A LOOK-UP-TABLE APPROACH TO INVERTING REMOTELY SENSED OCEAN COLOR DATA

Curtis D. Mobley, Sequoia Scientific, Inc., 15317 NE 90th St., Redmond, WA 98052
Lydia K. Sundman, Sequoia Scientific, Inc., Redmond, WA 98052
Curtiss O. Davis, Naval Research Lab. Code 7212, Washington D.C. 20375
Marcos Montes, Naval Research Lab. Code 7212, Washington D.C. 20375
W. Paul Bissett, Florida Environmental Research Inst.,Tampa, FL 33611

ABSTRACT

We describe a new technique for extraction of environmental information such as water inherent optical properties (IOPs) and shallow-water bottom depth (z_b) and spectral reflectance (R_b) from remotely-sensed hyperspectral ocean-color data. Our look-up-table (LUT) approach compares a measured remote-sensing reflectance (R_{rs}) spectrum with tabulated R_{rs} spectra generated by a large number of Hydrolight runs having a range of water IOPs, bottom depths and reflectances, solar angles and viewing directions. The environmental conditions corresponding to the measured R_{rs} spectrum are assumed to be the same as the input conditions for the Hydrolight run that generated the database R_{rs} spectrum that most closely matches the measured R_{rs} spectrum. Here we show preliminary results based on water and bottom conditions at Lee Stocking Island, Bahamas during the Coastal Benthic Optical Properties (CoBOP) field experiments. The LUT method works very well in these initial studies. Non-uniqueness of R_{rs} spectra does not appear to be a problem, and wavelength-uncorrelated random noise does not affect the spectrum matching.

APPROACH

We used Hydrolight to generate a database of R_{rs} spectra corresponding to various combinations of measured and modeled water-column IOPs, bottom reflectance spectra, and bottom depths. Each R_{rs} spectrum in the database is indexed by labels that identify the IOPs, bottom reflectance R_b, and bottom depth z_b that were used as input to the Hydrolight run that generated the R_{rs} spectrum. We then generated test spectra $\tilde{\boldsymbol{R}}_{\mathbf{rs}}$ using other sets of IOPs, R_b, and z_b (sometimes similar to those in the database, and sometimes dissimilar). These test spectra were used to evaluate the performance of various possible spectrum-matching algorithms.

THE DATABASE

Water-column IOPs. Four sets of representative water-column IOPs measured during the May 2000 CoBOP field experiments at Lee Stocking Island, Bahamas were used in generating the database of R_{rs} spectra for initial evaluation of spectrum matching algorithms. The IOPs were assumed to be homogeneous. The particle backscatter fraction $B_p = b_{bp}/b_p$ was specified to be 0.02, which may be representative of the mixture

of small mineral particles and phytoplankton suspended in the LSI waters. The resulting total backscatter coefficients B (including water backscatter, which has $B_{water} = 0.5$) then are in the range of 0.025 to 0.05 for LSI waters (see Fig. 3). The total backscatter fraction B at the current wavelength was used to generate the total phase function from the Fournier-Forand family of phase functions as described in Mobley et al. (2002). We wanted to have both representative and non-representative R_{rs} spectra in the database for purposes of evaluating the spectrum matching algorithms for LSI. We therefore added six more sets of IOPs to the four measured sets. The first was simply pure water. The remaining five spectra were generated from the Hydrolight IOP model for Case 1 waters for chlorophyll concentrations of Chl = 0.05, 0.20, 0.50, 1.0 and 2.0 mg m^{-3}. In these runs, the particle backscatter fraction was set to $B_p = 0.005$, which is typical of phytoplankton in Case 1 waters. Although LSI waters typically have Chl concentrations of 0.05 to 0.3 mg m^{-3}, these Case 1 IOPs are not expected to be representative of the waters around LSI, which are Case 2 waters because of the additional CDOM, whose origin is believed to be corals and seagrass. The LSI water also may contain resuspended sediments and small calcite particles that precipitate out of solution. The five total Case 1 absorption spectra are shown in Fig. 1 as dotted lines. The total scattering coefficients b corresponding to Fig. 1 are shown in Fig. 2. Figure 3 shows the corresponding total backscatter fraction B.

