What a Multispectral Measurement Determines: Response Summaries and Consistency Repair

Liam Chalcroft, Eduardo Camarillo Abad, Peter Christopher, Thomas Albrow-Owen, Oliver Burton
Prospectral, Cambridge, UK
{liam, eduardo, peter, tom, oli}@prospectral.tech

***Abstract*—Each multispectral channel averages the incoming spectrum according to its spectral response function (SRF). Reconstruction methods often summarise responses by a centre wavelength or Gaussian and enforce exact measurement consistency after reconstruction. We ask how much response detail is needed and which errors consistency correction can remove. Holding the estimator, prior and regularisation selection fixed while varying only the response description, we find that on Sentinel-2 and Landsat responses a moment-based Gaussian gives almost the same spectral error as the full curves (RMS reflectance error $\times 100$ of 2.805 against 2.791 for a linear minimum mean-square error, LMMSE, estimator), whereas on a broader population of simulated sensors it adds 8.3%. A classical row-space decomposition separates the error the sensor measures from the error in its null space. Without noise, undamped correction removes the measured part: it restores colour, which a three-channel colour sensor determines, but not narrow-band vegetation and water indices, which Sentinel-2 and Landsat largely leave undetermined. Reference-based excess misfit ranks the relative change in colour error after correction (Spearman $\rho = -0.54$), but its size alone does not reliably identify which cases benefit.**

***Index Terms*—spectral response functions, hyperspectral reconstruction, measurement consistency, operator approximation, LMMSE**

I. Introduction

A multispectral measurement integrates a spectrum against a band response. Describing that channel by its nominal wavelength discards its finite width; describing it by a Gaussian also discards asymmetry and higher-order structure. Whether those omissions matter is a quantitative question about the spectrum, the sensor and the estimator: a visibly non-Gaussian band need not produce large reconstruction error, and a richer summary need not improve a fixed regularised estimator.

Sensor-conditioned models make the question practical. STARS encodes responses through averaged sinusoidal features [1], SenPa-MAE uses sampled response vectors [2], and Copernicus-FM uses centre and bandwidth [3]. Published scores cannot isolate the value of response detail, because the estimator, the prior and the response description change together. We fix the estimator family, the training prior and the regularisation selection, and vary only the response description, using per-pixel linear estimators standard in spectral imaging [4], [5].

We address two questions. **Q1:** how much response detail does reconstruction need? **Q2:** which part of the reconstruction error can a correction that enforces consistency with the measurement remove? Our contributions are (i) a controlled comparison of response descriptions from a centre wavelength to the full curve; (ii) a decomposition of each description's cost into the error the sensor measures and the error in its null space, which explains why the same summary is cheap for one sensor population and costly for another; (iii) an extension of the noiseless Matrix-R result [6] to noisy measurements and derived quantities, with a visibility measure, evaluated on real instruments, that states when consistency guarantees a quantity such as colour or a vegetation index; and (iv) a prespecified test of whether reference-based excess misfit ranks the effect of correction on colour.

Section II defines the measurement model and the estimators, Section III the data and sensors, and Section IV answers Q1. Section V develops the decomposition behind those results and uses it to answer Q2, and Section VI concludes. A companion paper [7] uses the same projection as a network's output constraint.

II. Measurement and Reconstruction

Each channel reports one weighted average of the spectrum. A change to the spectrum that leaves every one of these averages unchanged is invisible to the sensor: spectra that differ only by such a change are metamers and give identical measurements (Fig. 1). That part of a spectrum cannot be recovered from the measurement and must come from prior knowledge.

