Inspiring Great British Manufacturing

Report Title: **D3 – Trials**

Version Number: 1.1

Project Title: 38157 - *Hyperspectral Imaging for Soils*

Prepared For: Electricity North West

Author: Bethan Nye

Date: 11 May 2023

Executive Summary

The Environmental Agency (EA) Regulatory Position Statement 211 applies, as of March 2023, to business that handle excavated waste from unplanned utilities installation and repair works. Under RPS211, small amounts of spoil from unplanned works can be disposed of or recycled as non-hazardous waste. Electricity North West (ENWL) and fellow utility organisations use this exemption to dispose of or recycle spoil which cannot be reused as backfill for thousands of emergency works activities across the UK each year.

RPS211 has been under review, and will be withdrawn by April 2023. Withdrawal for this RPS will require companies such as ENWL to classify all of its waste. Without the ability to classify waste on-site, they will be required to either dispose of all waste as hazardous, or sort and segregate waste at an off-site location. The Energy Innovation Centre has quoted an average disposal cost of £40 for 10 m^3 of non-hazardous spoil, and £1,200 for 10 m^3 of hazardous spoil, with additional costs depending on the type of waste.

ENWL therefore wishes to explore opportunities for technologies which can classify waste on-site. The Manufacturing Technology Centre (MTC) has proposed Hyperspectral Imaging as a potential solution and is looking to prove the technology as an option to address the problem statement.

This first phase of work aims to demonstrate a proof-of-concept in a laboratory environment, establishing key contaminants and required sensitivities, and trial selected Hyperspectral Imaging hardware with standardised contaminant samples. The first deliverable of this project defined inspection requirements, while the second reviewed commercially-available hyperspectral cameras against these.

This report constitutes the third deliverable, describing laboratory trials undertaken with Specim hyperspectral cameras in the very-near-, near-, and short-wave-infrared ranges. The trials have produced spectra consistent with that expected from literature, showing the potential for detection of contaminants in spoil. This report also includes recommendations for further trials, and identifies a company, Tellux, for future collaboration.

Contents

1 Introduction

As of March 2023, the Environmental Agency (EA) Regulatory Position Statement 211 (RPS211) applies to businesses that handle excavated waste from unplanned utilities installation and repair works. Under RPS211, small amounts of spoil (≤ 10 m^3, which is approx. 13 tonnes) from unplanned works can be disposed of or recycled as non-hazardous waste. Electricity North West (ENWL) and all fellow utility organisations have used this exemption to dispose of or recycle spoil which cannot be reused as backfill for thousands of emergency works activities across the UK each year.

RPS211 has recently been under review, and will be withdrawn by April 2023 [1]. Withdrawal of this RPs will require companies such as ENWL to classify all of its waste. Without the ability to classify waste on site, they will be required to either dispose of all waste as hazardous, or sort and segregate waste at an offsite location. The Energy Innovation Centre has quoted an average disposal cost of £40 for 10 m^3 for non-hazardous spoil, and £1,200 for 10 m^3 of hazardous spoil, with additional costs depending on the type of waste.

ENWL therefore wishes to explore opportunities for technologies which can classify waste on site. The MTC has offered Hyperspectral Imaging as a potential solution and is looking to prove the technology as an option to address the problem statement. Spoils should be tested for contaminants in line with *Technical Guidance WM3* [2].

The first phase of work aims to produce a proof of concept, establishing key contaminants and required sensitivities, and trialling selected hyperspectral imaging hardware in a laboratory environment with standardised contaminant samples. The first deliverable of this work outlined the inspection required from both performance (e.g. what contaminants must be detected and at what sensitivity) and practical (e.g. size and cost of hardware) perspectives [3]. The second deliverable reviewed commercially-available hyperspectral cameras against the technical requirements defined in D1, based on manufacturers' specifications [4].

This report constitutes the third deliverable, in which three hyperspectral cameras – the Specim FX10, FX17, and SWIR – are used in laboratory trials to image reference materials representing uncontaminated and contaminated soils. Processing of the image data is described, and characteristic features in spectra discussed with respect to detecting and characterising defects.

1.1 Objectives

Objectives of this report are:

- Describe practical trials of hyperspectral imaging and their results, including:
    - Variability in measured spectra;
    - Identification of characteristic features of contaminants and soil types, with correspondence to features identified in literature.
- Describe processing of hyperspectral images and spectra to characterise spectral features;
- Review cameras with respect to technical requirements;
- Provide recommendations for further work towards deploying hyperspectral imaging on worksites.

2 Trial Imaging

2.1 Reference Samples

Hyperspectral images of the soil reference materials were taken with Specim FX10 and FX17 cameras by the MTC and with a Specim SWIR camera by Tellux. As Tellux's facility is based in Petit-Couronne, France, different reference material samples were used. Due to issues with the supply of reference materials, some were available only to the MTC, and some only to Tellux. To compensate for unavailable reference materials, Tellux prepared closest-possible-equivalent samples by doping of the 'clean' reference materials CLNSOIL2 and CLNSOIL5 [5].

Table 1 lists which reference materials were imaged by which cameras. The compositions of Sigma-Aldrich samples are detailed in Appendix A, and those prepared by Tellux in Appendix B.

Table 1: Summary of reference materials scanned by each of Specim FX10, FX17, and SWIR cameras. Materials in *italic green* are those prepared by Tellux.

	FX10	FX17	SWIR
CRM353 – TPH – Sandy Loam	✓	✓	
CRM359 – TPH – Clay Loam	✓	✓	
TX359 – TPH – Clay Loam			✓
CRM141 – PAH – Loamy Clay	✓	✓	
TX141 – PAH – Clay Loam			✓
TX141+ - PAH – Clay Loam			✓
CRM170 – PAH – Clay Soil	✓	✓	
TX170 – PAH – Clay Soil			✓
CRM005 – Trace Metals – Sewage-Amended Soil	✓	✓	✓
SQC001 – Metals in Soil			✓
PB3000 – Lead – Soil	✓	✓	✓
CLNSOIL2 – Clean Clay Loam	✓	✓	✓
CLNSOIL5 – Clean Clay	✓	✓	✓

2.2 Cameras Trialled

Originally it was intended to use near-infrared (NIR) cameras from a number of manufacturers – specifically the Specim FX17 [6], Headwall Micro-Hyperspec Extended VNIR or SWIR 640 [7], and Resonon Pika IR+ or IR-L+ [8, 9], as discussed in Deliverable 2 [4]. However, of those recommended, only the Specim FX17 was found available for trial. As such, the Specim FX 10 was also trialled. Following test with both the Specim FX17and the FX10 [10], a preliminary review of imaged spectra suggested a lack of specificity in identifiable features, related to contaminants, in the near-infrared range.

To address this, an additional camera, operating in the shortwave infrared (SWIR) range, was added into the trials. While more expensive than NIR cameras, imaging in longer wavelengths increases the potential characteristic features that may be captured. To provide the most direct comparison of imaging in NIR and SWIR ranges, the Specim SWIR camera was selected for trial. The Specim SWIR, shown in Figure 1, captures images in 384 wavebands in the 1000 – 2500 nm range, overlapping the range of the FX17 and otherwise similar in functionality and operation [11].

Quantum Design UK, who facilitated hire of the FX10 and FX17, identified Tellux as having the capability to carry out imaging with a Specim SWIR camera as it was unavailable for hire. Tellux are a French startup specialising in hyperspectral imaging of soils for identification and analysis of pollutants [12, 13].

Figure 1: Specim SWIR hyperspectral camera. (Image: Specim)

2.3 Data Acquisition

All hyperspectral images were captured indoors with each sample in a circular container and imaged collectively on a linear stage, as shown in Figure 2 and Figure 3 (MTC; FX10 and FX17) and Figure 4 (Tellux; SWIR). Broad-spectrum (halogen) light sources were placed either side of the camera, illuminating the samples from a high angle. To investigate repeatability, additional images were taken with the FX10 and FX17 cameras, with each sample shaken to redistribute the material. From each image, approx. 4,000–6,000 pixels were sampled for each material.

To account for potential discrepancies in illumination and camera noise, two references are taken for each detector pixel: white and dark. The white reference is a capture of light reflected from a sample with (nominally) uniform reflectance, and allows disparities in illumination to be accounted for. The dark reference is captured with the camera shutter closed. The reflectance spectrum $R(\lambda)$, of each pixel, p, is calculated as

$$R_p(\lambda) = \frac{I_p(\lambda) - D_p(\lambda)}{W_p(\lambda) - D_p(\lambda)}\ , \qquad 1$$

where $I_p(\lambda)$ is the measured intensity of light of wavelength λ, $W_p(\lambda)$ is the white reference, and $D_p(\lambda)$ the dark reference. In this way, any differences between samples as measured by the MTC and by Tellux should be limited to differences between the samples themselves.

Figure 2: Soil samples scanned by FX17 and FX10: conventional digital photograph with annotations (top) and false-colour preview image generated by FX17 (bottom). Wavelengths in false colour image are R = 1573 nm, G = 1273 nm, B = 1001 nm, all in near-infrared.

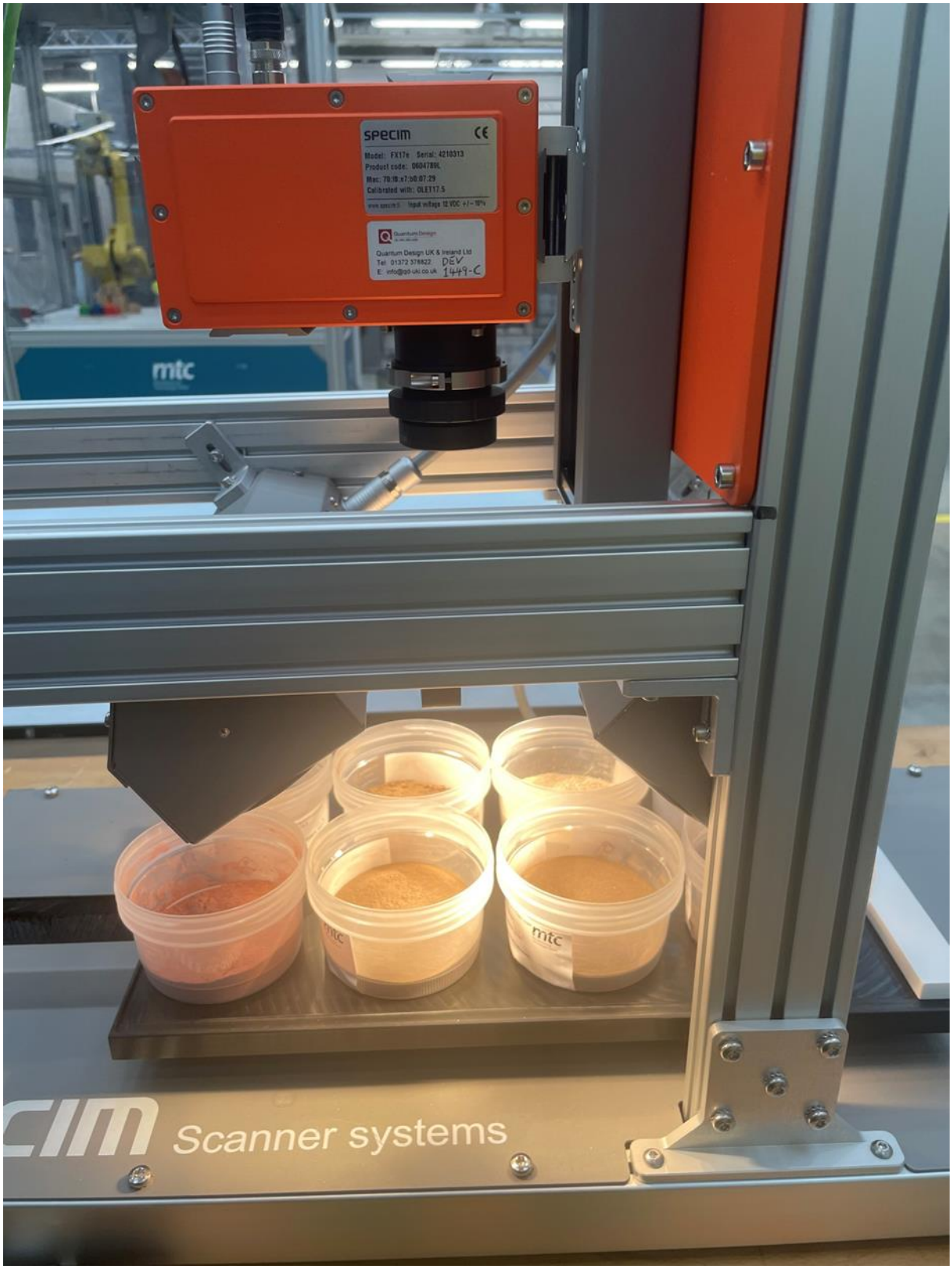

Figure 3: Photo of soil samples (as in Figure 2) being imaged by Specim FX17e camera.

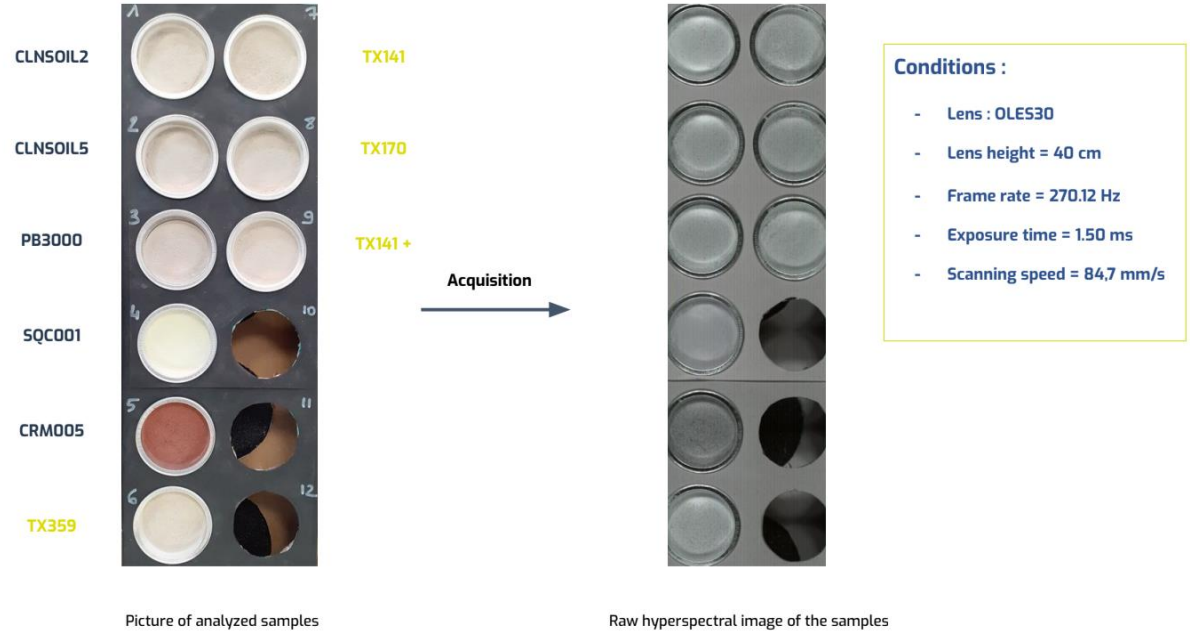

Figure 4: Image from Tellux showing labelled photo of samples (left) and raw hyperspectral (false colour) image (right), along with acquisition settings [5].

The angle of lighting in Specim scanning stage meant sample illumination was partially blocked by and reflected from the containers (visible as shadows and brighter reflections on the surface of the samples. As this would contribute to the measured reflectance, measurements of samples' reflectance were taken only from the central region of each where this contribution is minimal.

2.4 Captured Images

Figure 2 shows the arrangement of soil samples as imaged by the FX10 and FX17 cameras, along with a false-colour image of the same produced by the FX17 camera. Spectra were extracted from regions of this image corresponding to each reference material, as illustrated for an FX10 image in Figure 5. The same is shown for an FX17 image in Figure 6 and for Tellux's SWIR image in Figure 7.

These false-colour images do not match what is seen by eye (or conventional RGB camera) – for example, the sewage-amended soil (CRM005) is visibly red but yellow in the FX10's default false colour profile – the TPH- and PAH-contaminated reference materials produced by Sigma-Aldrich are darker when seen by eye and by FX10 and FX17. Tellux's samples TX141, TX141p, and TX170 are lighter than their counterparts imaged at MTC. This may be due to differences in specific contaminants, in the size and distribution of contaminant droplets, in the concentration of contaminants, and/or in differences in the soil matrices not specified by Sigma-Aldrich.

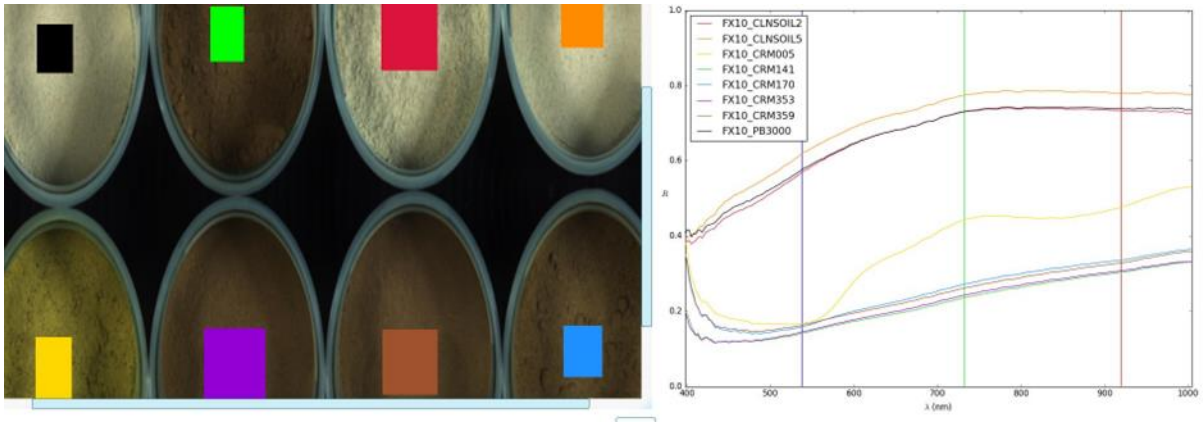

Figure 5: False colour image (left) produced by FX10 camera and mean spectra (right) measured for each reference material. Coloured rectangles in false-colour image indicate sampled regions. Vertical lines in spectra plot indicate wavelengths of R, G, and B channels in false-colour image.