Bottom reflectance spectra. Twenty-two bottom reflectance spectra were used to generate the R_{rs} database: 11 spectra measured on sediments in the vicinity of LSI (shown as solid lines in Fig. 4), 6 spectra generated as weighted averages of clean coral sand and green algae (dotted lines in Fig. 4), and 5 spectra that were weighted averages of green algae and a clean seagrass leaf (dashed lines). With the assumption that the bottom BRDF is Lambertian for the incident lighting conditions and viewing geometry, these spectra are equivalent to bottom irradiance reflectance spectra, $R_b = E_u/E_d$.

Bottom depths. The R_b spectra were applied at nine bottom depth values: z_b = 0.5, 1, 2, 5, 10, 15, 20, 30, and 50 m. A tenth option, infinitely deep water, was also included. In this case, the non-Lambertian "bottom" BRDF was computed from the water-column IOPs.

Remote-sensing reflectances. The database thus has 10 sets of IOPs, 22 bottom reflectances, and 10 bottom depths. When combined in Hydrolight runs, these yield 10×22×10 = 2,200 R_{rs} spectra, all else being the same. For the initial database, only one solar zenith angle was used: $\theta_s = 40$ deg, and only the polar cap water-leaving radiances were used to compute R_{rs}. Hydrolight runs were made for 60 wavelength bands from 400 to 700 nm; each bandwidth was 5 nm. Figure 5 shows the 2,200 spectra in this initial R_{rs} database.

THE TEST SPECTRA

We next generated a small set of R_{rs} spectra to use in developing and evaluating spectrum-matching algorithms. The intention was to have some R_{rs} spectra that corresponded closely to the IOPs, bottom reflectances, and bottom depths used to generate the R_{rs} database. These R_{rs} should be easy to match with nearly similar spectra

in the database, whose indices would then identify IOPs, reflectances, and depths close to the ones used to generate the test spectra. Other test spectra were generated with IOPs, bottom reflectances, and bottom depths unlike those in the original database. These spectra should, in principle, be hard to match and may lead to recovered IOPs, reflectances, or depths unlike the ones that actually generated the test spectrum. There of course may be R_{rs} spectra in the database that are very similar in shape and magnitude to the test spectra even though they were generated with greatly different IOPs, reflectances, and bottom depths; this is where non-uniqueness may cause problems.

The test spectra had IOPs defined as follows. IOP test spectrum 1 was a weighted average of two of the LSI IOP spectra (25% of LSI ac-9 spectrum 2 and 75% of LSI spectrum 3); these IOPs are thus similar to those used to create the database. IOP test spectrum 2 was generated using the Hydrolight Case 2 water IOP model, with Chl = 0.1 mg m^{-3}, B_p = 0.005; CDOM proportional to Chl as in Case 1 water; and 0.5 gm m^{-3} of calcareous sand with B_p = 0.03. IOP test spectrum 3 was similar to spectrum 2, except that the mineral component was 1 gm m^{-3} of red clay particles with B_p = 0.02. IOP spectra 2 and 3 are quite different from any of those in the database.

Three test bottom reflectance spectra were defined as follows. R_b test spectrum 1 was a weighted average of two LSI spectra (25 % of LSI spectrum 7 and 75% of LSI spectrum 3, both of which are sand spectra). R_b test spectrum 2 was a weighted average of 25% coral sand and 75% red algae. R_b test spectrum 3 represents gray mud and has a reflectance of 0.1 at all wavelengths. Again, R_b spectrum 1 is similar to spectra in the database, and spectra 2 and 3 are unlike any spectra in the database. Three bottom depths were used: z_b = 4.5 m (close to the database value of 5 m), 12.5 m (half way between database values of 10 and 15 m), and infinite. The test IOP, R_b, and z_b data were then used to generate 3×3×3 = 27 R_{rs} test spectra, which are shown in Fig. 6.

SPECTRUM-MATCHING ALGORITHMS

There are many criteria that can be used to match a given spectrum $\tilde{R}_{rs}$ with one in the R_{rs} database. Let $R_{rs}(i,\lambda_j)$ denote the i^{th} spectrum in the database, i = 1 to I (I =2200 at present), which is defined for discrete wavelengths λ_j, j = 1 to J (J = 60 at present). Let $w(j)$ be a weighting function that can be chosen to de-emphasize wavelengths where the data are of suspicious quality or carry less information. The weighting function satisfies $0 \le w(j) \le 1$ for each j. Setting $w(j) = 1$ for each j gives each wavelength equal importance. Setting $w(j) = 0$ for a particular j omits that wavelength from use in the matching algorithm. A few simple candidate spectrum-matching algorithms (SMAs) are then defined as follows.

