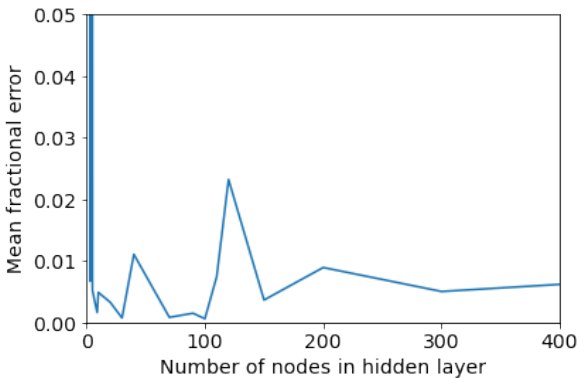


Fig. 11: Minimum reconstruction error obtained for MLP with 100 nodes in the hidden layer; SLSTR channel 1, mean for 192 surface spectra

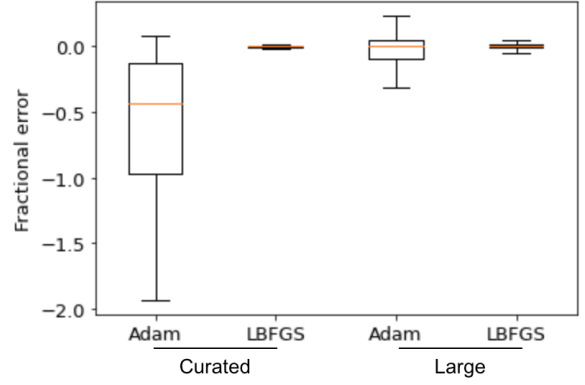


Fig. 12: Comparison of neural network reconstructions for SLSTR channel 1; outliers omitted.

neural network than for best-case reconstruction by the Adam-optimised one, further supporting this conclusion.

One possible explanation is that LBFGS estimates the curvature of the parameter space by evaluating the Hessian; Adam, by contrast, uses as first-order approximation to the Hessian. The full calculation of the Hessian means that LBFGS scales poorly to large datasets compared to Adam. For large datasets ( $> 10,000$  points), Adam is expected to outperform LBFGS, but for the purposes of enquiries of the sort conducted here, where datasets of this size are challenging to obtain, LBFGS is likely to remain the preferred optimisation method.

### 4.3 Comparison of SVD and Neural Network Reconstruction

When trained on the large dataset using the LBFGS-algorithm, the neural network's performance approached that of the SVD reconstruction, but failed to surpass it: comparing best-case performances of both methods, mean SVD reconstruction error was 53% lower, and consistently had a tighter distribution around a median of zero, as seen in Fig. 13. This is also indicated by the fact that SVD-based reconstructions typically fall below the  $y = x$  line in Fig. 14, i.e. reconstructed SLSTR measurement values are closer to the true SLSTR measurement than the unreconstructed MODIS measurements were; the converse is true for the neural network reconstruction.

As noted in Section 3.2, the main challenge faced by a neural network is the approximation of a function that is not known. The larger the training dataset, the better this approximation becomes, as indicated by the improvement in MLP reconstruc-



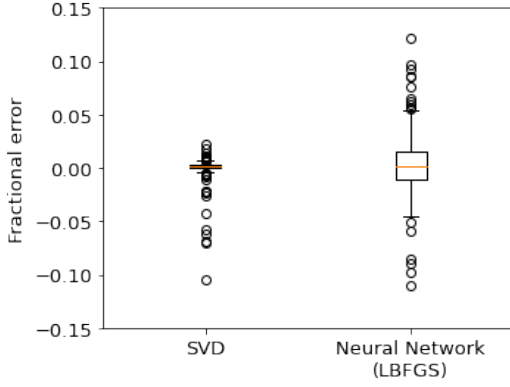


Fig. 13: Reconstruction of SLSTR channel 1 by neural network and SVD, trained on large dataset

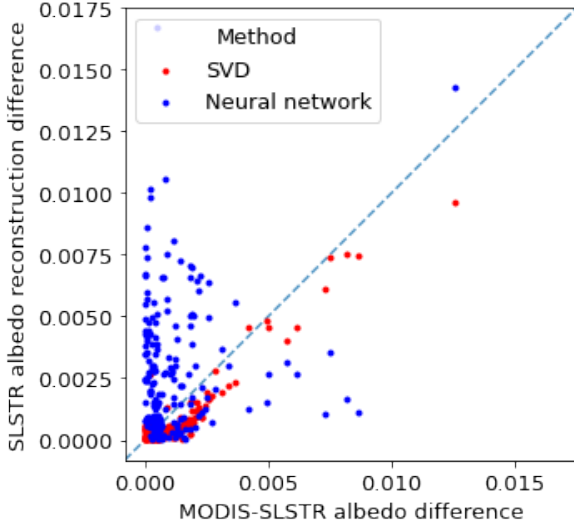


Fig. 14: Comparison of SLSTR Ch. 1 error before and after reconstruction for SVD and neural network

tion for the large training dataset. By contrast, the performance of the SVD reconstruction remained fairly constant with dataset size, suggesting that given a sufficiently large dataset, the neural network would outperform the SVD reconstruction.