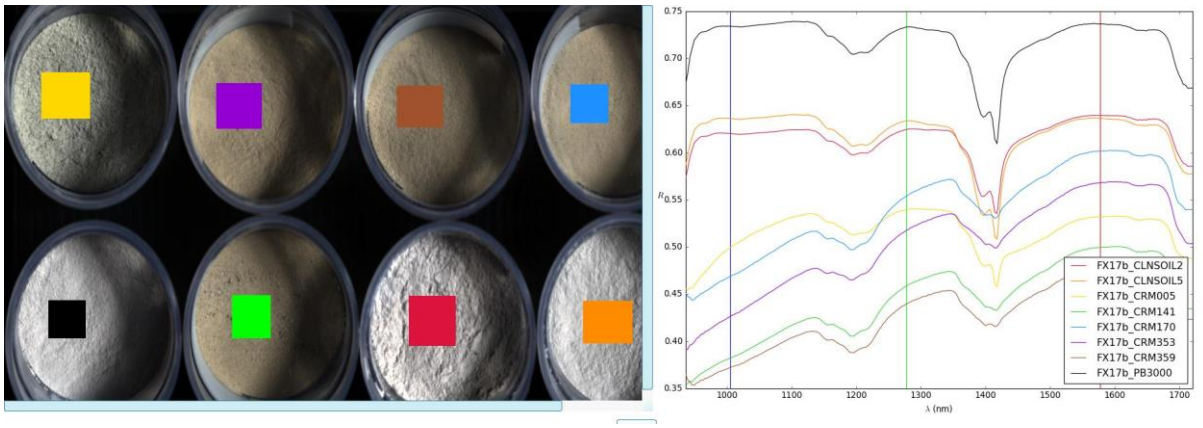

Figure 6: False colour image (left) produced by FX17 camera and mean spectra (right) measured for each reference material. Coloured rectangles in false-colour image indicate sampled regions. Vertical lines in spectra plot indicate wavelengths of R, G, and B channels in false-colour image. Note that samples are not in the same relative positions as in Figure 2 and Figure 5.

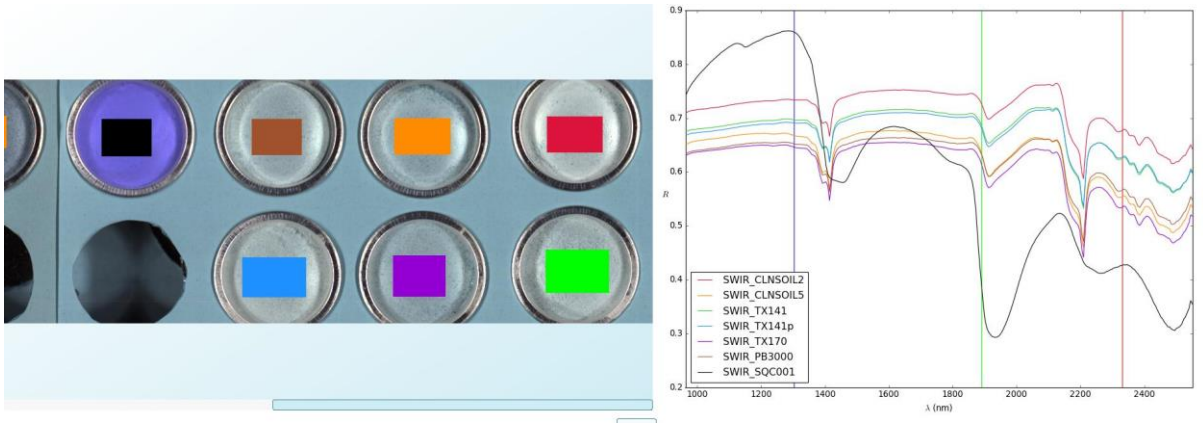

Figure 7: False colour image (left) produced by SWIR camera and mean spectra (right) measured for seven of the nine reference materials (the other two are out of view). Coloured rectangles in false-colour image indicate sampled regions. Vertical lines in spectra plot indicate wavelengths of R, G, and B channels in false-colour image.

2.5 Comparing Images and Confounding Factors

2.5.1 SWIR vs. FX17

The spectral range of the Specim SWIR camera (962–2551 nm) overlaps with that of the Specim FX17 (936– 1720 nm), allowing for direct comparison of the datasets collected by the MTC (FX17) and Tellux (SWIR). Figure 8 shows spectra captured by FX17 and SWIR for clay loam materials. While some variation is expected between the spectra of the contaminated materials, an absorption feature around 1200 nm is seen for all reference materials imaged by FX17 but none imaged by SWIR.

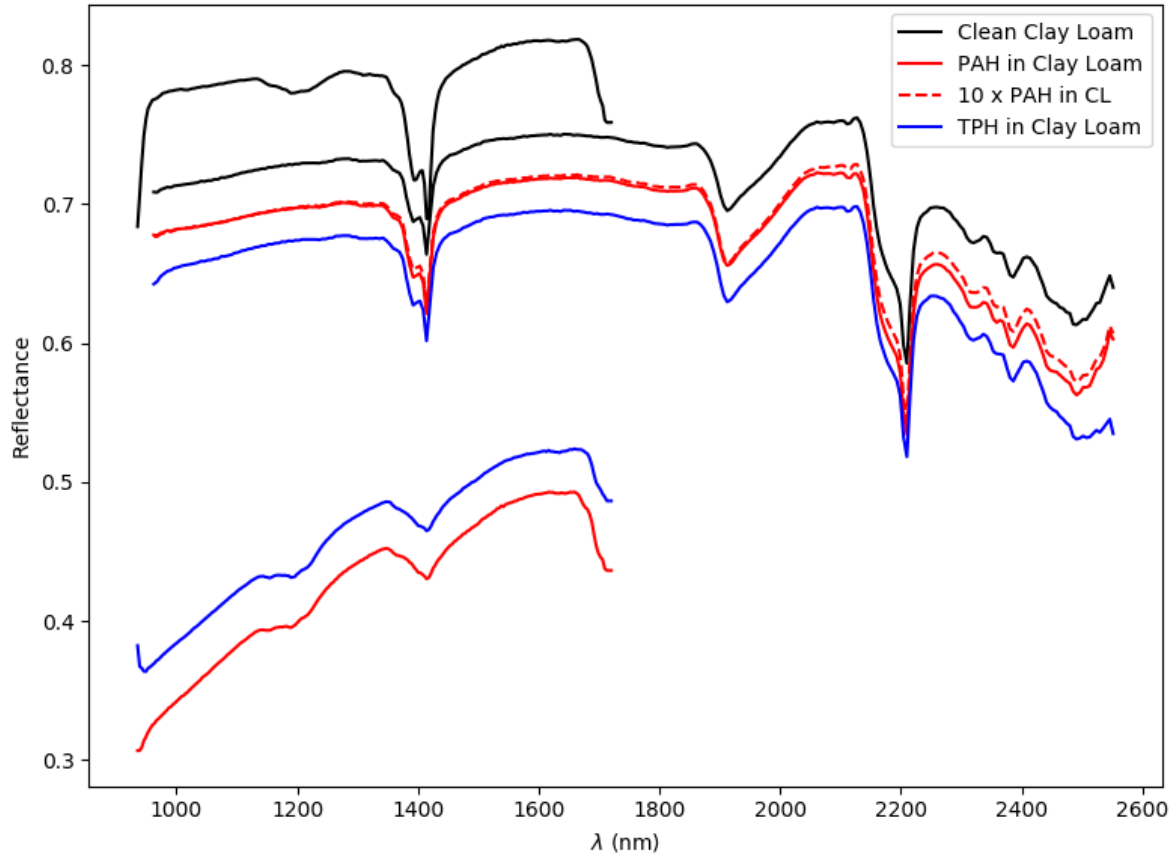

Figure 8: Mean spectra captured by Specim FX17 (low half of wavelength range) and SWIR (full range) cameras for selected reference materials: clean clay loam (CLNSOIL2), PAH in clay loam (CRM141, TX141, and TX141+), and TPH in clay loam (CRM359 and TX359).

Sampling reflectance spectra from the plastic (FX10, FX17) and glass (SWIR) containers reveals that absorption around 1200 nm is a feature of the plastic containers, shown in Figure 9. Absorption around 1200 and 1400 nm correspond to C–H peaks, with the latter visible in both plastics and soils [14]. Specifically, the measured spectrum of the plastic containers is consistent with that of polypropylene, shown in Figure 10.

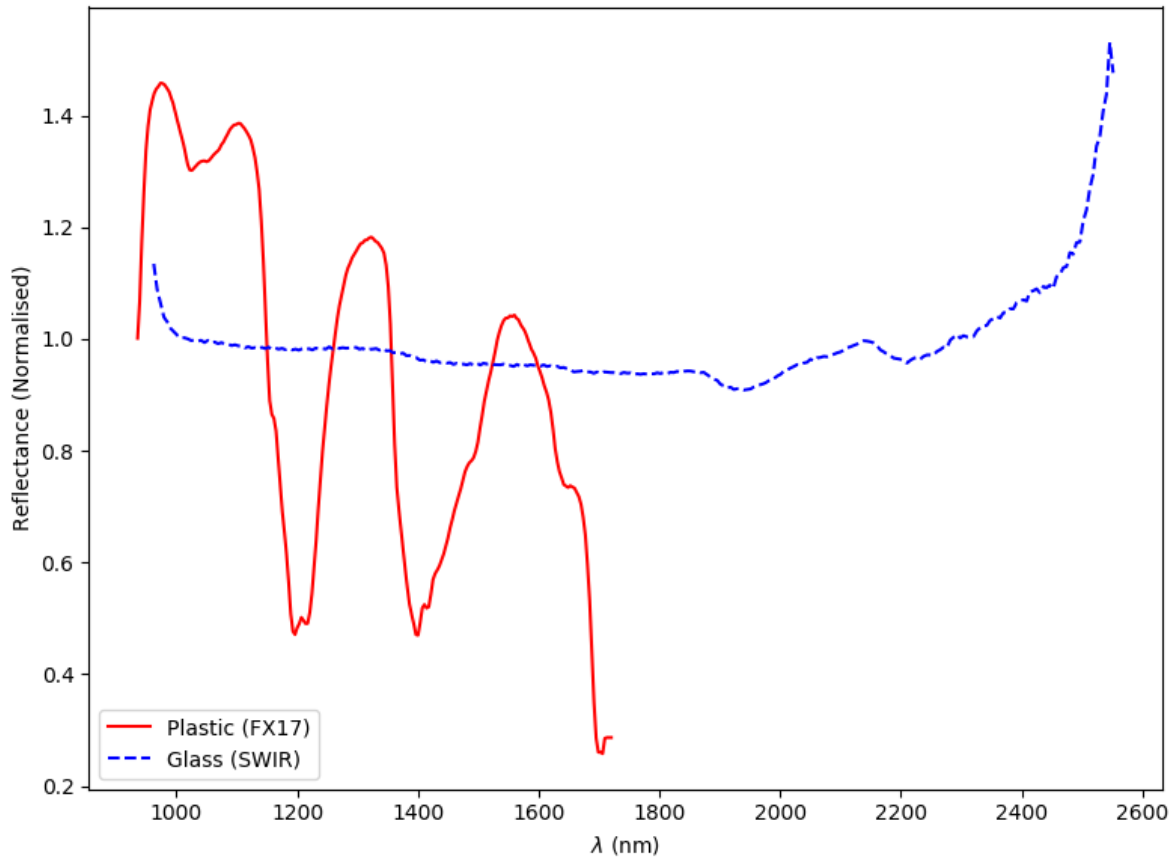

Figure 9: Normalised reflectance spectra from plastic containers imaged by FX17 camera and glass containers imaged by SWIR.

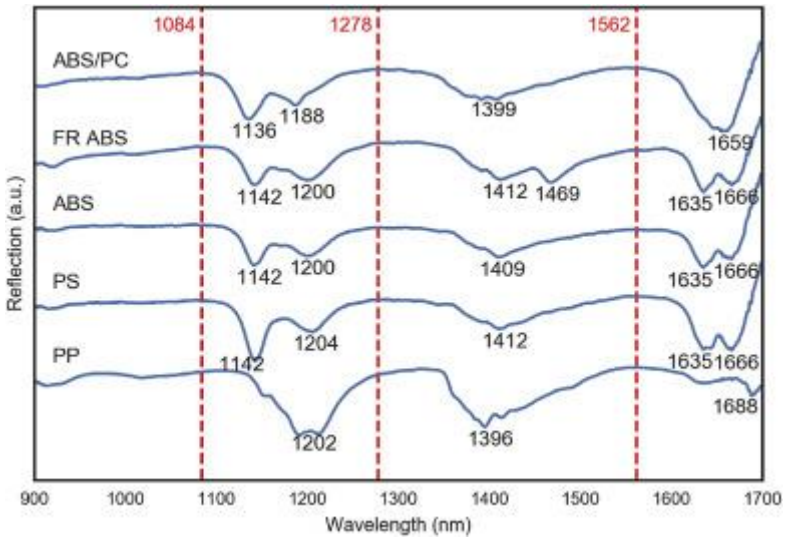

Figure 10: Reflectance spectra of common household plastics: ABS/polycarbonate blend (ABS/PC), flame-retardant ABS (FR ABS), ABS, polystyrene (PS), and polypropylene (PP). Figure from [15].

2.5.2 Repeat Imaging

Figure 11 shows the mean, normalised reflectance spectra for CLNSOIL2 and CRM141 materials captured with the FX17 in two imaging sessions. While the same features can be seen in each, their relative magnitude differs significantly. The biggest change for a single feature is the 1200 nm absorption corresponding to the plastic containers, despite the same lighting being used for both images.

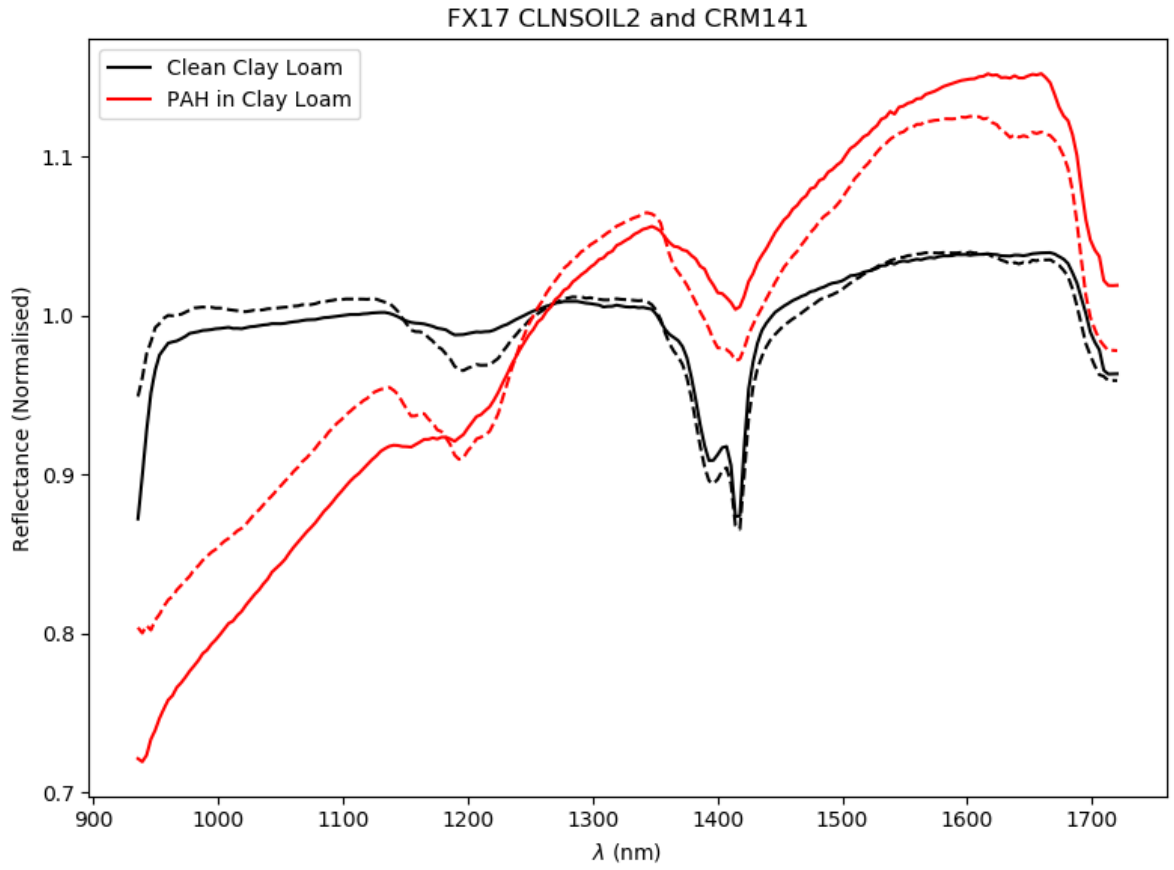

Figure 11: Mean, normalised reflectance spectra from clean clay loam (CLNSOIL2) and PAH in clay loam (CRM141) from first (solid line) and second (dashed line) imaging sessions.

The repeatability (i.e. consistency) in spectra between images should be considered with respect to the variation within each image. Figure 12 compares the same spectra but with the range ofone standard deviation ($\pm 1\sigma$) highlighted around the first spectra. Assuming the difference in 1200 nm absorption is due to a difference in lighting conditions, the variation between images is less than $\pm 1\sigma$ for much of the CLNSOIL2 spectra, but not for CRM141. A potential explanation for the variation in CRM141 is evaporation of volatile aromatic hydrocarbons, particularly with heating from the halogen lights. It would also be consistent in a loss of humidity in the CRM141 sample, as infrared absorption by water tends to increase with wavelength.