SMA1: Least-squares matching of the absolute spectra. The database spectrum that most closely matches the given spectrum $\tilde{R}_{rs}$ is the one that minimizes

$$LSQabs(i) = \sum_j \left\{ w(j) \left[R_{rs}(i,\lambda_j) - \tilde{R}_{rs}(\lambda_j) \right]^2 \right\}. \quad (1)$$

The value of i giving the minimum value of $LSQabs(i)$, call it i_{min}, is just the record index

of the database spectrum.

SMA2: Least-squares matching of length-normalized spectra. If the spectra are relative values (such as digital counts) or may have multiplicative errors, then it makes sense to normalize each spectrum by its euclidian length in *J*-dimensional wavelength space, defined as

$$|R_{rs}(i)| = \sqrt{\sum_j R_{rs}^2(i,\lambda_j)}\,. \tag{2}$$

The quantity to be minimized is then

$$LSQnorm(i) = \sum_j \left\{ w(j) \left[\frac{R_{rs}(i,\lambda_j)}{|R_{rs}(i)|} - \frac{\tilde{R}_{rs}(\lambda_j)}{|\tilde{R}_{rs}|} \right]^2 \right\}. \tag{3}$$

Algorithm (3) will regard two spectra as being the same if they differ only by a multiplicative factor.

SMA3: Angle-matching of length-normalized spectra. Using the same normalization as in algorithm (3) to covert the spectra to unit vectors, the cosine of the "angle" between the unit vectors is then

$$\cos(\theta_i) = \sum_j \left\{ w(j) \frac{R_{rs}(i,\lambda_j)}{|R_{rs}(i)|} \frac{\tilde{R}_{rs}(\lambda_j)}{|\tilde{R}_{rs}|} \right\}. \tag{4}$$

The closest spectrum in the database is the one that maximizes $\cos(\theta_i)$, i.e., that minimizes the angle between the two unit vector spectra. Two spectra that differ only by a multiplicative factor will have a cosine of 1, i.e., an angle of 0 between them. If $w(j) = 1$ for all j, algorithm (4) corresponds to the matching criterion used in the Spectral Image Processing System (SIPS; see Kruse, et al., 1993).

SMA4: Least-squares matching of offset spectra. If the spectra may have additive errors (such as from imperfect removal of a wavelength-independent dark current value), then it makes sense to shift the spectra to a common baseline value. For example, each spectrum can be shifted so that its minimum value is zero. Let

$$\min R_{rs}(i) = \min_j R_{rs}(i,\lambda_j) \tag{5}$$

Then the spectra are compared using

$$LSQoffset(i) = \sum_j \left\{ w(j) \left[\left(R_{rs}(i,\lambda_j) - \min R_{rs}(i) \right) - \left(\tilde{R}_{rs}(\lambda_j) - \min \tilde{R}_{rs} \right) \right]^2 \right\}. \tag{6}$$

SMA5: Least-squares matching of zero-offset, then length-normalized spectra. The additive and multiplicative normalizations can be combined. One can first shift the spectra to a common baseline value of zero, as in 5.4, and then normalize to unit length as in 5.2. The least-squares fit of the resulting shifted, normalized spectra is then used to

find the closest match.

SMA6: Least-squares matching of length-normalized, then zero-offset spectra. Alternatively, one can first normalize the spectra to unit length and then shift them to zero at their minimum value. The least-squares fit of the resulting normalized, shifted spectra is then used to find the closest match. Note that zero-shifting and unit-normalization are not commutative; thus algorithms 5.5 and 5.6 are in principle different.

There are other possible spectrum-matching algorithms, such as matching wavelength derivatives. We have not yet evaluated the performance of derivative-matching (or other) algorithms.