Formally, a spectrum $s \in \mathbb{R}^{204}$ is observed as $y = Rs + \epsilon$, where row c of the response operator R is the unit-mass SRF of channel c and ϵ is noise. The invisible directions form the null space, $\ker R = \{h : Rh = 0\}$. Its orthogonal complement $S = \mathrm{row}(R)$ is the subspace the sensor measures. The Moore–Penrose pseudoinverse R^+ [8] gives the orthogonal projector $\Pi_S = R^+R$ onto S; the metamer set is $\{s : Rs = y\}$ [9], [10]. Colour science calls $\Pi_S s$ the fundamental metamer and $(I - \Pi_S)s$ the metameric black of matrix-R theory [11]. Every reconstruction error $e = \hat{s} - s$ splits orthogonally into a measured component $\Pi_S e \in S$ and an unmeasured component $(I - \Pi_S)e \in \ker R$.

To isolate the value of response detail, the estimator sees only a description $\widetilde{R}$ of the responses, while the measurements, the consistency correction and the evaluation always use the true R. Spectra are modelled as $\widehat{s} = \mu + B\widehat{z}$, with mean μ and a principal component basis $B \in \mathbb{R}^{204\times 16}$ fitted only on training spectra, capturing 99.59% of the training variance

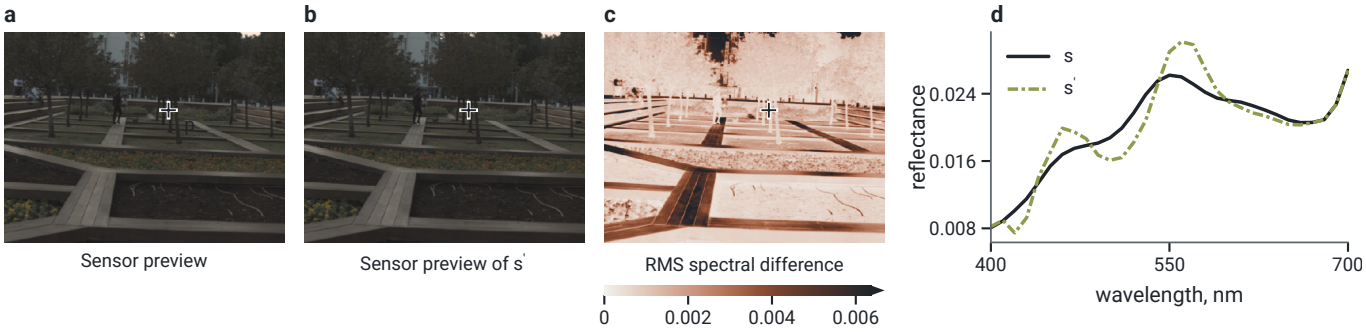


Fig. 1. Spectral differences a camera cannot see, illustrated on a cube from the public ARAD-1K hyperspectral dataset under the Nikon D3X response, separately from the BJTU-UVA evaluation. (a,b) Renders of the cube and of a perturbed copy s', identical to numerical precision. (c) RMS spectral difference between the two cubes. (d) Spectra at the marked pixel p. The perturbation, of nominal scale 0.1, lies in the null space of the camera's response operator.

(99.34% for 8 components, 99.75% for 24). Writing $A = \widetilde{R}B$ and $d = y - \widetilde{R}\mu$, ridge regression uses

$$\widehat{z} = (A^\top A + \alpha I)^{-1} A^\top d, \tag{1}$$

and the covariance-aware linear minimum mean-square error (LMMSE) estimator uses the training coefficient covariance Σ:

$$\widehat{z} = \Sigma A^\top (A\Sigma A^\top + \alpha I)^{-1} d. \tag{2}$$

Both are standard regularised estimators for multispectral reflectance estimation [4]. Each estimator, description and noise condition selects its own α from seven log-spaced values between 10^{-6} and 1, by scene-balanced RMSE on procedural validation sensors; α acts as a regulariser rather than a calibrated noise variance. Outputs are not clipped.

The response descriptions are the centre wavelength alone; a moment-based Gaussian, parameterised by the response's mean and variance; a Gaussian parameterised by its centre and full width at half maximum (FWHM); Fourier (harmonic) coefficients with $K = 4, 8, 16$; and the full response curve. Both Gaussians are sampled with trapezoidal weights and renormalised on the finite wavelength grid. A richer description need not improve a fixed regularised estimator, because approximation and regularisation interact; only for the best possible estimator can more information never hurt [12].