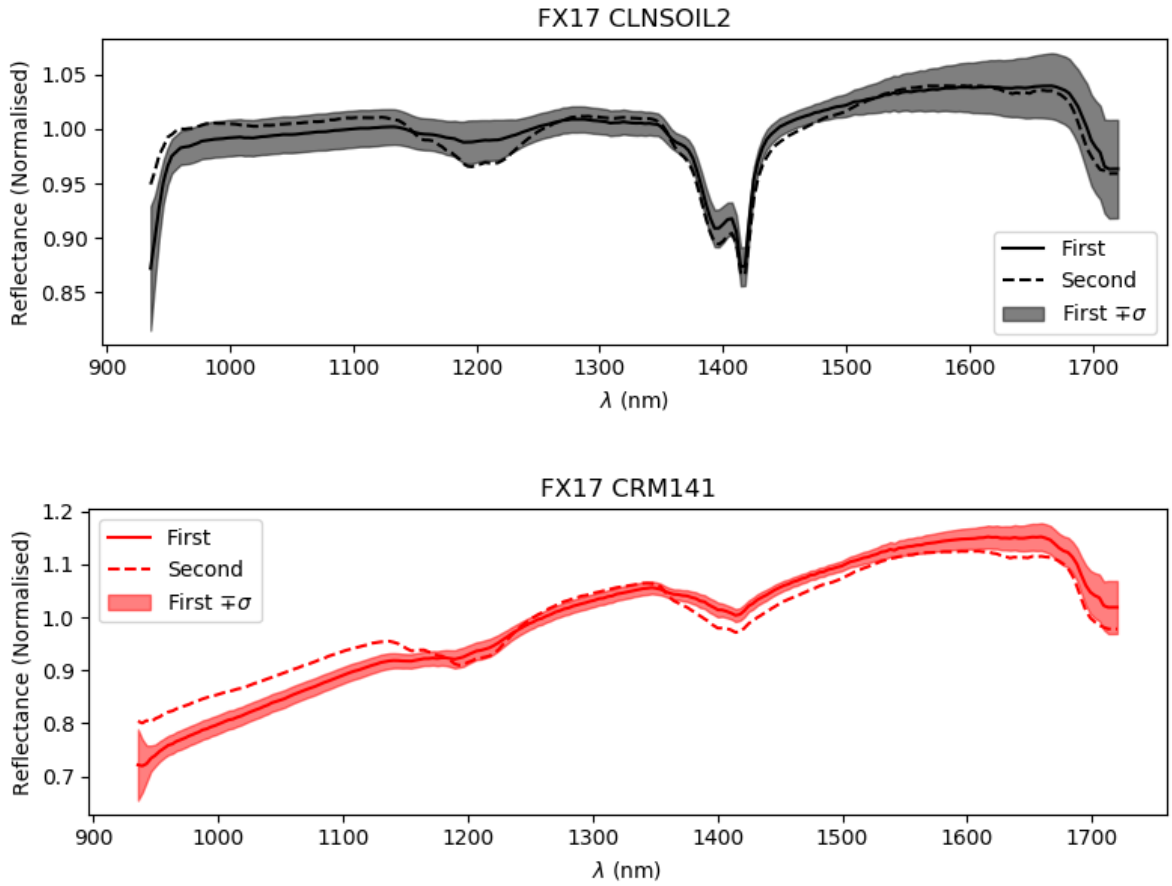

Figure 12: Mean, normalised reflectance spectra from clean clay loam (CLNSOIL2; top) and PAH in clay loam (CRM141; bottom) from first (solid line) and second (dashed line) imaging sessions, with shaded area around first spectrum indicating the standard deviation in pixels' spectra.

2.6 Trends in Spectra

Viewing mean spectra for each material, as in Figure 5, Figure 6, and Figure 7, a number of local features can be seem in common between most materials, particularly around 1200 nm (FX17), 1400 nm (FX17 and SWIR), and 2200 nm (SWIR). The most obvious differences, listed in Table 2, are generally on a larger scale. Overall, the nominally-clean and lead-contaminated soils have similar spectra, as do the TPH- and PAH-contaminated soils. The other metal-contaminated soils, CRM005 and SQC001, have spectra that diverge significantly from the others, though without a priori information it is not apparent whether this is due to the metal contaminants or unspecified differences in their soil matrices (or what combination of both).

Table 2: Summary of broad features in mean spectra for each material in comparison to general trend/other materials. Shading indicates materials with features in common.

	FX10	**FX17**	**SWIR**
General Trend	Increasing R with λ.	Absorption features around 1200 and 1400 nm.	Absorption features around 1400, 1900, and 2200 nm.
CLNSOIL2 Clean Clay Loam	Fast increase in R with λ, plateauing around 800 nm.	Overall uniform R.	*As general.*
CLNSOIL5 Clean Clay	Fast increase in R with λ, plateauing around 800 nm.	Overall uniform R.	*As general.*
CRM005 Trace Metals Sewage-Amended Soil	Far more varied, with broad absorption around 500 nm and 800 nm.	Initially increasing R with λ, uniform or slight decrease after 1300 nm.	Broader absorption around 1400 and 1900 nm.
CRM141, TX141 PAH – Loamy Clay	Slower increase in R with λ, with no plateau.	Increasing R with λ.	*As general.*
CRM170, TX170 PAH – Clay Soil	Slower increase in R with λ, with no plateau.	Increasing R with λ.	*As general.*
CRM353 TPH – Sandy Loam	Slower increase in R with λ, with no plateau.	Increasing R with λ.	*n/a*
CRM359, TX359 TPH – Clay Loam	Slower increase in R with λ, with no plateau.	Increasing R with λ.	*As general.*
PB3000 Lead – Soil	Fast increase in R with λ, plateauing around 800 nm.	Overall uniform R.	*As general.*
SQC001 Metals in Soil	*n/a*	*n/a*	Decreasing R with λ. Much broader absorption features. Reflectance feature around 2350 nm.

3 Processing

3.1 Normalisation and Small-Scale Features

While large-scale spectral features can be indicative of a number of factors (e.g. greater R with λ may correspond to greater water content in an imaged object), characterisation of specific chemicals often utilises narrower features that correspond to chemical composition (e.g. bonds within molecules). To increase the visibility of these features, both to a human observer and various processing techniques for automated classification, spectra are typically pre-processed to remove large-scale features.

Some figures earlier in this report refer to normalised reflectance. While the white reference in an image is used to account for spectral features of light sources (see Equation 1, pg. 6), the relative *amount* of this illumination (e.g. shadows) affects the overall intensity of light reflected. Normalising the reflectance spectra of each pixel removes this effect. The normalised spectra, $R_{n,p}(\lambda)$, for a pixel, p, is calculated as

$$R_{n,p}(\lambda) = \frac{R_p(\lambda)}{\sqrt{\sum_{i=o}^{N} R_p^2(\lambda_i)}}\,, \qquad \textbf{2}$$

where $R_p(\lambda_i)$ is the pixel's measured reflectance for waveband λ_i and N the total number of wavebands in the measured spectrum. This is sometimes referred to as the *spectral angle*, as it is equivalent to the angle of a spatial vector. As the number of wavebands measured by each camera differs, for comparison of FX17- and SWIR-measured spectra the normalisation is adjusted to

$$R'_{n,p}(\lambda) = \frac{R_p(\lambda)\sqrt{N}}{\sqrt{\sum_{i=o}^{N} R_p^2(\lambda_i)}}\,. \qquad \textbf{3}$$

Normalising the spectra shown previously in Figure 8 (pg. 11) yields the spectra shown in Figure 13. With the overall differences in illumination removed, the relative scale of different features is more accurately represented.

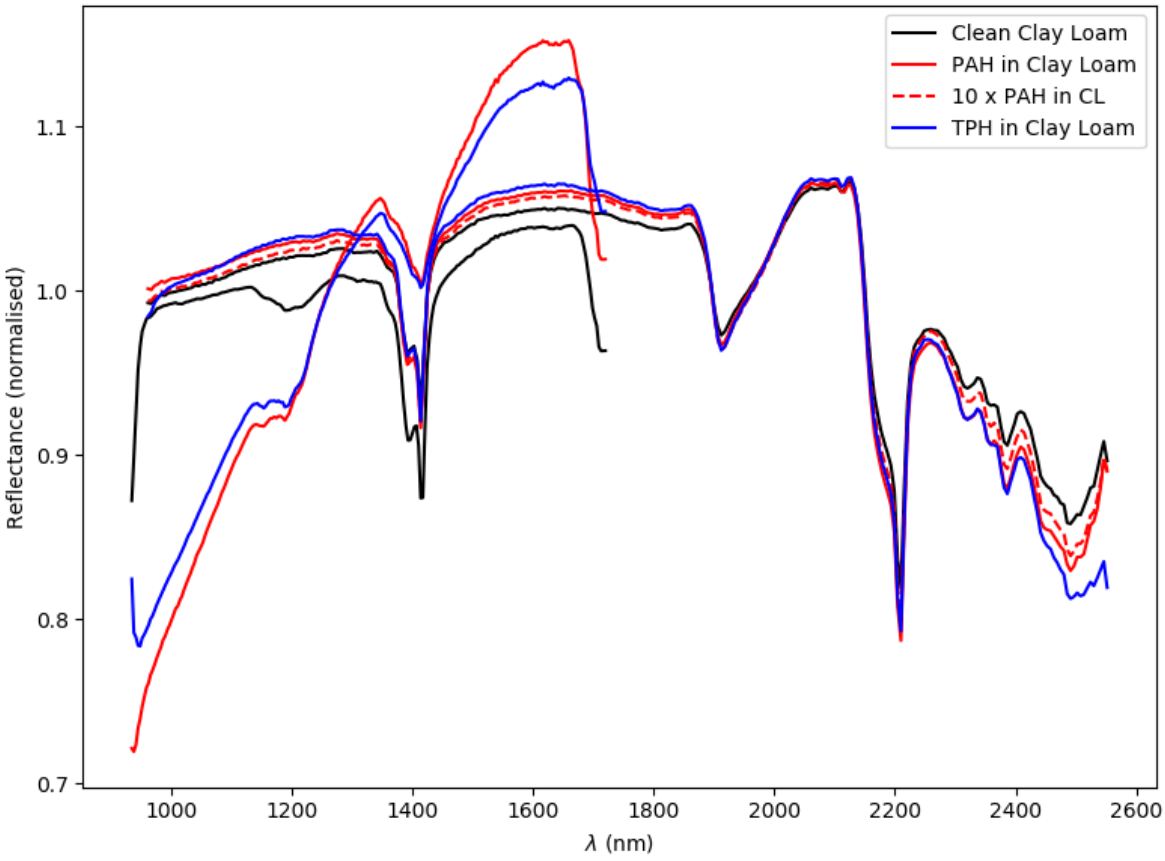

Figure 13: Normalised mean spectra captured by Specim FX17 (low half of wavelength range) and SWIR (full range) cameras for selected reference materials: clean clay loam (CLNSOIL2), PAH in clay loam (CRM141, TX141, and TX141+), and TPH in clay loam (CRM359 and TX359).

3.2 Continuum Removal

Visibility of small-scale features can be further improved by a process of *continuum removal* [16]. This process approximates the large-scale features of the spectrum by its *convex hull*, the outline of local maxima, an example of which is illustrated in Figure 14. This assumes all small-scale features of interest are absorption features, and the continuum-derived spectra $R'_{CR}(\lambda)$ is calculated by dividing the normalised spectra by the convex hull, $C(\lambda)$:

$$R'_{CR}(\lambda) = \frac{R'_n(\lambda)}{C(\lambda)} \ . \qquad (4)$$

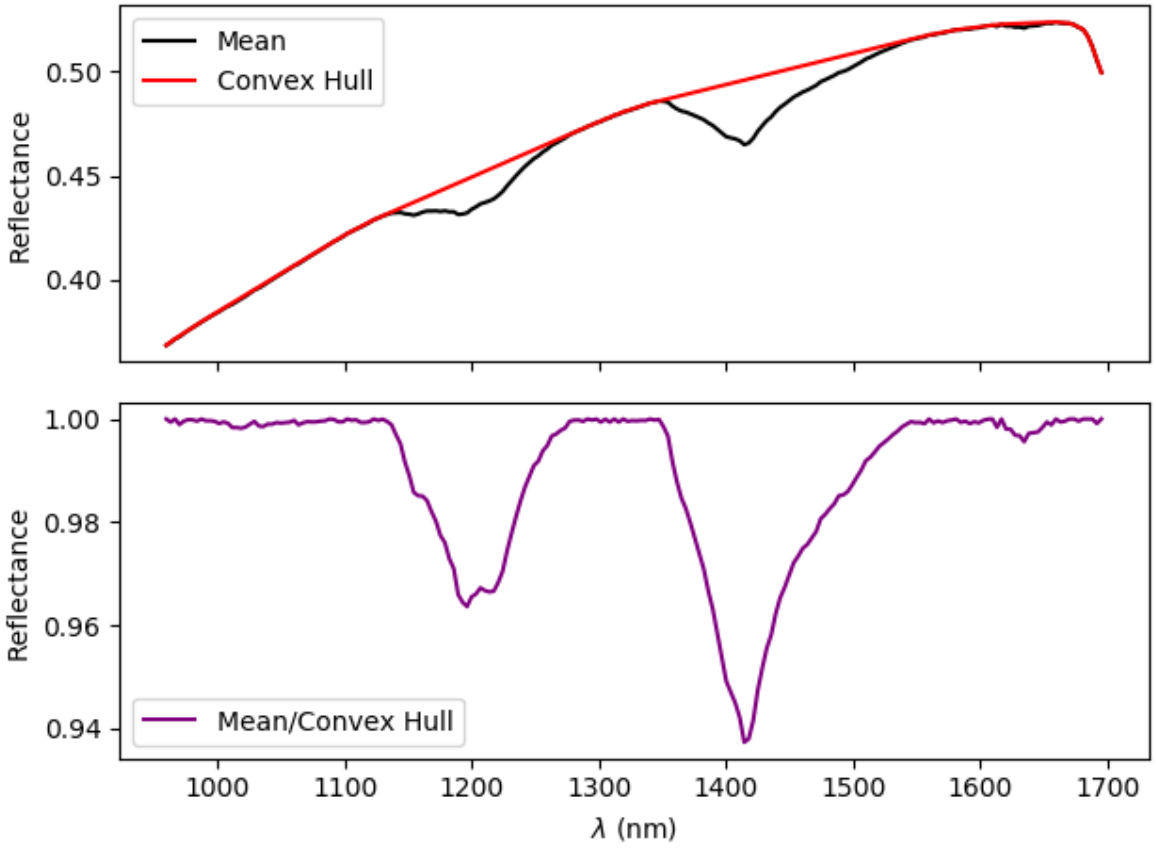

Figure 14: Illustration of continuum removal process: normalised mean reflectance spectrum of CRM141 from FX17 (top; black), its convex hull (top; red), and the continuum-removed spectrum (bottom; purple).

Figure 15 and Figure 16 show continuum-removed spectra for all materials imaged by FX17 and SWIR cameras respectively. Larger-scale figures of continuum-removed spectra for individual reference materials are included in Appendix C.

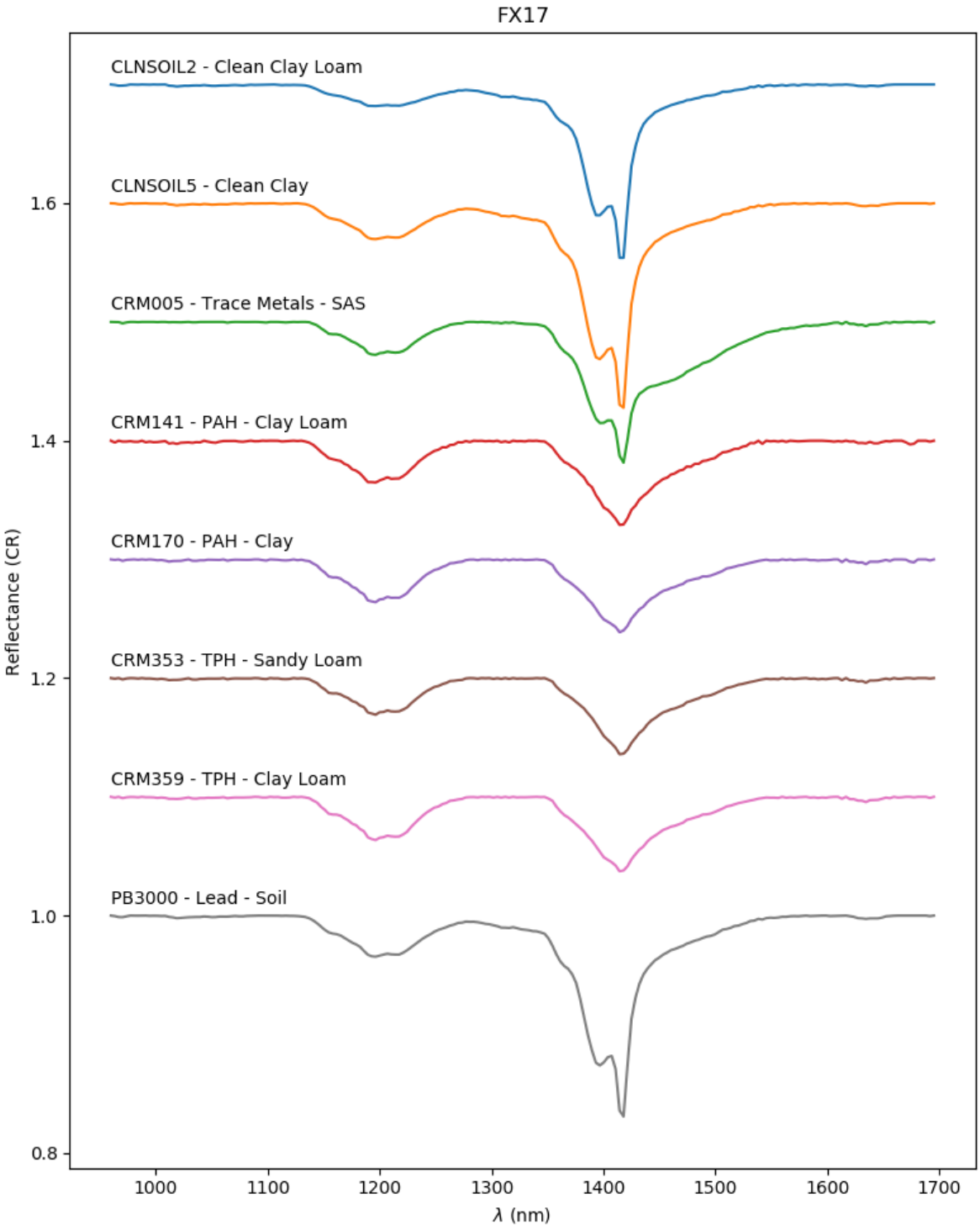

Figure 15: Continuum-removed spectra for reference materials imaged by Specim FX17 camera. Each spectrum is offset in y to visually separate spectra.