INITIAL RESULTS

Table 1 shows the results for three test R_{rs} spectra that are based on LSI IOP and R_b spectra, and which have z_b values close to those in the database. Recall that test IOP spectrum 1 (which is used here) is a weighted average of two of the LSI IOP spectra (25% of LSI ac-9 spectrum 2 and 75% of LSI spectrum 3). Test R_b spectrum 1 (used here) is a weighted average of two LSI sediment spectra (25 % of LSI spectrum 7 and 75% of LSI spectrum 3, both of which are sand spectra). The three bottom test depths were 4.5 m (near the database value of 5 m), 12.5 m (in between 10 and 15 m in the database), and infinitely deep. In retrieving the IOPs, bottom reflectance, and depth for these test spectra, we should be overjoyed if the spectrum-matching algorithm concludes that the IOPs were either database IOP spectrum 2 or 3, database R_b spectrum 3 or 7, and depth 5 m (for test spectrum 1, which has $z_b = 4.5$ m), 10 or 15 m (for test spectrum 2, which has $z_b = 12.5$ m), or infinitely deep. As seen in Table 1, each spectrum-matching algorithm concluded that the water IOPs were database spectrum 2, R_b spectrum 3, and bottom depth $z_b = 5$ m. The retrieval of environmental information is thus considered to be correct in all cases.

Test spectrum 2 was a bit harder because the bottom was deeper, making it harder to distinguish bottom reflectance effects in the R_{rs} spectrum. Nevertheless, four of the matching algorithms concluded that the IOPs were database spectrum 3, R_b spectrum 3, and bottom depth 15 m, which are all acceptable. Two of the matching algorithms gave IOP spectrum 2, and bottom depth 10 m, which are acceptable, but chose bottom reflectance spectrum 11. However, database R_b spectrum 11 is another sand spectrum, so this probably also should be regarded as an acceptable recovery.

Test spectrum 3 had infinitely deep water. Each matching algorithm gave IOP spectrum 3, which is acceptable. Each of the algorithms placed the bottom depth at either 50 m or infinite, either of which could be considered acceptable because 50 m is very deep water and not much different from infinite as regards the bottom effect on R_{rs}. Five of the algorithms gave bottom reflectance spectrum 1, presumably because no spectrum was ever better than the first one considered, since the bottom was infinitely deep and no bottom spectrum from the tabulated values was used in the Hydrolight runs to generate these R_{rs} spectra.

Thus all of the spectrum-matching algorithms performed acceptably well, i.e. they recovered environmental data from the database that was close to what was used to

generate the test spectra. Of course, these test cases were rather ideal. In particular, there was no noise in the database or test R_{rs} spectra. However, the addition of even large amounts of wavelength-uncorrelated random noise seldom changed the retrieval. This is because in comparing the spectra over the 60 wavelengths, e.g., via Eq. (1), we are essentially summing 60 zero-mean random deviates, which nearly sum to zero and thus do not change the match except in the occasional case with a large deviation at one wavelength is not balanced by a deviation of the opposite sign at another wavelength. Nevertheless, it is reassuring that noise of this type does not greatly affect the spectrum-matching. It remains for future research using real spectra to see which spectrum-matching algorithms perform best in various real-world situations (e.g, when wavelength-correlated noise may be present).

The test set also contained 24 R_{rs} spectra that were purposely constructed with IOPs and bottom reflectances that were not representative of the IOP and R_b spectra used to construct the database. These 24 spectra had either IOPs, or R_b spectra, or both, that did not correspond to any IOP and R_b spectra in the database. Nevertheless, the bottom depth was often accurately retrieved when the water was shallow. There were 48 recoveries of bottom depth for which the test spectrum had $z_b = 4.5$ m (8 spectra times 6 algorithms). In 35 of the 48 retrievals (78%), the bottom depth was correctly retrieved to be 5 m (the closest depth in the database). In 9 of 48 retrievals the 4.5 m bottom depth was recovered as 10 m, and in 4 cases as 15 m. For the 48 recoveries where the true bottom depth was 12.5 m, the recovered depth was 10 m (10 times), 15 m (11 times), 20 m (21 times), and 30 m (6 times). Thus 44% of the depth recoveries could be regarded as correct in the sense that the closest database values were chosen.

ONGOING WORK

The above results are encouraging. As we saw in Section 6, various spectrum-matching algorithms seem to perform comparably in finding the closest database spectrum. There are some minor differences, but all did rather well at extracting environmental information. Wavelength-uncorrelated random noise does not greatly impede the spectrum matching. Thus it appears that the choice of spectrum-matching algorithm may not be critical, and wavelength-uncorrelated random noise is not likely to be a problem. (However, it remains to be seen if there are significant differences in algorithm performance when applied to real data, which may have noise characteristics that favor particular search algorithms.)