#### 4.4 Impact of Reconstruction on ORAC Retrieval

Despite the accurate reconstruction of test spectra in the validation stage, the SVD-generated transformation matrix failed to reduce the error in aerosol retrieval, Fig. 15. Retrieval error was defined as the difference between the ground-based measurement of aerosol optical depth (Eq. 2) by AERONET and the average of all ORAC AOD retrievals within a 15km radius of the ground-based instrument; it was assumed that AOD varies little over this length-scale.

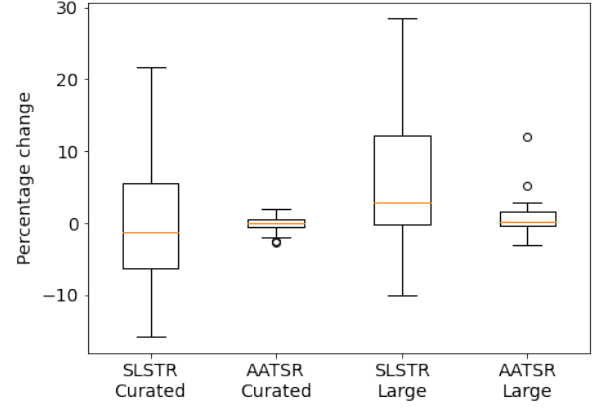


Fig. 15: Mean change in aerosol retrieval error for AATSR and SLSTR across  $\sim 20$  retrievals when SVD reconstruction used

ORAC retrievals using SLSTR measurements were more strongly affected by surface-signal transformation than those using AATSR, as seen in the greater SLSTR variance of Fig. 15. This may suggest that SLSTR retrieval is more sensitive to the land surface *a priori*. However, to draw definitive conclusions, more ORAC retrievals need to be run.

The curated dataset marginally outperformed the larger dataset, reducing the retrieval error by a mean value of 0.4%, while in the latter case, the mean error increased by 4% compared to unaltered retrieval across all instruments and channels; by contrast, the Sayer et al. reconstruction reduced this error by  $\sim 3\%$  for AATSR [4]. This suggests that the spectral dataset used here did not adequately represent the land surfaces encountered by the satellite-based radiometers, with certain (in particular, artificial) surfaces strongly over-represented, relative to their commonality on Earth; a diversification of the spectra used would improve future investigations. It would also be interesting to revisit assumption that AOD varies little over a 15km radius, by examining how these results change if the distance over which ORAC retrievals are averaged is adjusted.

Due to the scope of the project and the technical complexity of implementing a neural network in the main ORAC algorithm, it was not possible to directly assess the neural network's effect on aerosol retrieval.

## 5 Conclusion

The difference in the spectral responses of spectrometers used in aerosol retrieval is a source of uncertainty in the retrieval of aerosol properties. Here,

SVD and neural network methods were trialled in an attempt to reduce this error.

The new SVD algorithm outperformed the algorithm designed by Sayer et al. in validation, but its impact on ORAC retrieval was poor. One possible explanation is that the curated training dataset used was smaller, after segmentation into testing and training data, than that of Sayer et al, while the larger training dataset, which exceeded the one used in the Sayer paper, failed to accurately represent the Earth’s surface due to redundancy in the data and inclusion of surfaces that are not representative of the land captured by the satellites.

Although the MLP underperformed the SVD reconstruction here, it has shown its promise as a method for reconstructing land surface measurements between instruments. Both the LBFGS- and Adam-optimised networks benefited from a larger training dataset, suggesting that, given a sufficiently large training dataset, they would outperform the SVD reconstruction.

The superior performance of the curated SVD during ORAC retrieval indicates the importance of accurately representing the Earth’s surface in the training dataset. The surface spectra used here were failed to represent the Earth’s surface in part because each contained spectral data from only one surface type. In reality, numerous types of land cover may be present in a single satellite surface measurement. This could be simulated most simply by superposing one or more spectra from the existing dataset. By adjusting the relative weights of the superposed spectra, an arbitrarily large number of example surface spectra could, in principle, be generated [4]. It remains to be seen whether the resulting spectra would be sufficiently diverse to improve the reconstruction algorithms. Given more time, it would be illuminating to implement the neural network reconstruction into ORAC retrievals.

In addition to an expansion of training datasets, there is scope for significant development of neural network reconstruction methods. Here, the MLP was used as a fitting algorithm, however neural networks are commonly used in computer vision applications to identify patterns in an image. The ability to handle a large number of inputs in a versatile manner sets neural networks apart from SVD methods. In the real land scenes captured by satellite radiometers, pixels have additional information in the form of spatial context: neighbouring pixels are generally correlated, since terrain

types are roughly continuous. This information was not utilised in this study, since training and reconstruction are done on a pixel-by-pixel basis, rather than treating them together as an image. Moreover, the MODIS MCD43C1 surface albedo data product used in ORAC is time-varying. A neural network taking a time-series of MODIS land scenes as an input would benefit from spatiotemporal context, which would improve the neural network’s predictions.