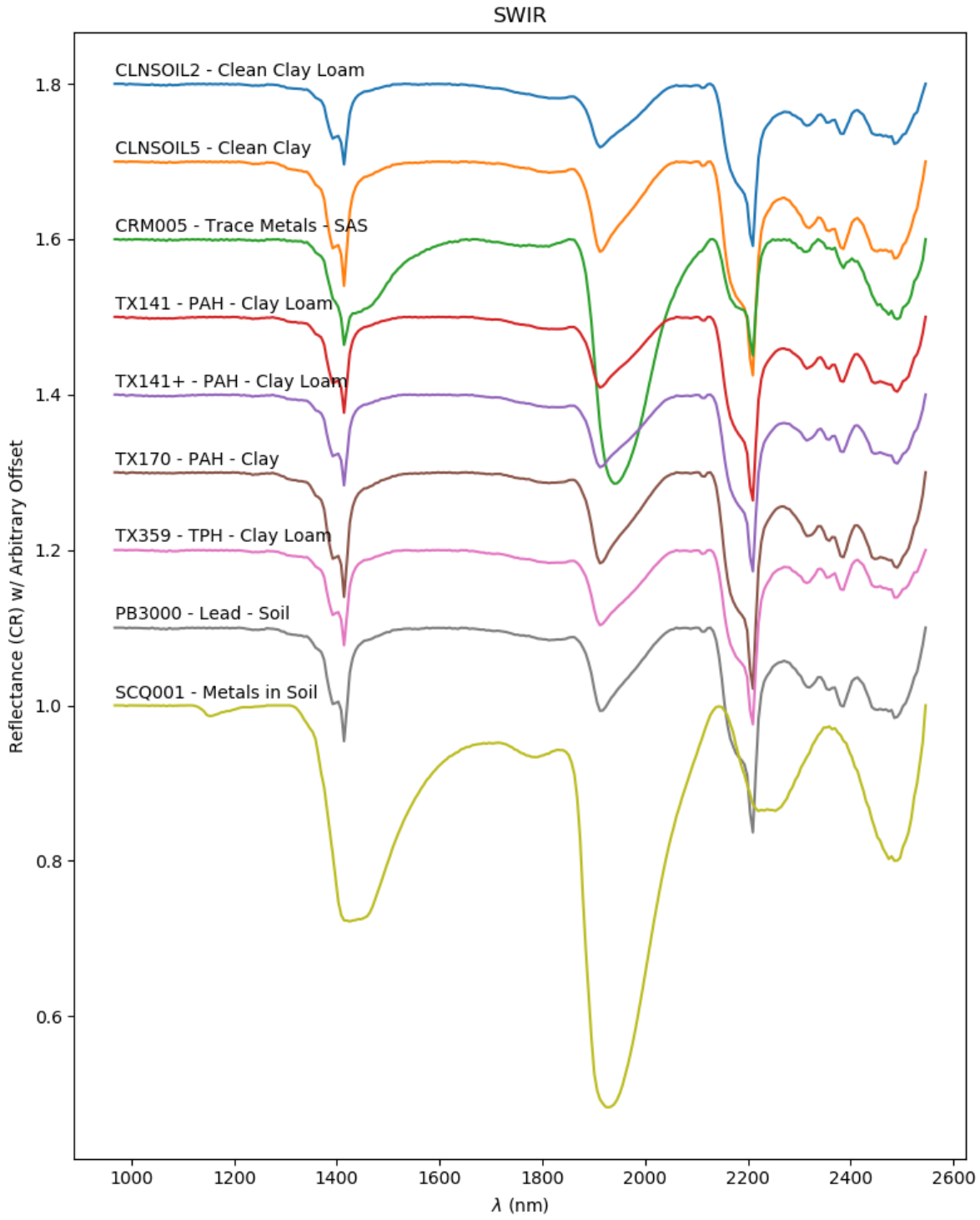

Figure 16: Continuum-removed spectra for reference materials imaged by Specim SWIR camera. Each spectrum is offset in y to visually separate spectra.

3.3 Variation and Inhomogeneity

Variation within an image – i.e. between reflectance spectra measured at each pixel – gives an indication of the inhomogeneity of an imaged sample. While some variation may be expected even in a perfectly homogeneous sample (e.g. due to detector noise, optical artefacts, variations in lighting, etc.), for an inhomogeneous material there will be additional variation in reflectance of wavelengths of characteristic features.

One method of quantifying this variation is by calculating the standard deviation of the reflectance over an image. It might be assumed that, all else being equal, the standard deviation would be proportional to reflectance. However, this is not the case, as shown in Figure 17. While some "local" trends can be seen – both standard deviation and reflectance varying continuously over certain ranges – there is no overall trend save a weak negative correlation (ρ = -0.38). That is, standard deviation is generally greater for wavelengths with less mean reflectance (i.e. for absorption features).

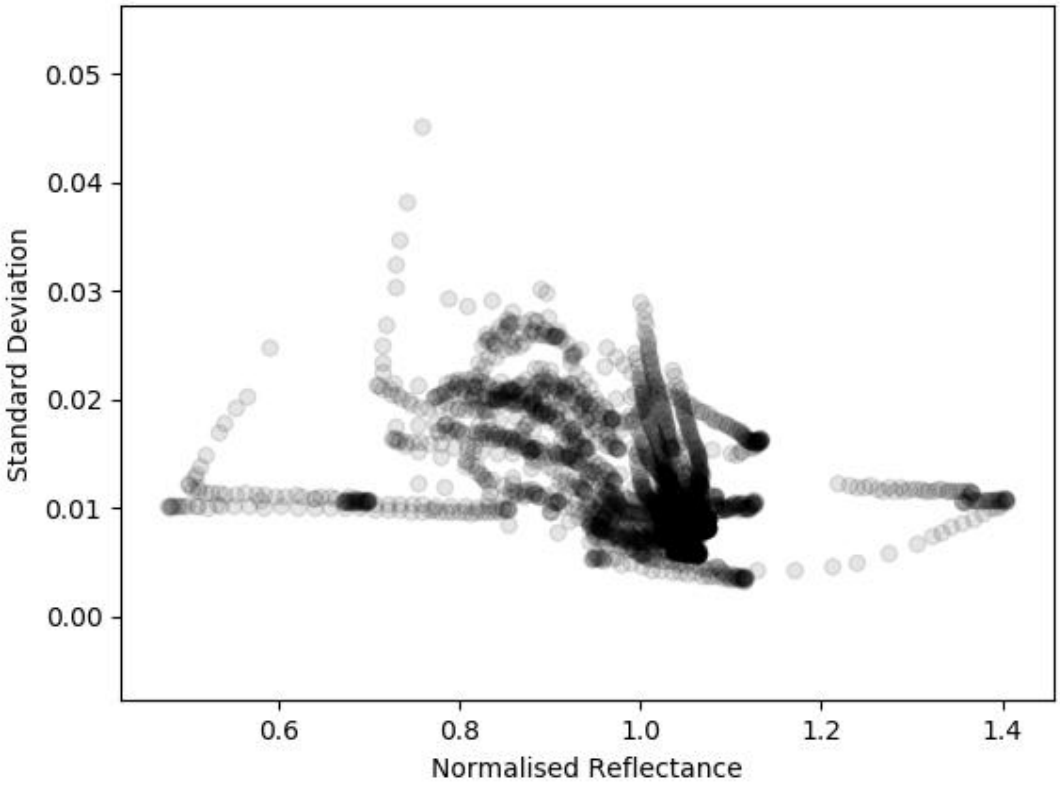

Figure 17: Standard deviation vs. normalised reflectance for all samples imaged by Specim SWIR camera. Each point represents standard deviation and normalised reflectance of one wavelength for one sample. Each point has an alpha of 0.1, hence darker regions in plot indicate multiple, overlapping points.

Figure 18 and Figure 19 show the standard deviation of reflectance for samples imaged by Specim FX17 and SWIR cameras respectively. In addition to being greater at reflectance features (for most samples, around 1200, 1400, 1900, and 2200 nm), standard deviation is greater towards the limits of each camera's spectral range (due to lower sensitivity, and therefore greater influence of noise). What can be seen comparing samples is that the range of standard deviations around features of interest is roughly the same for all samples, indicating that contaminated samples are approximately as homogeneous as 'clean' samples.

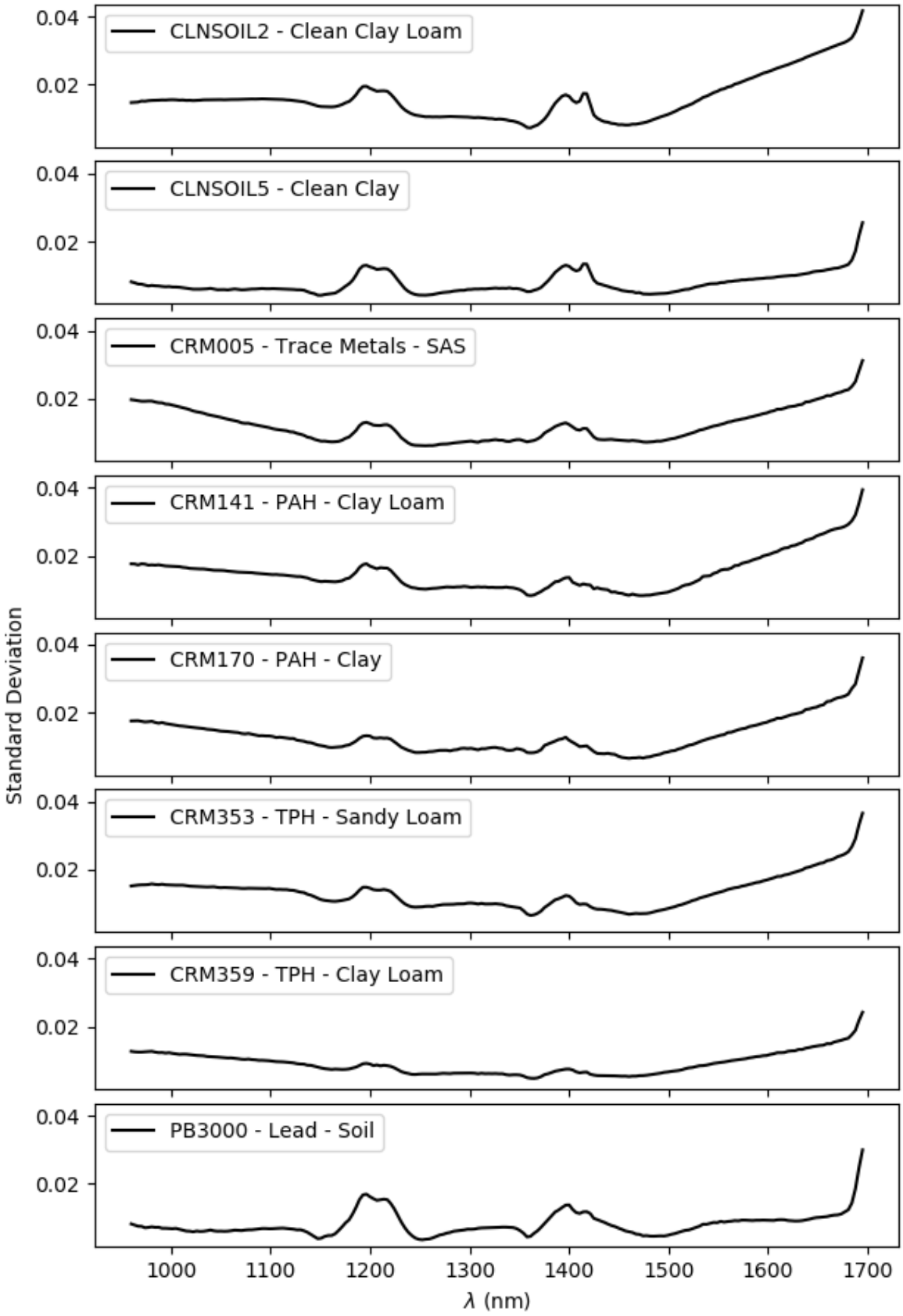

Figure 18: Standard deviation of reflectance for samples imaged by Specim FX17 camera. Note that *y*-axes have identical scaling.

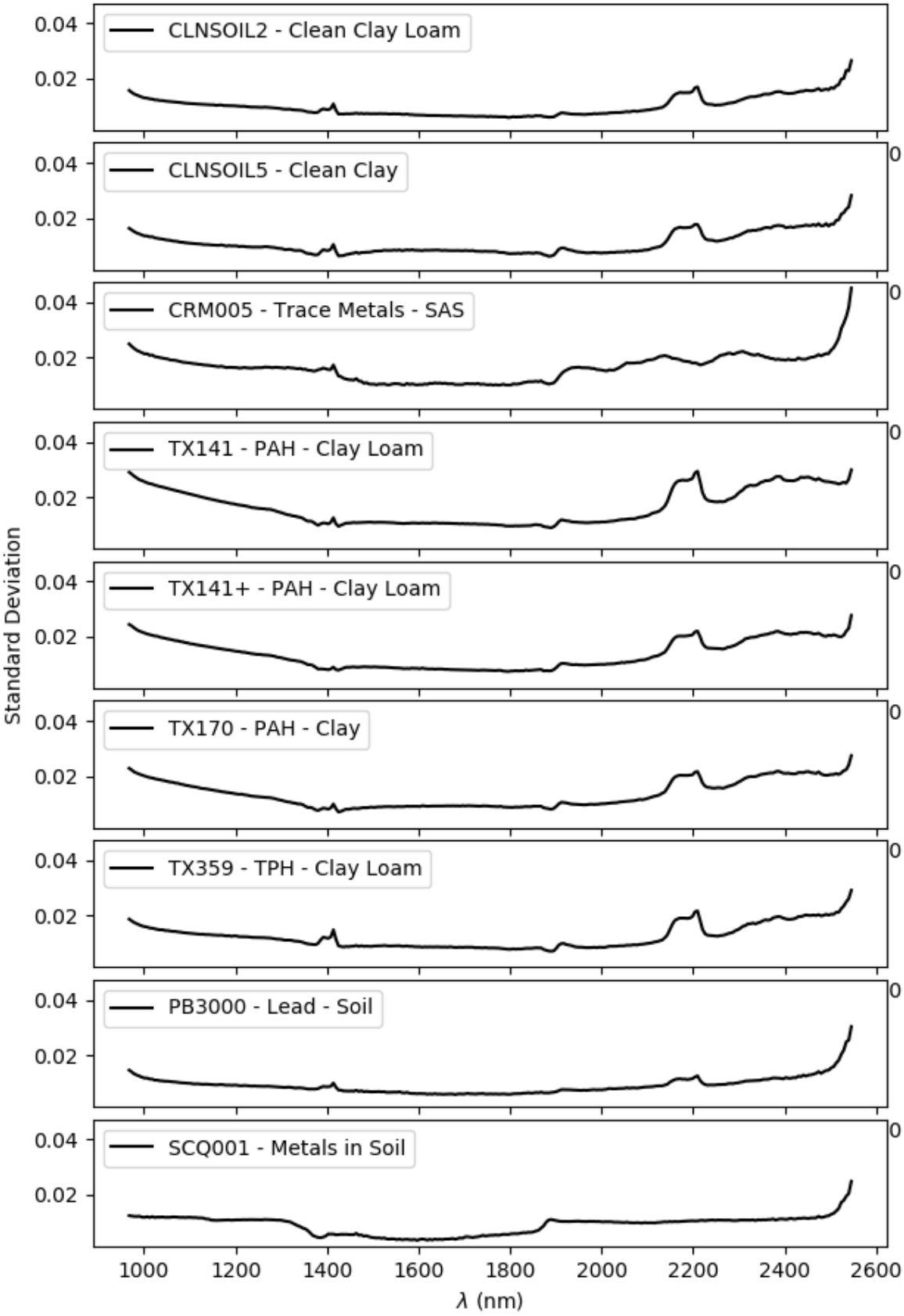

Figure 19: Standard deviation of reflectance for samples imaged by Specim SWIR camera. Note that *y*-axes have identical scaling.

3.4 Correlation

The relative abundances of major components between each reference material are given in Table 3. For the purpose of this analysis, all petroleum hydrocarbons (TPH), polycyclic aromatic hydrocarbons (PAHs, representing coal tar), and heavy metal compounds ("metals") are considered collectively, while lead (II) nitrate and nickel dinitrate hexahydrate (NDH) are considered separately as they are the only listed contaminant in PB3000 and CRM005 respectively. For soils of specified texture, relative abundance of clay, silt, and sand was determined from the geometric centre of each, e.g. the centre of the 'clay loam' region is 34% clay, 34% silt, and 32% sand [17, 18].

For soil components, the relative abundance is given as the percentage by mass of the reference material. For contaminants, the scale is arbitrary; this does not affect analysis. For example, the nominal concentration of NDH in CRM005 is < 0.01%, but its relative abundance is given as 1. For Sigma-Aldrich materials, the median listed concentration has been used to derive these values. For the Tellux-prepared materials, the values in Table 9 have been used [5].