We are currently applying the LUT methodology to a PHILLS image taken during the May 2000 Closure Experiment at LSI. Results will be reported at the Ocean Optics conference.

ACKNOWLEDGMENTS

This work is supported by the Environmental Optics Program of the U. S. Office of Naval Research. We thank Emmanuel Boss and Ron Zaneveld for use of their IOP data from LSI, and Carol Stephens for use of her LSI bottom reflectance spectra.

REFERENCES

Kruse, F. A., A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Brloon, and A. F. H. Goetz, 1993. The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ*. **44**, 145-163.

Mobley, C. D., L. K. Sundman, and E. Boss, 2002. Phase function effects on oceanic light fields. *Appl. Optics* **41**, 1035-1050.

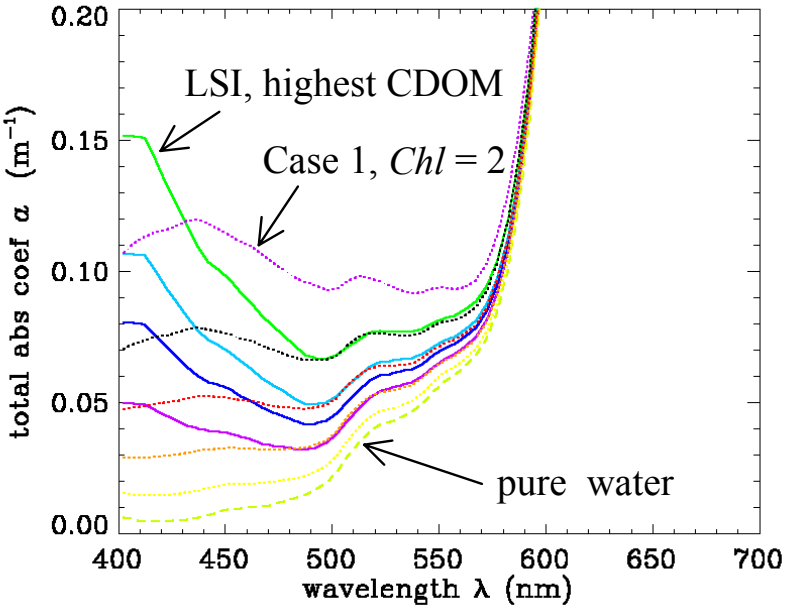

Fig 1. The ten total absorption spectra in the LUT database. Solid lines, from LSI ac-9 measurements; dotted lines, from Case 1 bio-optical model for Chl = 0.05 to 2 mg m^{-3}; dashed line, pure water.

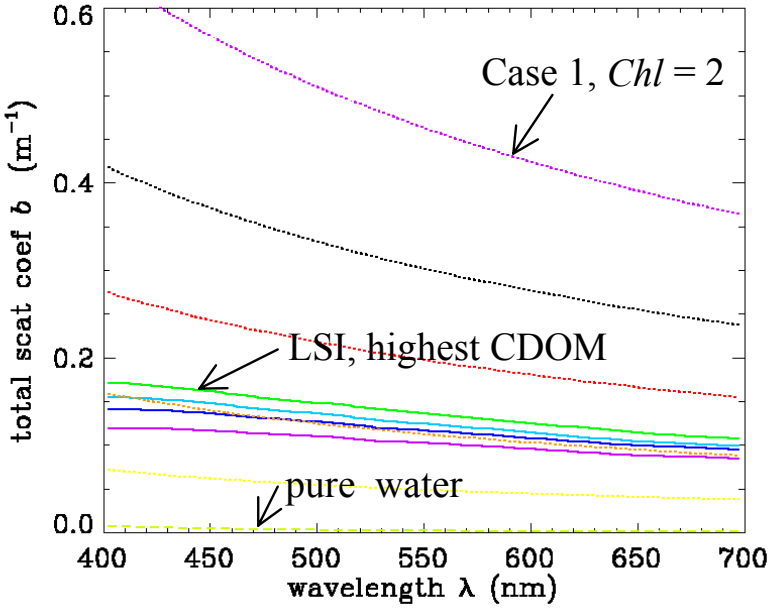

Fig 2. Total scattering spectra. Line patterns and colors are the same as in Fig. 1.