Finally, there is scope for utilising the neural network in combination with SVD analysis by using an SVD reconstruction to preprocess data, aiding convergence on optimal network parameters. In short, neural networks remain a promising method for constraining land signal in aerosol retrievals.

## References

- [1] Masson-Delmotte, V., P. Zhai, A. Pirani et al. IPCC, 2021: *Climate Change 2021: The Physical Science Basis*. Cambridge University Press.
- [2] Sparrow, S., Millar, R., Yamazaki, K. et al. (2018). *Finding Ocean States That Are Consistent with Observations from a Perturbed Physics Parameter Ensemble*. Journal Of Climate, 31(12), 4639-4656.
- [3] Levy, R., Mattoo, S., Munchak, L., Remer, L., Sayer, A., Patadia, F., & Hsu, N. (2013). *The Collection 6 MODIS aerosol products over land and ocean*. Atmospheric Measurement Techniques, 6(11), 2989-3034.
- [4] Sayer, A., Thomas, G., Grainger, R. et al. (2012). *Use of MODIS-derived surface reflectance data in the ORAC-AATSR aerosol retrieval algorithm: Impact of differences between sensor spectral response functions*. Remote Sensing Of Environment, 116, 177-188.
- [5] NASA Ocean Color. (2022). Retrieved 11 January 2022, from [https://oceancolor.gsfc.nasa.gov/docs/rsr/rsr\\_tables/](https://oceancolor.gsfc.nasa.gov/docs/rsr/rsr_tables/)
- [6] Kokaly, R.F., Clark, R.N., Swayze, G.A. et al., 2017, *USGS Spectral Library Version 7*: U.S. Geological Survey Data Series 1035.
- [7] Fao.org. 2022. *FAOSTAT Land Cover*. [online] Available at: <https://www.fao.org/faostat/en/#data/LC> [Accessed 24 March 2022].

- [8] Cybenko, G., *Approximations by superpositions of sigmoidal functions*, Mathematics of Control, Signals, and Systems, 2(4), 303–314, 1989.

## Bibliography

- Bermeitinger, B., Hrycej, T., & Handschuh, S. (2019, September). *Singular value decomposition and neural networks*. In International Conference on Artificial Neural Networks (pp. 153-164). Springer, Cham.
- Dawson, A., 2016. eofs: A Library for EOF Analysis of Meteorological, Oceanographic, and Climate Data. *Journal of Open Research Software*, 4(1), p.e14.
- European Space Agency, 2019, *ADAM Surface Reflectance Database* v4.0
- EUMETSAT, 2022. *Spectral response functions — NWP SAF*. [online] Nwp-saf.eumetsat.int. Available at: <https://nwp-saf.eumetsat.int/site/software/rttov/download/coefficients/spectral-response-functions/> [Accessed 10 March 2022].
- Giles, D., Sinyuk, A., Sorokin, M. et al. (2019). *Advancements in the Aerosol Robotic Network (AERONET) Version 3 database* Atmospheric Measurement Techniques, 12(1), 169-209.
- Hulley, G., Hook, S., 2022, ECOSTRESS Level-2 Land Surface Temperature and Emissivity Algorithm Theoretical Basis Document (ATBD)
- Kingma, D. P., & Ba, J. (2014). *Adam: A Method for Stochastic Optimization*.
- Kommalapati, Raghava R.; Valsaraj, Kalliat T. (2009). *Atmospheric aerosols: Characterization, chemistry, modeling, and climate*. Vol. 1005. Washington, DC: American Chemical Society. pp. 1–10.
- Kozak, D. *The reason of superiority of Limited-memory BFGS over ADAM solver*, URL (version: 2020-05-11): <https://stats.stackexchange.com/q/315632>
- Pedregosa et al., *Scikit-learn: Machine Learning in Python*, JMLR 12, pp. 2825-2830, 2011.
- Robson, J., Aksenov, Y., Bracegirdle, T., Dimdore-Miles, O., Griffiths, P., & Grosvenor, D. et al. (2020). *The Evaluation of the North Atlantic Climate System in UKESM1 Historical Simulations for CMIP6*. Journal Of Advances In Modeling Earth Systems, 12(9).
- Schaaf, C., Wang, Z. (2015). textitMCD43C1 MODIS/Terra+Aqua BRDF/AlbedoModel Parameters Daily L3 Global 0.05Deg CMG V006[Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2022-02-20 from <https://doi.org/10.5067/MODIS/MCD43C1.006>
- Spectral Response Function Data - Sentinel Online. (2022). Retrieved 11 January 2022, from <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-3-slstr/instrument/measured-spectral-response-function-data>
- Stevens, B., Feingold, G. *Untangling aerosol effects on clouds and precipitation in a buffered system*. Nature **461**, 607–613 (2009).
- Thomas, G. E., Poulsen, C. A., Sayer et al. (2009a). *Oxford-RAL Aerosol and Cloud (ORAC): Aerosol retrievals from satellite radiometers*. In A. A. Kokhanvosky, & G. de Leeuw (Eds.), Satellite Aerosol Remote Sensing Over Land. Berlin: Springer.