Table 3: Relative abundances of major components of each reference material. Materials in *italic green* are those prepared by Tellux. Asterisks indicate unknown quantities where soil texture is not specified.

	TPH – Sandy Loam	**TPH – Clay Loam**	***CLNSOIL2 + TPH***	**PAH – Loamy Clay**	***CLNSOIL2 + PAH***	***CLNSOIL2 + PAH***	**PAH – Clay Soil**	***CLNSOIL5 + PAH***	**Trace Metals – Sewage Amended Soil**	**Metals in Soil**	**Lead – Soil**	**Clean Clay Loam**	**Clean Clay**
	CRM353	CRM359	*TX359*	CRM141	*TX141*	*TX141+*	CRM170	*TX170*	CRM005	SQC001	PB3000	CLNSOIL2	CLNSOIL5
Soil	99	99	100	99	100	98	99	100	100	100	99	100	100
of which Clay	10	34	34	34	34	34	60	60	*	*	*	34	60
of which Silt	25	34	34	34	34	34	20	20	*	*	*	34	20
of which Sand	64	32	32	32	32	32	20	20	*	*	*	32	20
Clay	10	34	34	34	34	34	60	60	*	*	*	34	60
Silt	25	34	34	34	34	34	20	20	*	*	*	34	20
Sand	64	32	32	32	32	32	20	20	*	*	*	32	20
TPH	10	10	1	0	0	0	0	0	0	0	0	0	0
PAHs	0	0	0	5	1	10	5	1	0	0	0	0	0
Metals	0	0	0	0	0	0	0	0	1	100	10	0	0
Lead (II) nitrate	0	0	0	0	0	0	0	0	0	1	10	0	0
NDH	0	0	0	0	0	0	0	0	1	1	0	0	0

The relationship between these abundances and the (continuum-removed) spectra can be quantified using the *correlation coefficient*. For a pair of variables, x and y, the correlation coefficient, $\rho_{x,y}$, is defined as

$$\rho_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}, \quad (5)$$

where $\bar{x}$ is the mean value of x and σ_x its standard deviation. Correlation here is calculated between two sets of variables: the relative abundances and continuum-removed reflectance values at each wavelength. Calculated correlation coefficients from FX17 and SWIR images are shown in Figure 20 and Figure 21 respectively.

In NIR (imaged by FX17), the correlation of reflectance to the presence of TPH and PAH are very similar. This is consistent with both being a mix of hydrocarbons. As C–H bonds have multiple absorption lines around 1400 nm, the variety of contributions makes differences between the two families of hydrocarbons difficult to distinguish in this range of wavelengths. The disordered correlation to metal is likely amplified noise, with both CRM005 (trace metal in sewage-amended soil) and PB3000 (lead (II) nitrate in soil) having very low concentrations of their respective contaminants. The broader features seen in CRM005's spectrum likely correspond to the sewage-amended soil matrix itself.

TPH and PAH are far more distinct in their SWIR-imaged spectra, not just at higher wavelengths but around 1400 nm. The very narrow features at 2200 nm and 2207 nm correspond to known spectral features of petroleum hydrocarbons [19]. Metals – both as a category and lead (II) nitrate and NDH individually – also have more pronounced differences in correlated features, with clear, broad features visible in the range 1900–2100 nm, similarly consistent with literature [20].

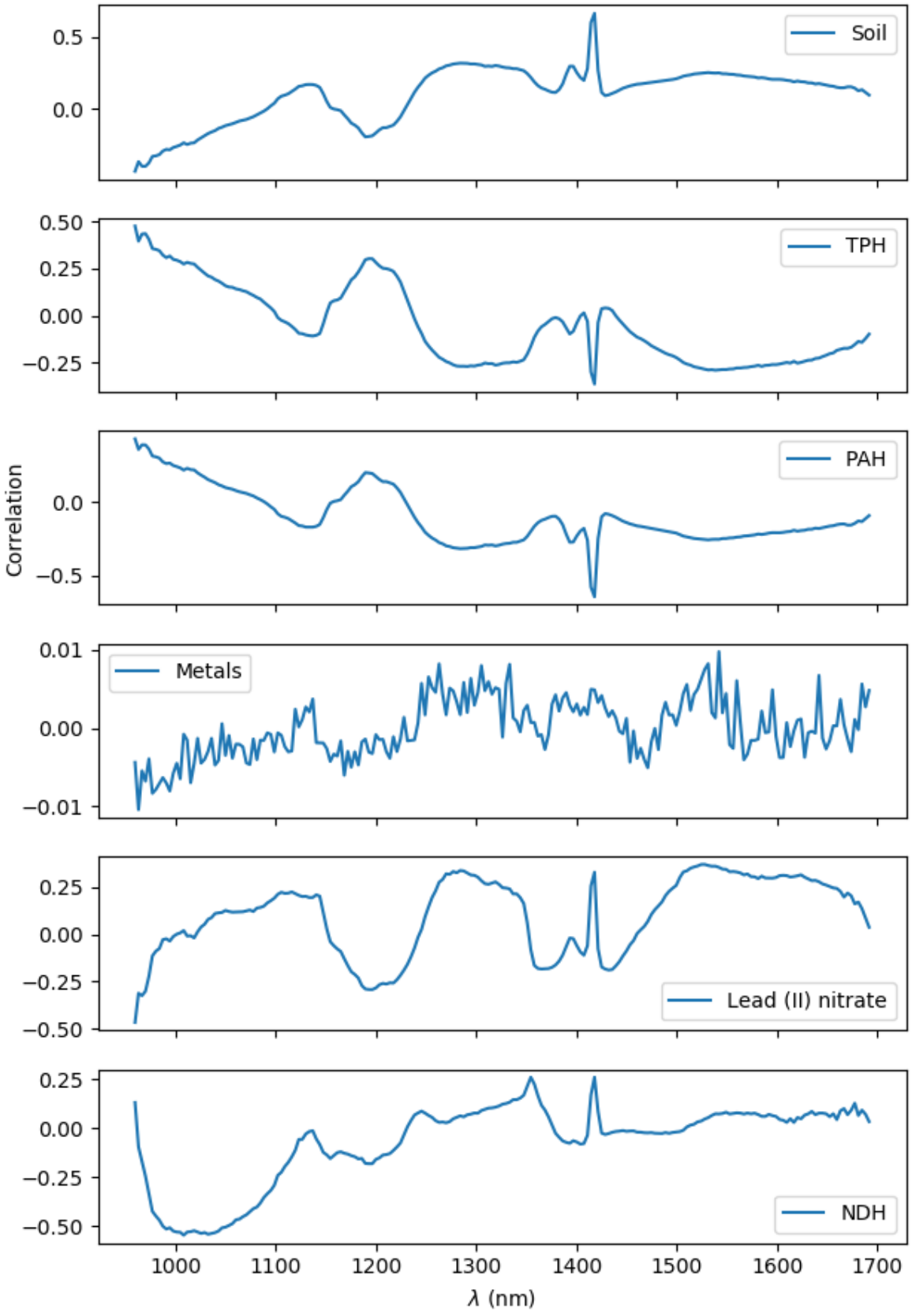

Figure 20: Correlation between reflectance values at various wavelengths and abundance of major components, from FX17 images.

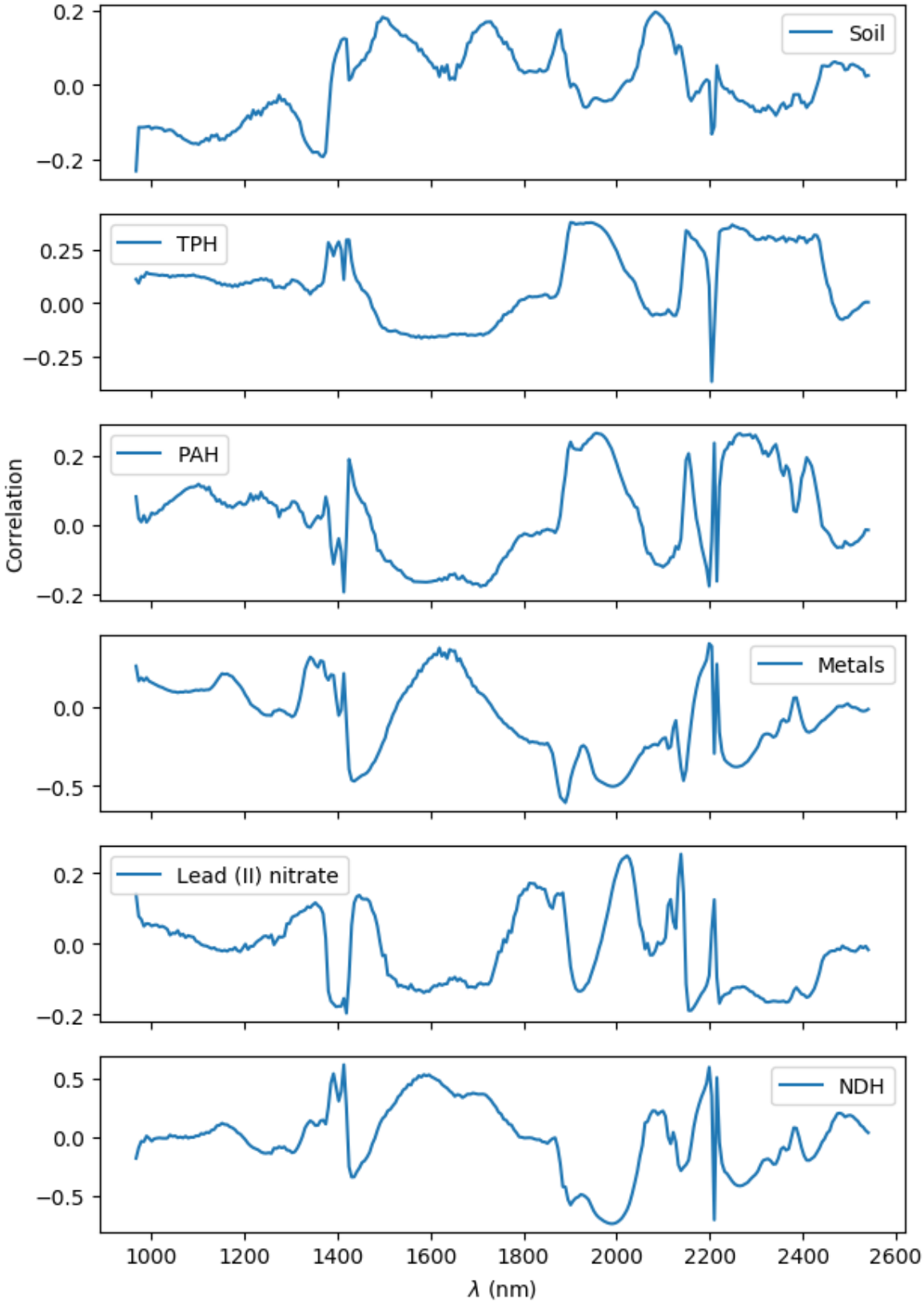

Figure 21: Correlation between reflectance values at various wavelengths and abundance of major components, from SWIR image.

4 Discussion

4.1 Comparison to Literature

While a number of spectral features observed are consistent with those expected from literature, there are nonetheless significant differences in measured spectra and the results of analysis. For example, Figure 22 shows a comparison of the spectra of PAH-contaminated and uncontaminated soil as measured by Douglas et al. (2017) [19] and as measured in this project. Douglas' et al.'s soil samples are sandy clay loam and clay loam from sites in the Niger Delta, with some contaminants in common with Tellux's TX141 and TX141+ (specifically: acenaphtylene, fluorene, phenanthrene, and pyrene). The overall mean concentration in Douglas' et al.'s samples is 0.65%, compared to 0.24% and 2.39% in TX141 and TX141+ respectively.

Of six characteristic features identified by Douglas et al., three are in common between both sets of samples: 1415, 1914, and 2200 nm, though the relative strength of the 2200 nm feature is far greater in our data. Conversely, absorption features at 1712 and 1758 nm and a reflectance feature at 2207 nm are visible in PAH-contaminated samples from [19] but not in our own.

This may be due to PAHs present in Nigel Delta soil but not TX141/+, such as benzo[a]pyrene, but absorption around 1700 nm is common to many polycyclic aromatic hydrocarbons, including acenaphtylene, corresponding to C—H bond bending [21].

4.2 Specificity

Classification of spoil should be capable not just of distinguishing contaminated from uncontaminated spoil, but one contaminant from another. This is particularly important where different contaminants present different hazards, and hence have different permissible concentrations. This is made more challenging where contaminants are not expected to exist in isolation.

While some similarity is expected in correlations etc. – as the presence of *any* contaminants and their spectral features is also the partial absence of soil and its features – the observed effects of TPH and PAH contamination are markedly similar in both NIR and SWIR. While even similar chemicals have differing spectral features at sufficient resolution, it does not appear possible to identify a specific chemical (e.g. benzo[a]pyrene) with NIR—SWIR hyperspectral cameras, at least at the concentrations examined here.

Hyperspectral imaging may have greater utility in identifying TPH and/or PAH contamination *in general* than in identifying specific compounds and their concentrations within a mixture. Future work should explicitly consider the separability of contaminants from "benign" materials other than the soil matrix.

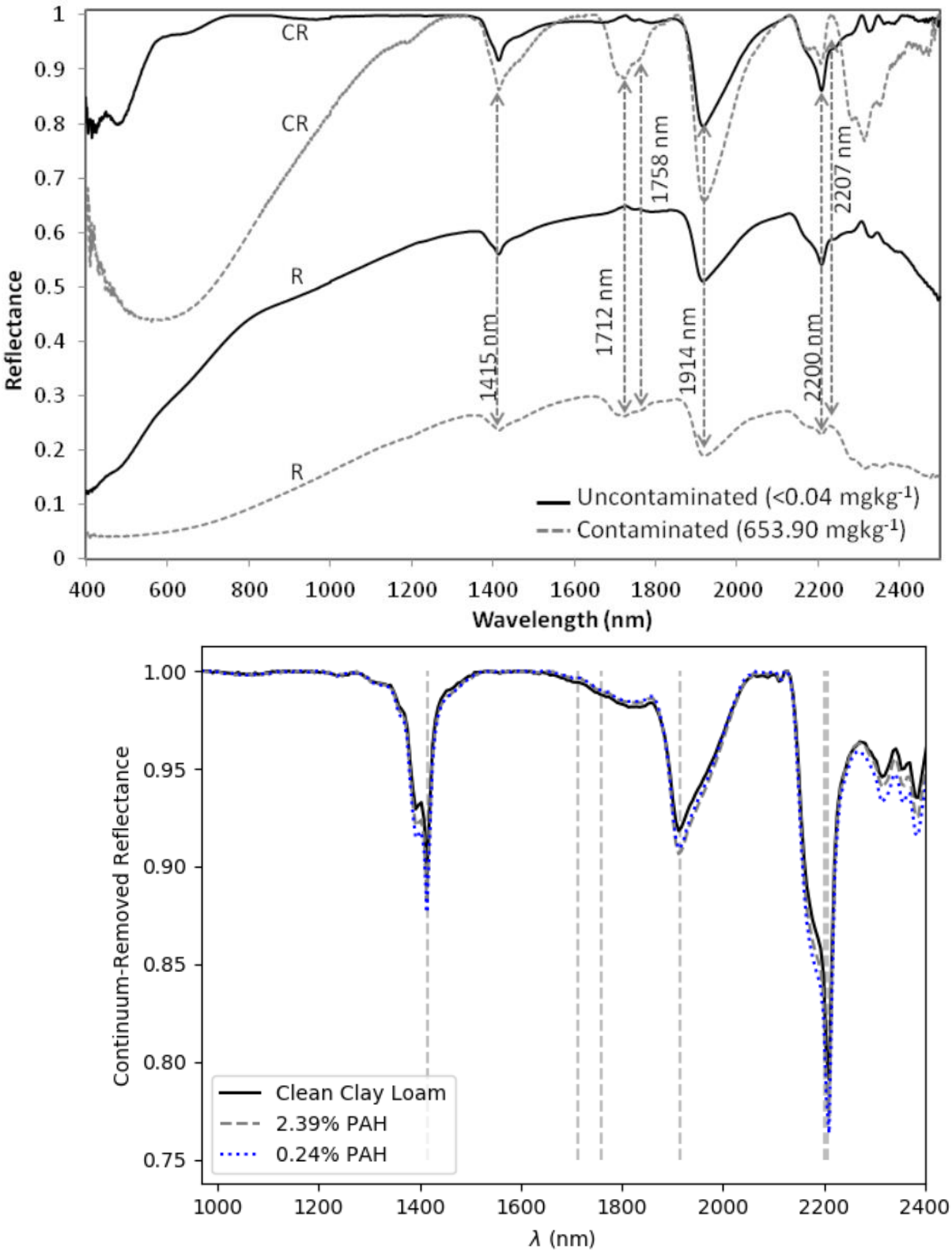

Figure 22: Comparison of reflectance spectra of uncontaminated and PAH-contaminated soils as measured by Douglas et al. and as measured in this project. Top: Average of raw (R) and continuum-removed (CR) spectra of PAH-contaminated (dashed grey line) and uncontaminated (solid black line) samples, reproduced from [19]. Bottom: Continuum-removed spectra of PAH-contaminated (TX141 and TX141+; dashed grey and blue lines) and uncontaminated (CLNSOIL2; solid black line) clay loam samples as imaged by Tellux and analysed by MTC. Vertical dashed lines indicate wavelengths of characteristic features identified by Douglas et al.

4.3 Concentration

The difference between spectra of uncontaminated and contaminated soils in [19] (see above) is far greater than between CLNSOIL2, TX141, and TX141+, despite TX141+ being approximately four times as contaminated as the Niger Delta soil (2.39% vs. 0.65%). While spectral features are not necessarily linearly proportional to concentration, they are generally greater with greater concentration.

Figure 23 shows the difference in reflectance spectra between TX141/+ and CLNSOIL2. With the exception of absorption around 1900 nm, there is *less* difference in spectral features in TX141+ despites having ten times the PAH concentration of TX141.