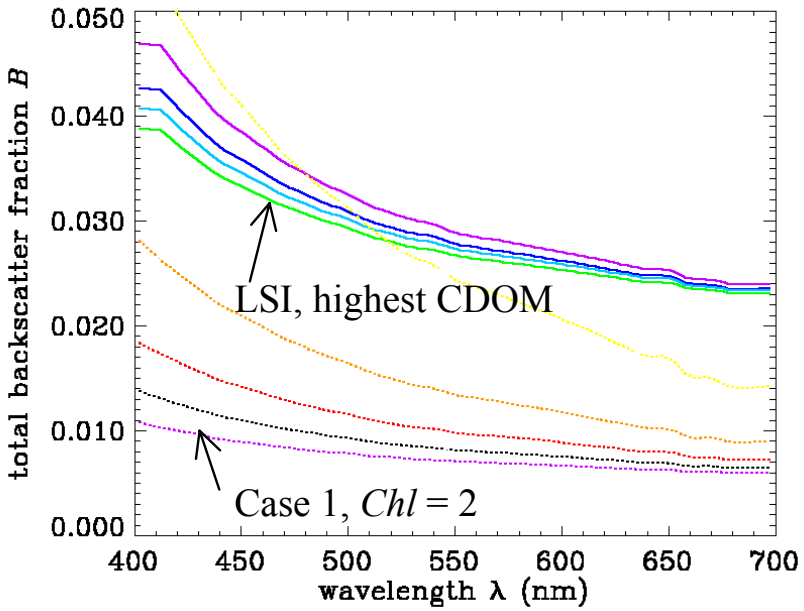

Fig 3. Total backscatter fractions. Line patterns and colors are the same as in Fig. 1. The value for pure water is 0.5 at all wavelengths.

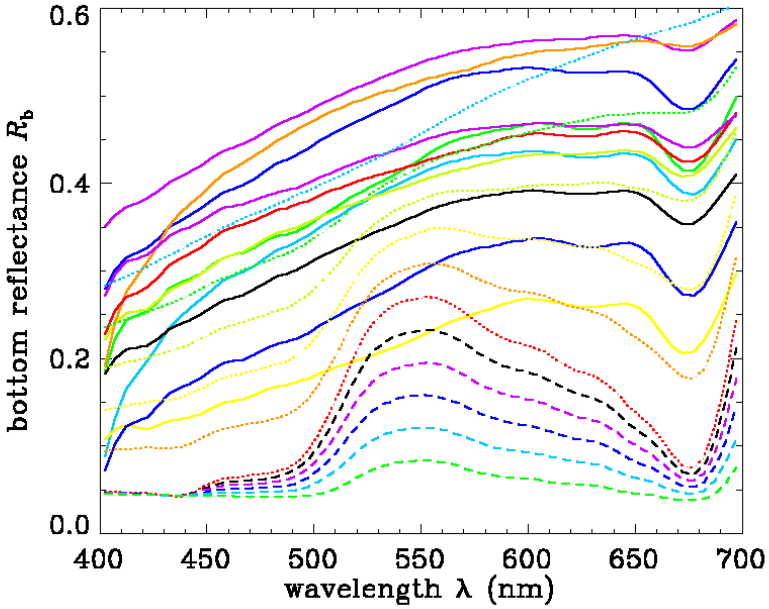

Fig. 4. Bottom reflectance spectra included in the database.

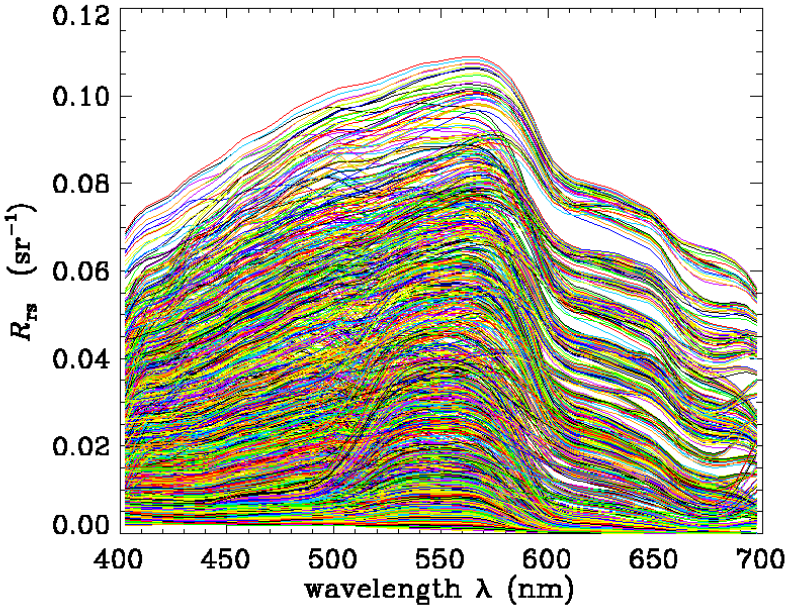

Fig. 5. The 2,200 R_{rs} spectra in the initial database.