III. Experimental Setup

BJTU-UVA [13] provides 535 training, 114 validation and 116 test scenes on its 204-band grid, 397.32–1003.58 nm. The basis and covariance use all 34,238 training pixels; evaluation uses eight fixed pixels per test scene, 928 spectra in all, shared by every condition.

The satellite responses are the published Sentinel-2 MSI and Landsat 8/9 OLI/OLI-2 tables [14], [15]. We use a band only if its whole published response lies within the grid. This keeps 10 of 13 MSI and 6 of 9 OLI/OLI-2 bands, dropping only the cirrus and shortwave-infrared bands, which lie entirely beyond 1331 nm. The procedural population consists of simulated sensors with 4 to 24 channels whose bands are wider and overlap more than those of the validation sensors used to choose α. We evaluate noiseless measurements and additive Gaussian noise with standard deviation 0.005, with the same noise draws for every method. This is a controlled experiment rather than a satellite noise model, and the satellite responses are applied to the BJTU-UVA spectra; no satellite images are reconstructed.

For spectral RMS we average squared error over bands and pixels within each scene, then equally over scenes, sensors within each instrument family and families, before taking the square root. The bootstrap and spectral-RMSE correlations instead average per-pixel RMSE with the same balancing. Paired intervals resample test scenes with replacement and average the per-scene differences between two descriptions; the 2.5th and 97.5th percentiles over 9,999 resamples give a 95% interval. The vegetation (NDVI), water (NDWI [16]) and red-edge (NDRE) indices are normalised differences of the grid bands nearest 665/842, 560/842 and 705/750 nm. Colour uses the CIE 1931 2-degree observer under D65 over 400–700 nm and the CIEDE2000 difference [17].

A separate colour experiment uses a three-channel sensor whose responses are the CIE XYZ functions under D65; these generate the observations and are given to every estimator. Besides ridge and LMMSE it evaluates a learned reconstructor, our implementation of the STARS universal spectral reconstruction network [1] (a set transformer over channels with a DeepONet-style decoder), trained for squared spectral error with 5 seeds at noise levels 0 and 0.005. Figure 2c also reports 2,304 spectra from a central 48×48 crop of the public AVIRIS Cuprite reflectance flightline (NASA/JPL), interpolated onto the BJTU-UVA grid as a transfer population.

The repair criteria of Section V were fixed in advance and evaluated on a reserved set of procedural sensors (response banks).

IV. Results

Response detail (Q1). On the satellite responses a moment summary costs almost nothing. With full curves, noiseless LMMSE reaches RMS $\times 100$ of 2.791; with moments it reaches 2.805, a 0.5% increase, smaller than the 1.7% added by noise at $\sigma = 0.005$. Under that noise the 95% bootstrap interval for the moments-minus-full difference contains zero,

TABLE I

SPECTRAL RECONSTRUCTION ERROR AND ITS MEASURED AND NULL-SPACE COMPONENTS (RMS $\times 100$). MEAS. AND UNM. ARE $\Pi_S e$ AND $(I-\Pi_S)e$ FOR THE ERROR $e=\hat{s}-s$, WITH $\Pi_S = R^+R$. MEAN SQUARES ARE AVERAGED WITHIN SCENE, THEN OVER SENSORS WITHIN AN INSTRUMENT FAMILY, THEN OVER FAMILIES, BEFORE TAKING SQUARE ROOTS, SO TOT.^{2} = MEAS.^{2} + UNM.^{2} BEFORE ROUNDING. PROJ. IS THE ERROR AFTER UNDAMPED CONSISTENCY CORRECTION; IT EQUALS UNM. IDENTICALLY AT $\sigma = 0$ AND IS OMITTED ON THE PROCEDURAL POPULATION, WHERE GRAM CONDITIONS REACH 6.8×10^7 AND THE CORRECTION CAN REMOVE THE OBSERVATION RESIDUAL WHILE INCREASING SPECTRAL ERROR THROUGH AMPLIFIED MEASUREMENT NOISE.