These differences in reflectance spectra are of a similar order to the standard deviation. For example, the greatest deviation is 2204 nm is −0.028, where the corresponding standard deviation is 0.029 (TX141) and 0.022 (TX141+).

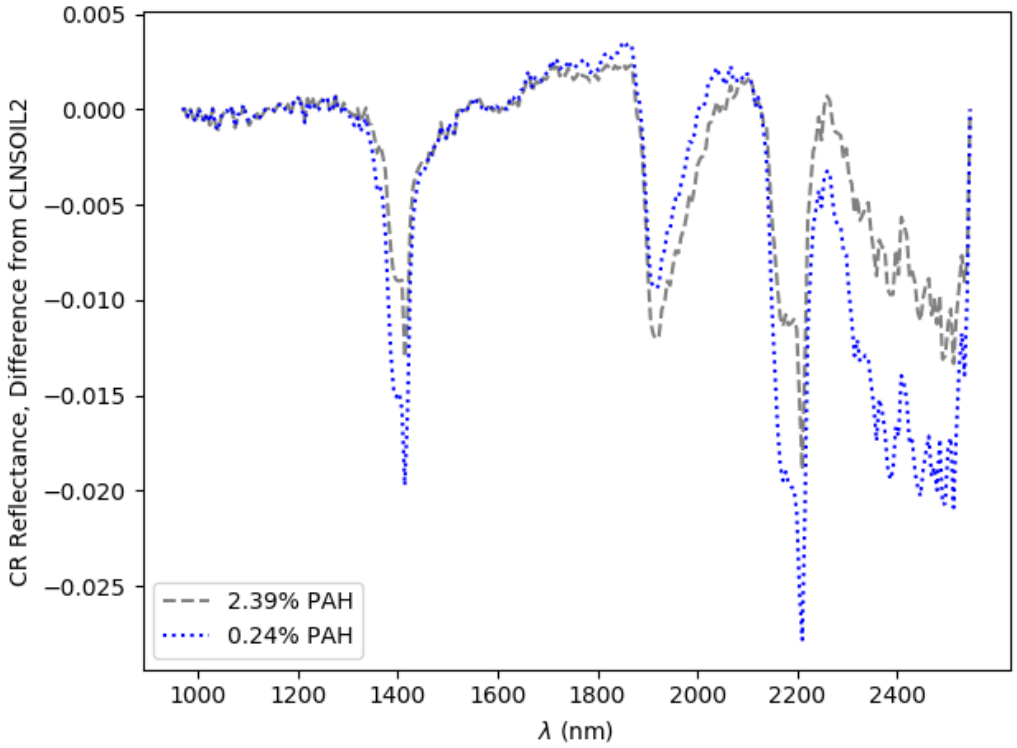

Figure 23: Difference between TX141+ (grey dashed line) and TX141 (blue dotted line) and CLNSOIL2 spectra, as measured with Specim SWIR.

While heat from illumination for hyperspectral imaging can affect some materials, polycyclic aromatic hydrocarbons are generally not volatile (i.e. unlikely to evaporate) and solid at standard temperature and pressure [22], so it is unlikely that imaging conditions have adversely affected the samples. A potential explanation is differing particle sizes, which can affect spectra due to differences in relative absorption and scattering [23]. It is possible to determine particle size with similar processing to chemical identification, though simultaneous determination of chemical, concentration, and particle size may not be practically possible.

5 Conclusions and Recommendations

Results of this project's hyperspectral imaging of uncontaminated and contaminated soil samples with Vis-VNIR, NIR, and SWIR cameras are generally consistent with literature. However, not all potentially-observable features, with respect to identification of specific contaminants, have been reproduced to the extent that would be desired for robust classification of hazardous spoil.

Fully evaluating (and/or developing) hyperspectral classification is expected to require substantially more samples than examined here, particularly with the diversity of contaminant types that may be encountered in works. Many articles and applications of hyperspectral imaging consider a comparatively limited range of contaminants, and their results may not be generally reproducible with a greater range and/or greater required specificity.

Pragmatically, it may be preferable to develop and deploy hyperspectral imaging for screening of spoil – predicting presence of hazardous materials, but not identifying specific chemicals or their concentrations. While not a complete solution for classifying spoil for appropriate disposal, it would reduce the volume of spoil needing analysis by more precise, but more expensive and time-intensive, means.

5.1 Camera Suitability and Spectral Range

While differences in uncontaminated and contaminated spectra imaged using the FX17 camera are consistent with those found in literature, the range of characteristic features may be insufficient to identify contaminants with sufficient specificity for categorising spoil. In particular, the features correlated to the presence of total petroleum hydrocarbons and that of polycyclic aromatic hydrocarbons are highly similar to each other. Very-near-infrared (VNIR) as imaged by the FX10 is unlikely to be suitable.

Specific identification of contaminants, therefore, is likely to require a camera capable of imaging SWIR wavelengths. Cameras operating in longer wavelengths may be worth considering. The increased cost of such systems must be considered in the context of ENWL's use case. If multiple cameras are required to inspect spoil at many worksites each day, SWIR cameras may be prohibitively expensive. While full characterisation of contaminants may not be possible with NIR cameras such as the FX17, these could be deployed to perform preliminary analysis of spoil, discriminating between uncontaminated and contaminated spoil, reducing the volume of spoil requiring characterisation by other (more expensive or time-consuming) means.

A significant factor in camera suitability of ENWL's use case is cost. While cost also depends on factors such as resolution and framerate (e.g. "snapshot" cameras are more expensive than equivalent line-scan cameras), cost increases significantly with longer wavelengths due to differences in detector construction. Line-scan VNIR, NIR, and SWIR cameras may cost on the order of £20,000, £45,000, and £150,000, respectively. While SWIR cameras are more desirable with respect to specificity of detection, this may be prohibitively expensive. While many hyperspectral cameras are designed for field as well as industrial/laboratory use, their performance does depend on careful operation and being kept in good condition. As such, they may be more suited to use by dedicated personnel travelling between work sites. The practicality of such deployment will depend on number of work sites to be assessed each day, their geographic distribution, etc.

5.2 Further Work

The extent to which these trial results can be extrapolated to identification of contaminants in the field is limited by a number of factors:

- The reference materials contain only soil matrix and contaminants of a single type, and are not representative of genuine spoil which may contain other (benign) materials and/or multiple contaminants.
- Notwithstanding differences in composition and concentration between Sigma-Aldrich's and Tellux's reference materials, contaminants in each sample were limited in the range of concentrations.
- Trials were conducted in laboratory conditions, with artificial lighting. Artificial lighting ensures consistency in the spectrum used to illuminate samples, reducing unintended spectral artefacts and as well as more even lighting, reducing shadows in the images. Ideally, measurements would be carried out on-site and without needing bulky lighting equipment.

Further trials should be conducted before developing and deploying hyperspectral imaging for soil inspection. These should investigate genuine spoil samples and samples with varying concentrations of a known contaminant. Due to the inherent lack of control in contamination of genuine spoil, this may require two series of trials – one with genuine spoil and one with doped samples of various concentration – conducted in parallel or series. These trials may be supplemented by drawing from literature or other existing work.

Should ENWL wish to continue to develop a soil characterisation solution we would recommend that MTC would carry out the trial planning, detailing the number and composition of the samples to be tested, as well as provide an impartial analysis of the outcome of the trials. We would also suggest further engaging with Tellux to perform further trials, due to their expertise in hyperspectral technologies.

5.3 Next Steps

Immediate next steps should be consideration of which method of deployment should be pursued:

- Low-specificity assessment of spoil contamination (e.g. with NIR),
- High-specificity classification of individual contaminants (e.g. with SWIR),

as this will determine the kind and extent of trials required in future work.

Alternatively, both could be pursued in tandem. This would provide some measure of contingency should one method prove unviable. This could also enable a third option for deployment, where (e.g.) initial screening is conducted with less precise, more robust cameras deployed to all teams and specific classification of pre-screened hazardous spoil conducted with more precise cameras by dedicated teams.

This work should consider ENWL's current practices (e.g. number, size, duration, and geographic distribution of works each day) as well as costs (monetary, time, labour) associated with alternatives to hyperspectral classification of spoil.

6 References

[1] Environment Agency, “Excavated Waste from Utilities Installation and Repair: RPS 211,” Environment Agency, 7 July 2022. [Online]. Available: https://www.gov.uk/government/publications/excavated-waste-from-utilities-installation-and-repair-rps-211/excavated-waste-from-utilities-installation-and-repair-rps-211. [Accessed July 2022].

[2] Natural Resources Wales, Scottish Environment Protection Agency, Environment Agency, “Guidance on the Classification and Assessment of Waste,” 1st Edition v1.2.GB, Technical Guidance WM3, 2021.

[3] L. Moore and B. Nye, “Hyperspectral Imaging for Soils - Problem Definition,” MTC, 37716-REP001, 2022.

[4] L. Moore and B. Nye, “Hyperspectral Imaging for Soils - First-Stage Downselection,” the MTC, 37716-REP002, 2022.

[5] F. Corentin, “Acquisition d'Image,” Tellux, Petit-Couronne, 2023.

[6] Specim, “FX17 Datasheet,” Specim, 2019.

[7] Headwell Photonics Inc., “Micro-Hyperspec Hyperspectral Imaging Systems - Product Datasheet,” Analytik Ltd., 2019.

[8] Resonon, “Pika IR+ Technical Datasheet,” Resonon Inc., 2022.

[9] Resonon, “Pika IR-L+ Datasheet,” Resonon Inc., 2022.

[10] Specim, “FX10 Datasheet,” Specim, 2020.

[11] Specim, “SWIR Datasheet,” Specim, 2020.

[12] CEMEX Ventures, “Tellux,” 2023. [Online]. Available: https://www.cemexventures.com/tellux/. [Accessed March 2023].

[13] Tellux, “Analyse des sols pollués - TELLUX,” 2023. [Online]. Available: https://tellux.fr/. [Accessed March 2023].

[14] P. Mishra, A. Chauhan and T. Pettersson, “Seeing Through Plastics: a Novel Combination of NIR Hyperspectral Imaging and Spectral Orthoganalisation for Detecting Fresh Fruit Inside Plastic Packaging to Support Automated Barcodeless Checkouts in Supermarkets,” *Food Control,* 2023.

[15] X. Wu, J. Li, L. Yao and Z. Xu, “Auto-Sorting Commonly-Recovered Plastics from Waste Houshold Appliances and Electronics using Near-Infrared Spectroscopy,” *Journal of Cleaner Production,* vol. 246, p. 118732, 2022.

[16] N. Kayet, K. Pathak, V. Chowdray, C. Singh, B. K. Bhattacharya, S. Jumar, S. Kumar and I. Shaik, “Vegetation Health Conditions Assessment and Mapping using AVIRIS-NG Hyperspectral and

Field Spectroscopy Data for -Environmental Impact Assessment in Coal Mining Sites," *Ecotoxicology and Environmental Safety,* vol. 239, p. 113650, 2022.

[17] Cranfield Soil and Agrifood Institute, "Soilscapes Map," [Online]. Available: http://www.landis.org.uk/soilscapes/. [Accessed May 2022].

[18] United States Department of Agriculture, "Soil Taxonomy - a Basic System of Soil Classification for Making and Interpreting Soil Surveys," Natural Resources Conservation Service, 1999.

[19] R. K. Douglas, S. Nawar, M. C. Alamar, A. M. Mouazen and F. Coulon, "Rapid Prediction of Total Petroleum Hydrocarbon Concentration in Contaminated Soil using vis-NIR Spectroscopy and Regression Techniques," *Science of the Total Environment,* Vols. 616-617, pp. 147-155, 2017.

[20] Y. Jeong, J. Yu, L. Wang and K.-J. Lee, "Bulk Scanning Method of a Heavy Metal Concentration in Tailings of a Gold Mine using SWIR Hyperspectral Imaging System," *International Journal of Applied Earth Observations and Geoinformation,* vol. 102, no. 102382, 2021.

[21] M. Izawa, D. Applin, L. Norman and E. Cloutis, "Reflectance Spectroscopy (350-2500 nm) of Solid-State Polycyclic Aromatic Hydrocarbons," *Icarus,* vol. 237, pp. 159-181, 2014.

[22] I. H. Abdel-Shafy and M. S. Mansour, "A Review on Polycyclic Aromatic Hydrocarbons: Source, Environmental Impact, Effect on Human Health and Remediation," *Egyptian Journal of Petroleum,* vol. 25, no. 1, pp. 107-123, 2016.

[23] L. Pieszczek and M. Daszykowski, "Near-Infrared Hyperspectral Imaging for Polymer Particle Size Estimation," *Measurement,* vol. 186, no. 110201, 2021.

[24] Sigma-Aldrich, "TPH - Sandy Loam 3 Safety Data Sheet," Merck, CRM353, 2022.

[25] Sigma-Aldrich, "TPH - Clay Loam 1," Merck, CRM359, 2022.

[26] Sigma-Aldrich, "PAHs - Loamy Clay 1 Safety Data Sheet," Merck, CRM141, 2022.

[27] Sigma-Aldrich, "PAHs - Clay Soil Safety Data Sheet," Merck, CRM170, 2022.

[28] Sigma-Aldrich, "Trace Metals - Sewage Amended Soil Safety Data Sheet," Merck, CRM005, 2022.

[29] Sigma-Aldrich, "Metals in Soil Safety Data Sheet," Merck, SQC001, 2022.

[30] Sigma-Aldrich, "Lead - Soil Safety Data Sheet," Merck, PB3000, 2022.

[31] Sigma-Aldrich, "Trace Metals - Sandy Loam 6 Safety Data Sheet," Merck, CRM043, 2022.

Appendix A – Reference Materials – Sigma Aldrich

A.1 – Total Petroleum Hydrocarbons (TPH)

The reference materials selected to represent TPH-contaminated spoil are CRM353 – TPH – Sandy Loam 3 and CRM359 – TPH – Clay Loam 1 [24, 25]. The contaminants in each are nominally identical, and listed in Table 4.

Table 4: Contaminants in reference materials CRM353 and CRM359.

Contaminant	Concentration
Baseoil	0.1 – 1%
Fuel oil no. 2	0.1 – 1%

A.2 – Coal Tar

The reference materials selected to represent coal-tar-contaminated spoil are CRM141 – PAHs – Loamy Clay 1 and CRM170 – PAHs – Clay Soil [26, 27]. Rather than coal tar itself, these are contaminated with polycyclic aromatic hydrocarbons (PAHs) found in coal tar, and which may be used as "marker" chemicals for identifying the presence of coal tar. Of particular importance is benzo[a]pyrene, which has its own concentration limit. The contaminants in both CRM141 and CRM170 are nominally identical and listed in Table 5.

Table 5: Contaminants in reference materials CRM141 and CRM170.

Contaminant	Concentration
2,2,2,o,p'-Pentachloroethylidenebisbenzene	0.025 – 0.1%
2,2,o,p'-Tetrachlorobinylidenebisbenzene	0.025 – 0.1%
Heptachlor	0.025 – 0.1%
Heptachlor endo-epoxide isomer	0.025 – 0.1%
Endrin	0.025 – 0.1%
Hexachlorobenzene	0.025 – 0.1%
Benz[a]anthracene	0.025 – 0.1%
Acenaphthene	0.025 – 0.1%
Benzo[ghi]perylene	0.025 – 0.1%
Anthracene	0.025 – 0.1%
Benzo[jk]fluorene	0.025 – 0.1%
Benzo[k]fluoranthene	0.025 – 0.1%
Chrysene	0.025 – 0.1%
Benz[e]acephenanthrylene	0.025 – 0.1%
2,2-bis(4-Chlorophenyl)-1,1-dichloro-ethane	0.025 – 0.1%
Dieldrin	0.025 – 0.1%
Aldrin	0.025 – 0.1%
1,1,1-Trichloro-2,2-bis(4-chlorophenyl)ethane	0.025 – 0.1%
2,2-bis(p-Chlorophenyl)-1,1-dichloroethylene	0.025 – 0.1%
2,2',5,5'-Tetrachlorobiphenyl	0.025 – 0.1%
2,4,4'-Trichlorobiphenyl	0.025 - 0.1%
2,2',4,5,5'-Pentachlorobiphenyl	0.025 – 0.1%
2,2',3,4,4',5,5'-Heptachlorobiphenyl	0.025 – 0.1%
2,2',3,4,4',5'-Hexachlorobiphenyl	0.025 – 0.1%
2,2',4,4',5,5'-Hexachlorobiphenyl,	0.025 – 0.1%
Dibenz[a,h]anthracene,	0.025 – 0.1%
Benzo[a]pyrene,	0.025 – 0.1%

A.3 – Heavy Metals

The reference materials examined to represent heavy-metal-contaminated soil are CRM005 – Trace Metals – Sewage Amended Soil, SQC001 – Metals in Soil, and PB3000 – Lead – Soil [28, 29, 30]. The contaminant(s) in each are listed in Table 6, Table 7, Table 8 respectively.

Table 6: Contaminant in reference material CRM005.

Contaminant	Concentration
Nickel dinitrate hexahydrate	0.0025 – 0.1%

Table 7: Contaminants in reference material SQC001.

Contaminant	Concentration
Calcium chloride	1 – 10%
Vitreous silica	1 – 10%
Barium nitrate	0.1 – 1%
Nickel dinitrate hexahydrate	0.025 – 0.1%
Cobalt dichloride hexahydrate	0.025 – 0.1%
Lead (II) nitrate	0.025 – 0.1%
Cadmium chloride	0.025 – 0.1%
Arsenic	0.025 – 0.1%
Selenium dioxide	0.025 – 0.1%
Silver (II) nitrate	0.0025 – 0.25%

Table 8: Contaminant in reference material PB3000.

Contaminant	Concentration
Lead (II) nitrate	0.5 – 1%

Appendix B – Reference Materials – Tellux

Four samples were prepared by Tellux to substitute for unavailable reference materials for TPH- and PAH-contaminated soils: CRM359 (total petroleum hydrocarbons) and CRM141 and CRM170 (coal tar). These materials consist of soil matrixes from CLNSOIL2 and CLNSOIL5 and contaminants as listed in Table 9. Diethylether was used as an intermediary in the doping process [5].