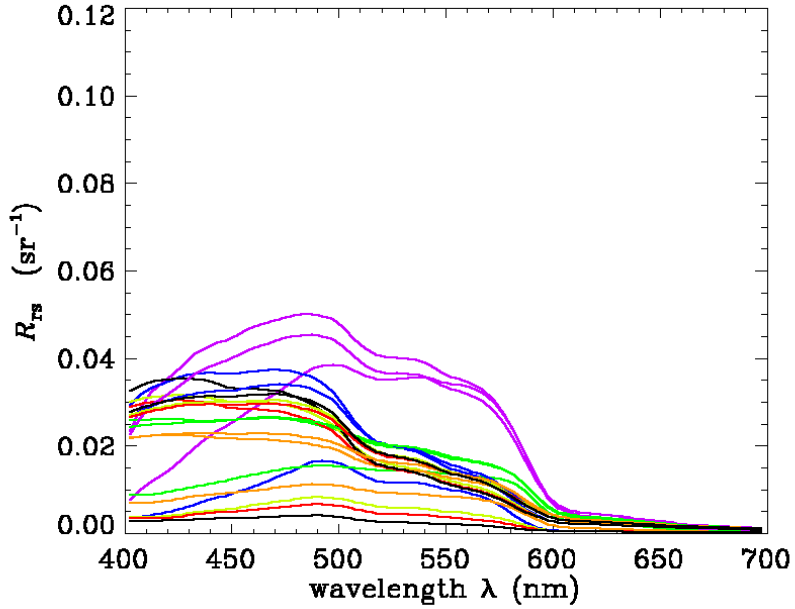

Fig. 6. The 27 test R_{rs} spectra.

Table 1. Performance of spectrum-matching algorithms SMA1 to SMA6 for three noise-free test R_{rs} spectra that are similar the LSI-based spectra in the database. IOP = 2 means that database IOP spectrum 2 was found to be the best match for the test IOP spectrum, etc. Incorrect retrievals are shown in red.

	test spectrum 1	test spectrum 2	test spectrum 3
	IOP = 0.25(#2) + 0.75(#3) R_b = 0.75(#3) + 0.25(#7), z_b= 4.5 m	IOP = 0.25(#2) + 0.75(#3) R_b = 0.75(#3) + 0.25(#7), z_b= 12.5 m	IOP = 0.25(#2) + 0.75(#3) R_b = 0.75(#3) + 0.25(#7), z_b= ∞
SMA1: LSQ diff of unnorm R_{rs} (Eq. 1)	IOP = 2; R_b = 3; z_b = 5 m	IOP = 2; R_b = 11; z_b = 10 m	IOP = 3; R_b = 1; z_b = 50 m
SMA2: LSQ diff of length norm (Eq. 3)	IOP = 2; R_b = 3; z_b = 5 m	IOP = 3; R_b = 3; z_b = 15 m	IOP = 3; R_b = 1; z_b = ∞
SMA3: min angle between length norm (Eq. 4)	IOP = 2; R_b = 3; z_b = 5 m	IOP = 3; R_b = 3; z_b = 15 m	IOP = 3; R_b = 22; z_b = ∞
SMA4: LSQ of zero offset (Eq. 6)	IOP = 2; R_b = 3; z_b = 5 m	IOP = 2; R_b = 11; z_b = 10 m	IOP = 3; R_b = 1; z_b = 50 m
SMA5: LSQ of zero offset, then length norm	IOP = 2; R_b = 3; z_b = 5 m	IOP = 3; R_b = 3; z_b = 15 m	IOP = 3; R_b = 1; z_b = ∞
SMA6: . LSQ of length norm, then zero offset	IOP = 2; R_b = 3; z_b = 5 m	IOP = 3; R_b = 3; z_b = 15 m	IOP = 3; R_b = 1; z_b = ∞