Table 9: Concentration of doped contaminants in samples prepared by Tellux. Values from [5].

Analyte	Concentration			
	TX359	TX141	TX141+	TX170
CLNSOIL2	99.87%	99.76%	97.61%	
CLNSOIL5				99.76%
Decane	0.02%			
Dodecane	0.02%			
Pentodecane	0.02%			
Docosane	0.02%			
Teracontane	0.00%			
Pyrene	0.05%	0.06%	0.54%	0.05%
Acenaphthene		0.05%	0.44%	0.04%
Chrysene		0.02%	0.22%	0.02%
Fluorene		0.04%	0.35%	0.04%
Naphthalene		0.06%	0.57%	0.06%
Phenanthrene		0.03%	0.27%	0.03%
Total Petroleum	0.13%			
Total PAH		0.24%	2.39%	0.24%

Appendix C – Material Spectra

The figures on the following pages display the continuum-removed, normalised, mean spectra of each reference material, as imaged by Specim FX17 and SWIR cameras. The scale in x (wavelength) and y (reflectance) are uniform throughout this appendix.

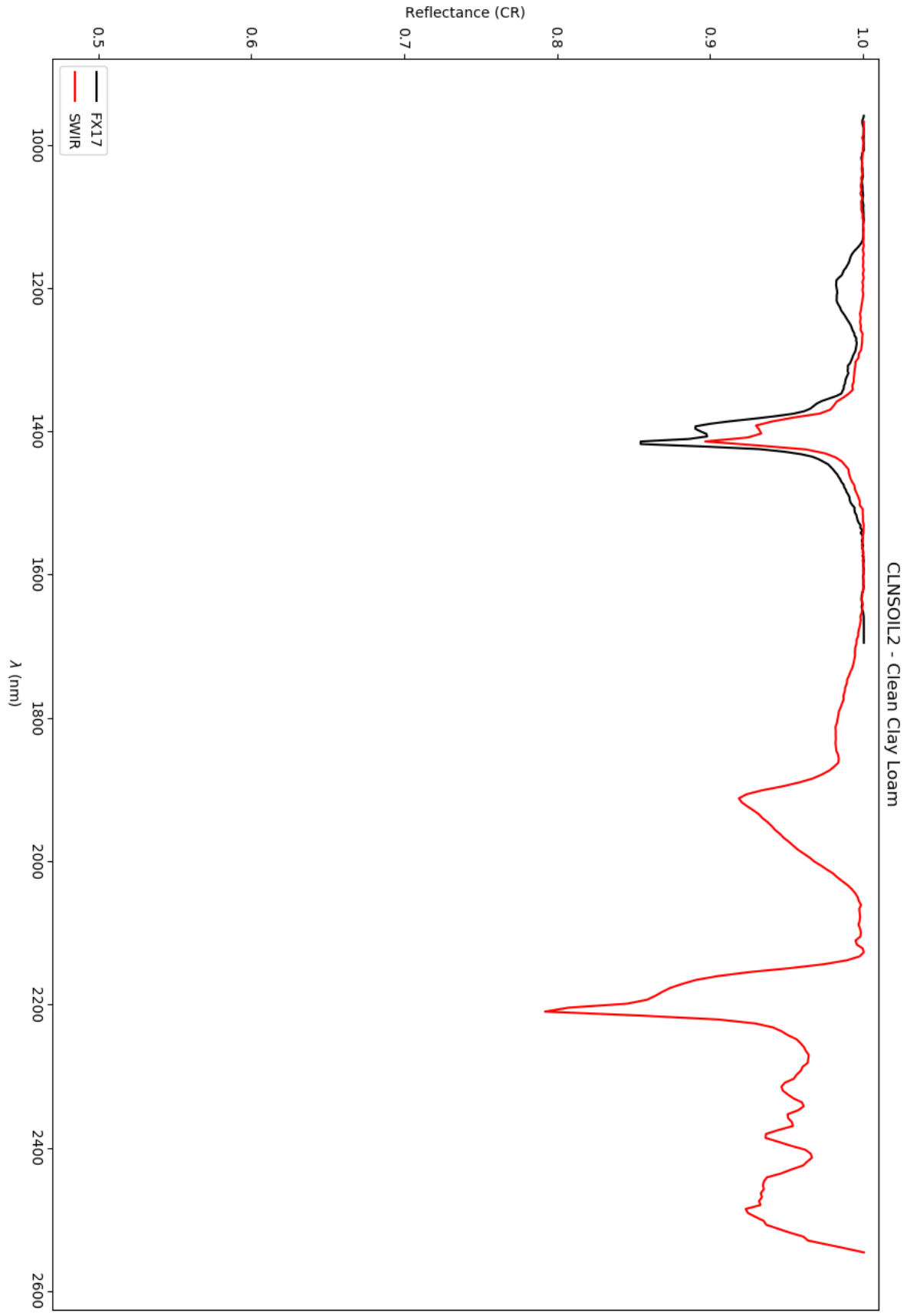

Figure 24: Continuum-removed reflectance spectra for clean clay loam reference material.

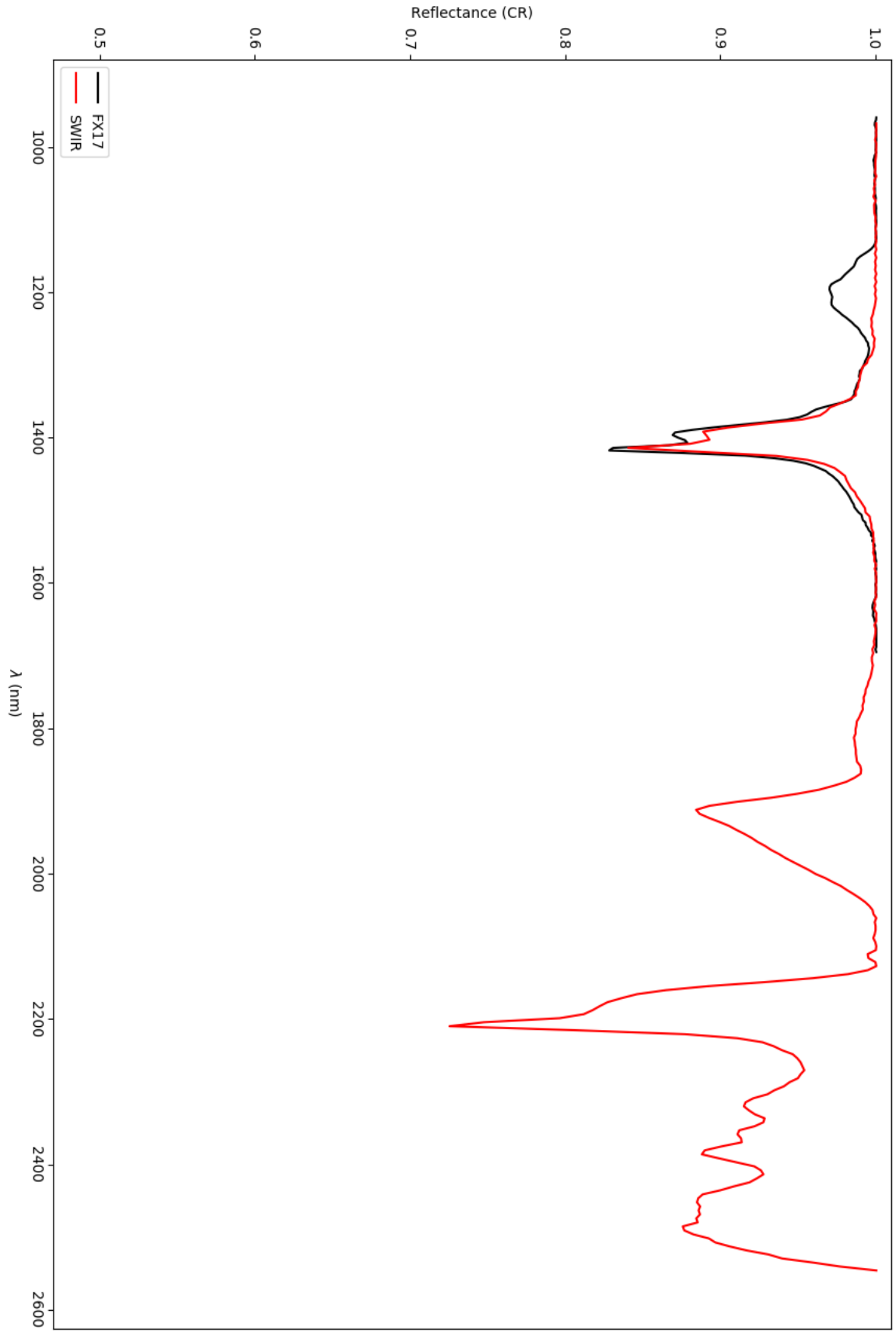

Figure 25: Continuum-removed reflectance spectra for clean clay reference material.

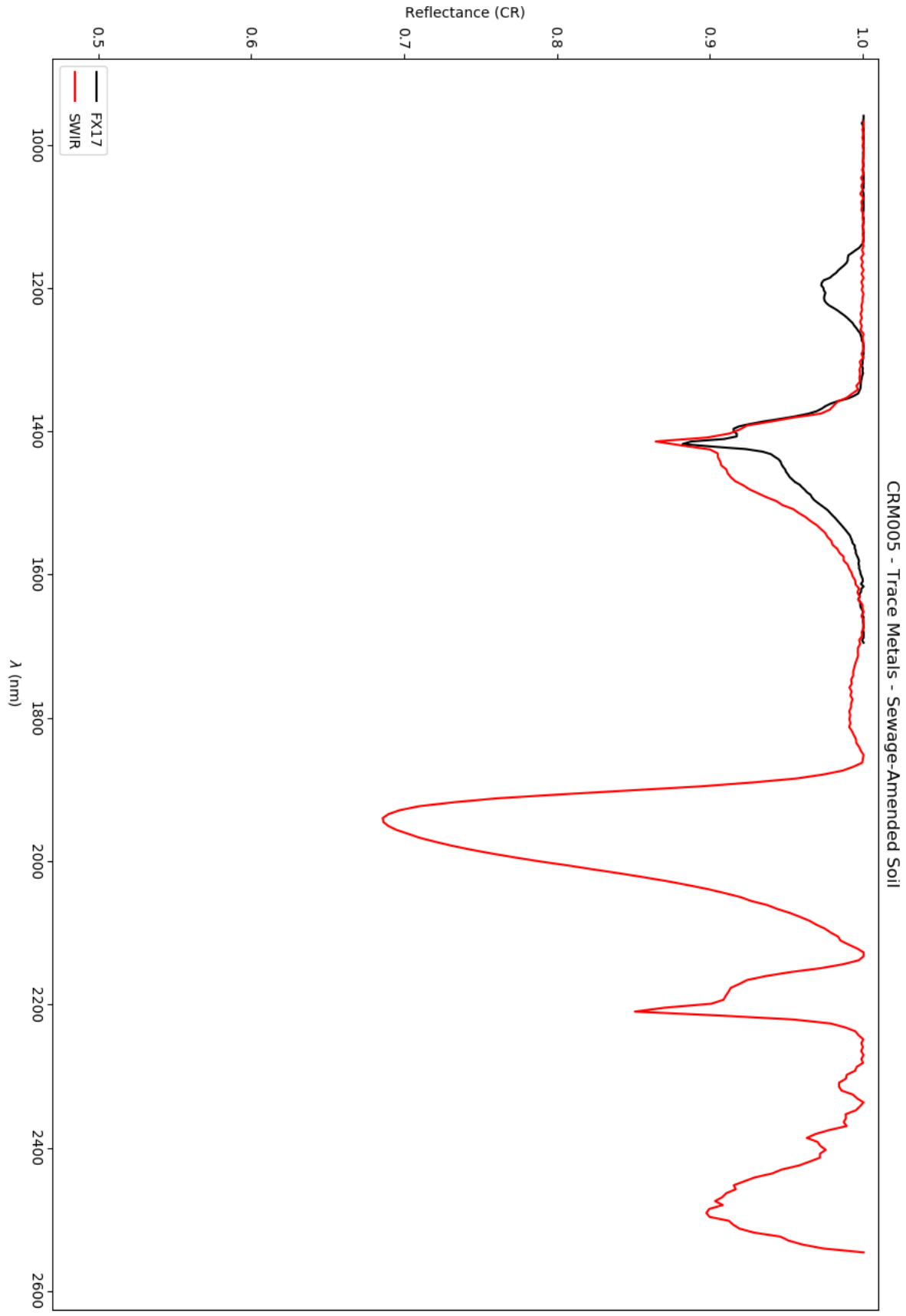

Figure 26: Continuum-removed reflectance spectra for trace metals in sewage-amended soil reference material.

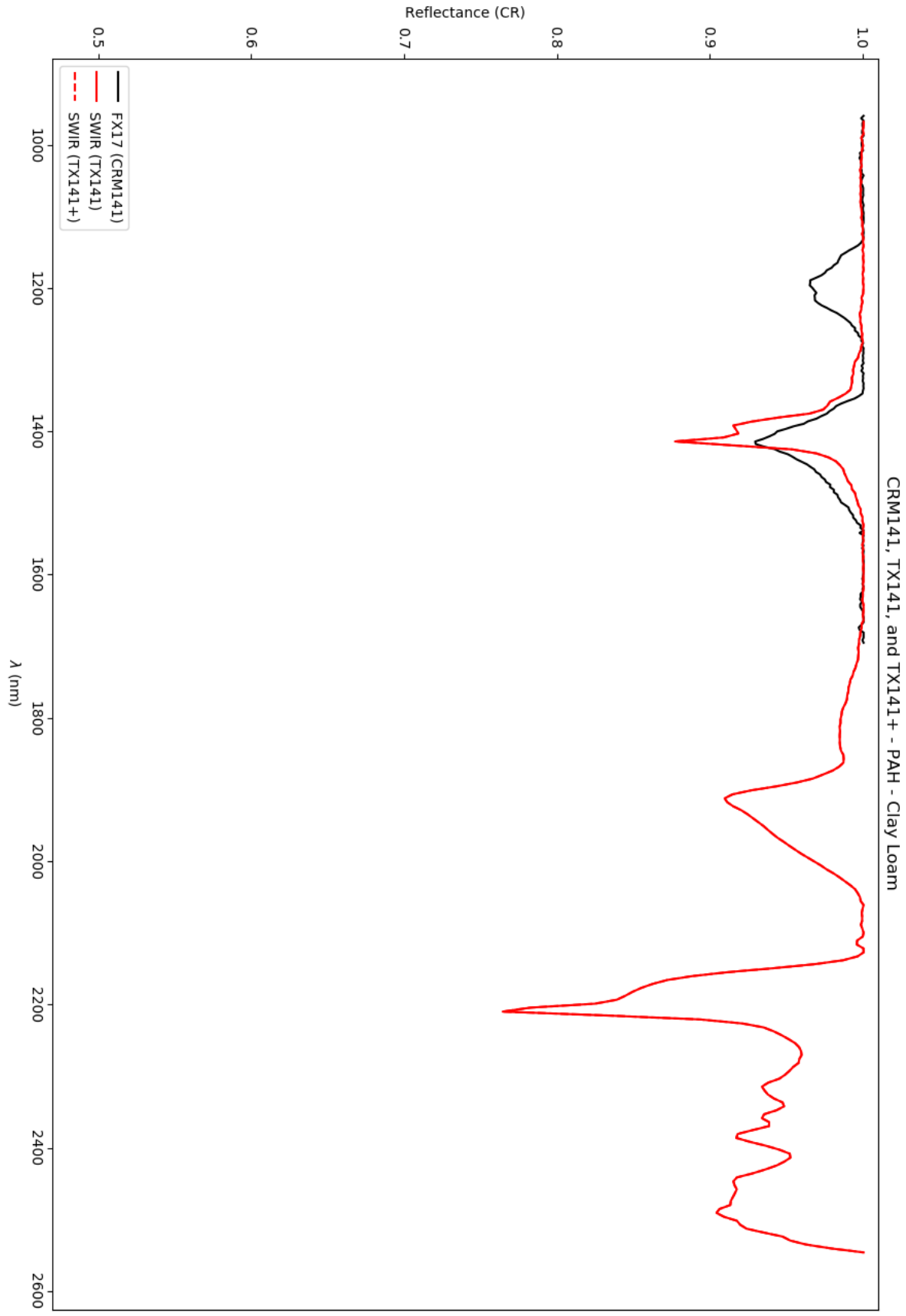

Figure 27: Continuum-removed reflectance spectra for PAH in clay loam reference material.

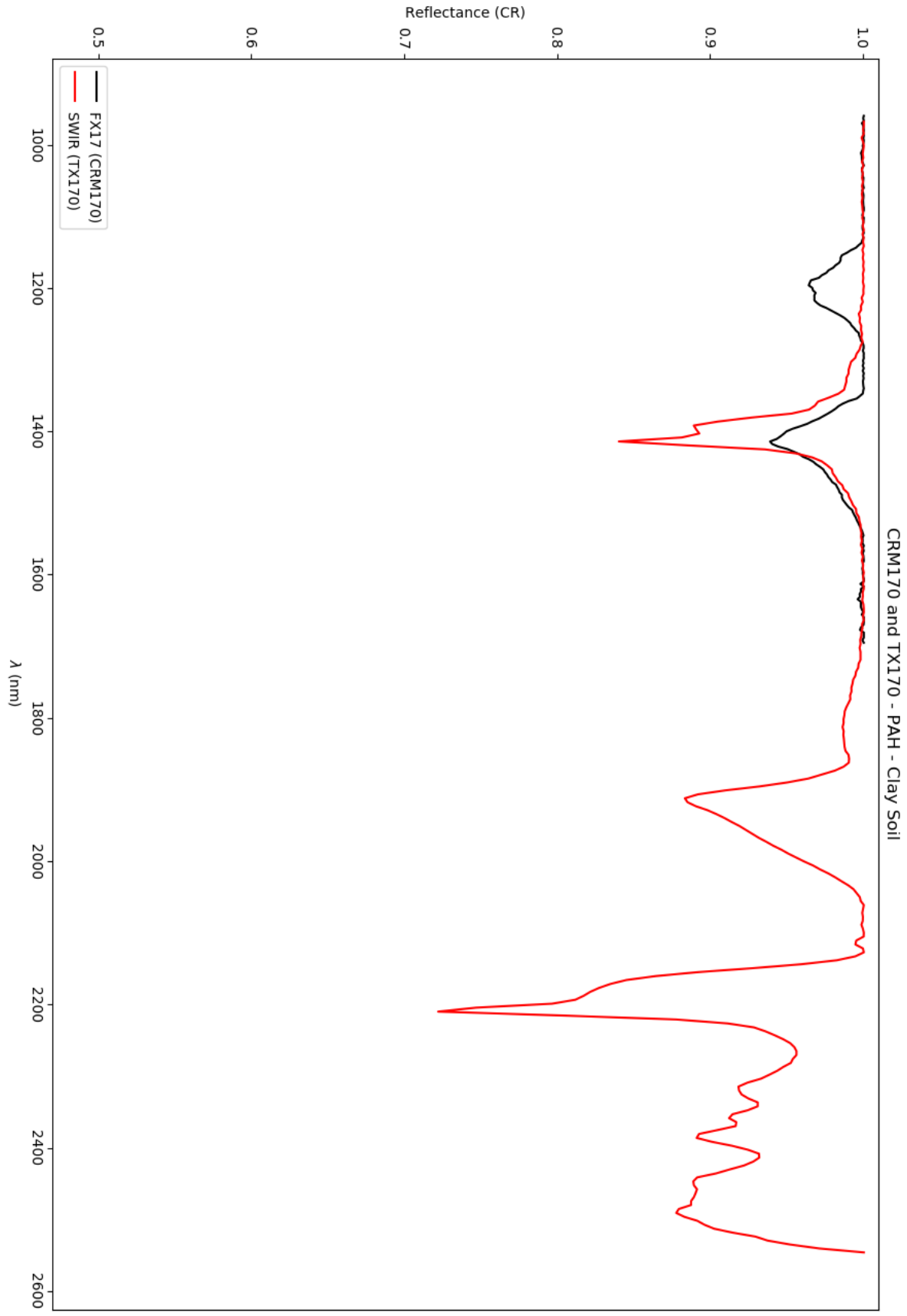

Figure 28: Continuum-removed reflectance spectra for PAH in clay reference material.

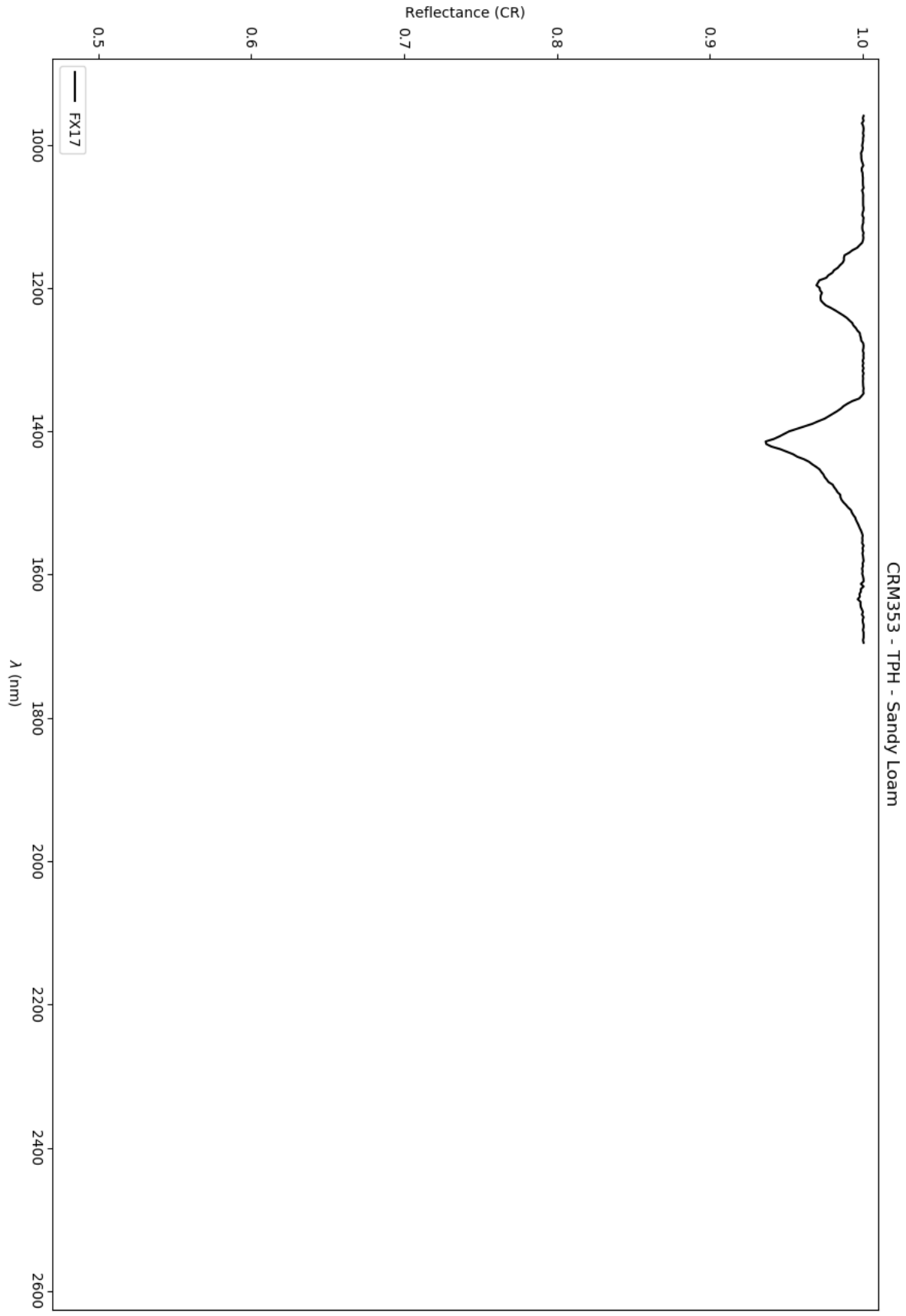

Figure 29: Continuum-removed reflectance spectrum for TPH in sandy loam reference material.

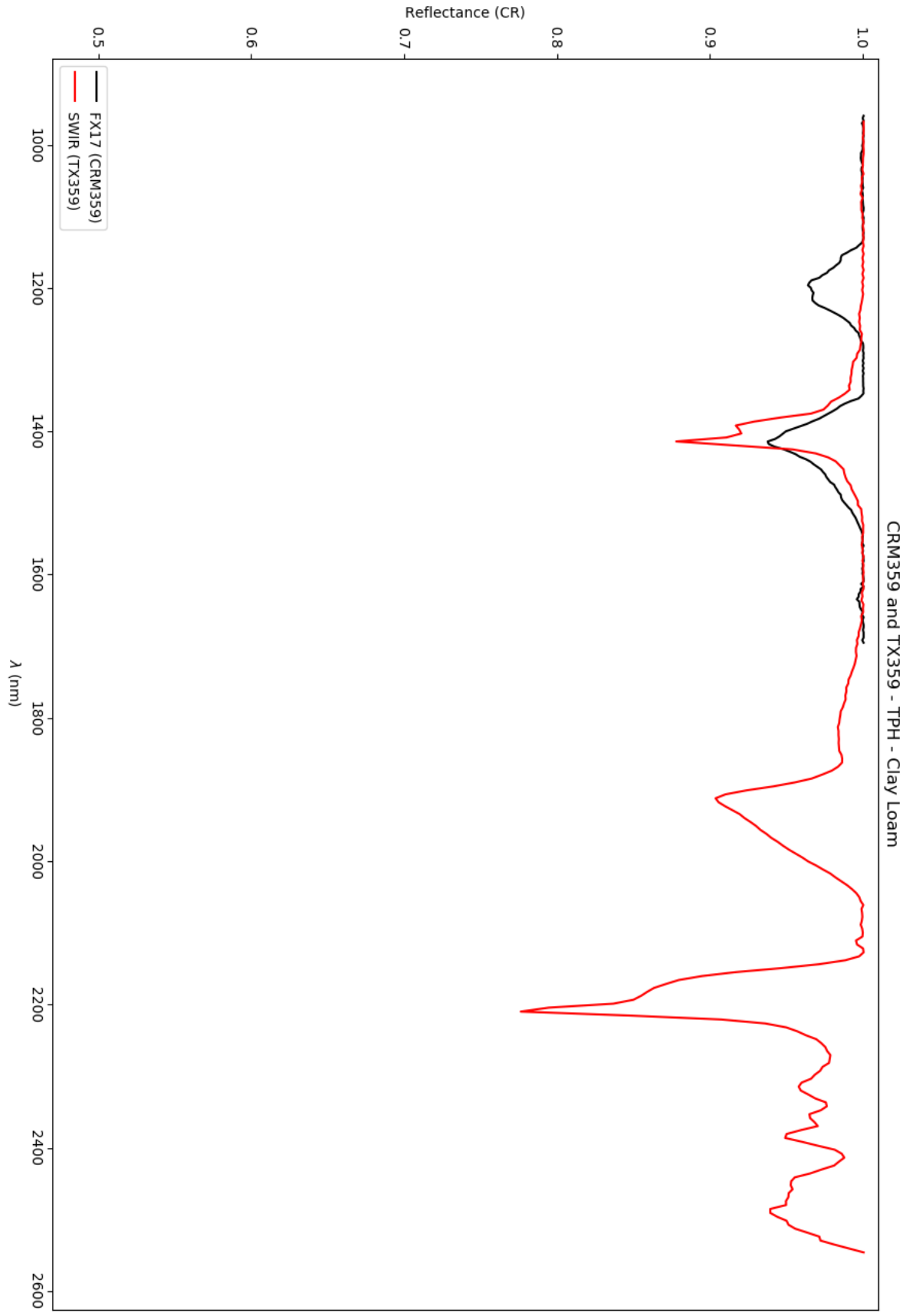

Figure 30: Continuum-removed reflectance spectra for TPH in clay loam reference material.

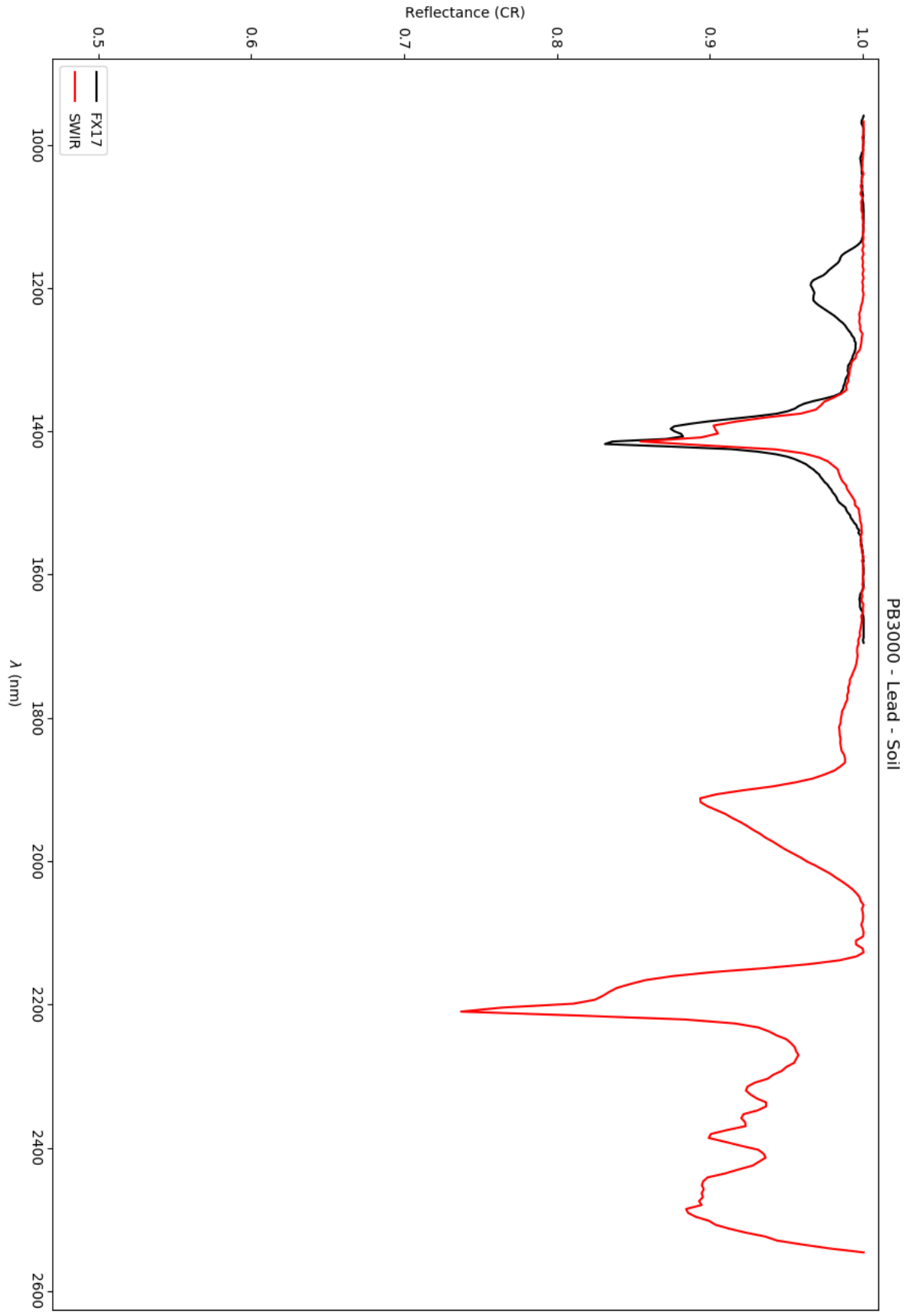

Figure 31: Continuum-removed reflectance spectra for lead (II) nitrate in soil reference material.

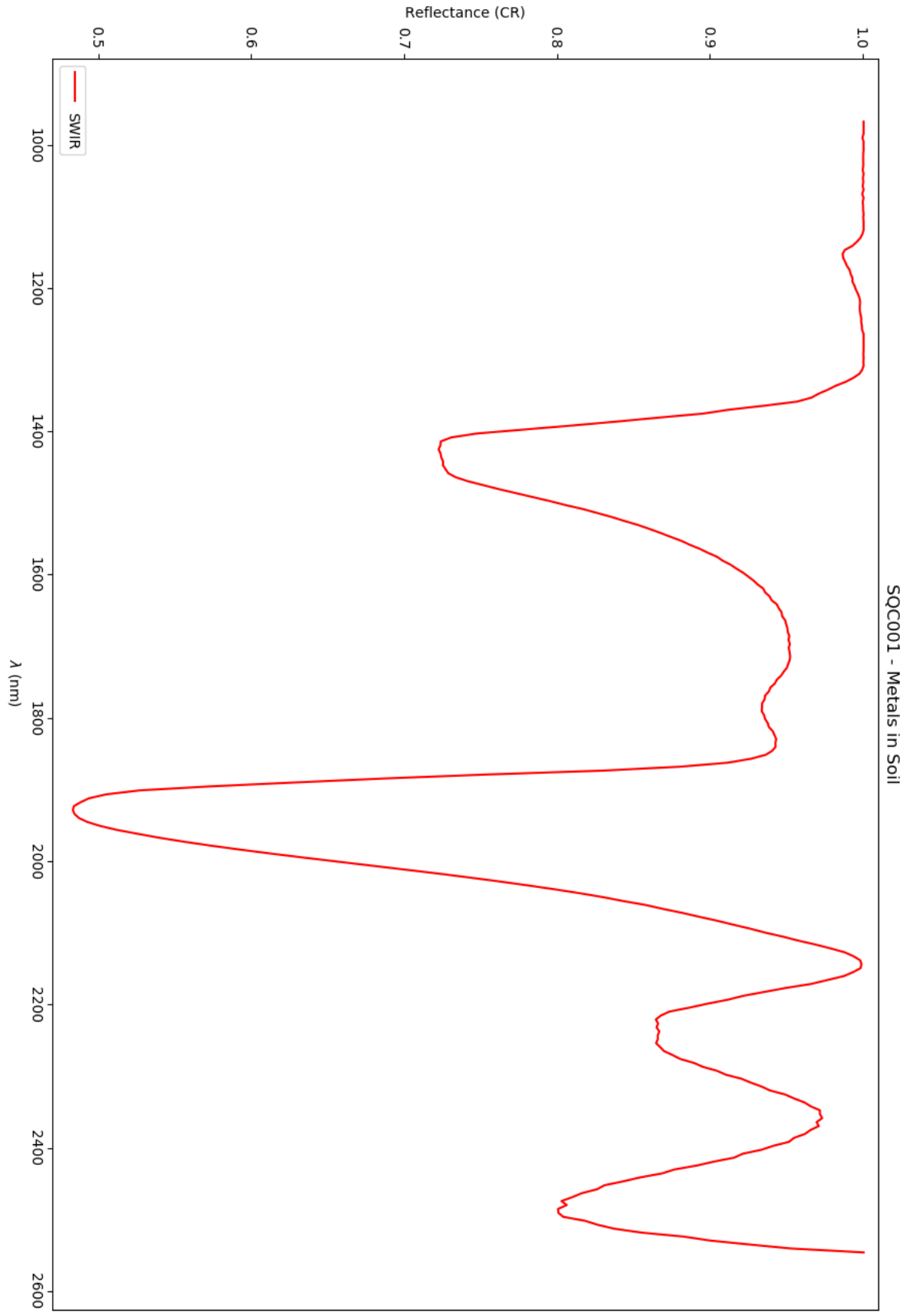

Figure 32: Continuum-removed reflectance spectrum for metals in soil reference material.