		$\sigma = 0$			$\sigma = .005$			
Method	Summary	Tot.	Meas.	Unm.	Tot.	Meas.	Unm.	Proj.
Measured SRFs								
Ridge	Moments	6.875	0.298	6.869	6.913	0.418	6.900	6.920
	FWHM	6.602	0.262	6.597	6.643	0.404	6.631	6.651
	H8	7.161	0.474	7.145	7.195	0.560	7.173	7.191
	Full	6.798	0.112	6.797	6.936	0.414	6.924	6.943
LMMSE	Moments	2.805	0.189	2.799	2.837	0.312	2.820	2.867
	H8	3.056	0.257	3.045	2.903	0.384	2.877	2.923
	Full	2.791	0.108	2.789	2.838	0.311	2.821	2.867
Procedural parameter range								
Ridge	Moments	3.378	1.493	3.031	3.495	1.635	3.089	–
	H8	2.987	1.031	2.804	3.120	1.247	2.860	–
	Full	2.768	0.377	2.743	3.005	1.147	2.778	–
LMMSE	Moments	1.771	0.710	1.623	1.836	0.830	1.637	–
	H8	1.703	0.587	1.598	1.841	0.846	1.635	–
	Full	1.636	0.420	1.581	1.817	0.799	1.632	–

so the data cannot separate the two descriptions there. On the procedural population the same summary raises the error from 1.636 to 1.771, by 8.3%. Across 8 to 24 PCA components this noiseless gap stays at -0.2 to 0.5% on the satellite responses and 6.9–8.5% on the procedural population. Whether a summary suffices thus depends on the sensor population, not on the summary alone.

Table I splits each error into its measured and unmeasured components (Section V). On the satellite responses almost all of the error is unmeasured: the measured component is 0.15% of LMMSE's full-response mean square without noise and 1.20% with it, and the small extra error of the moment summary lies mostly (69%) in the null space, where the prior governs the reconstruction. On the procedural population the measured component is 20% of ridge's moment-summary mean square and 56% of the error that summary adds: with wide, overlapping bands, a coarse description misrepresents what the sensor measures. That error is in principle correctable, but these operators are ill-conditioned and the undamped correction amplifies measurement noise, reaching RMS $\times 100$ of up to 174.7 across the procedural cells against 1.6–5.7 uncorrected, so it is not applied there. Even on the satellite responses, undamped correction of noisy reconstructions barely changes or raises spectral error (Proj. column), which is why the correction is damped in practice (Section V).

A richer description is not uniformly better: on the satellite responses FWHM ridge has lower error than full-curve ridge in both conditions, and eight harmonics (H8) have higher error than moments for LMMSE. These orderings most likely reflect how each description interacts with the prior and the regulariser, and should not be read as differences in the physical information each description carries.

Downstream and colour results. Downstream indices track the spectral results: across the ridge and LMMSE cells for every description, 95th-percentile NDVI, NDWI and NDRE errors have Pearson correlations with spectral RMSE of 0.84–0.89 on the satellite responses and 0.84–0.86 on the procedural population.

In the colour experiment the learned reconstructor reaches a mean noiseless CIEDE2000 of 8.65, far above both linear estimators. The difference is measurement consistency. Without noise, ridge and LMMSE reproduce the observed XYZ channels to within 4.9×10^{-4} relative error, so their spectral error is almost invisible in colour, whereas the learned reconstructor leaves 11–36% of the observation unexplained. Under noise the linear estimators' misfit relative to the clean channels rises to about 0.026 and their colour error approaches that of the learned reconstructor, which is barely affected by noise. Consistency is best imposed after training: a consistency loss, training through the projection and a five-times-longer budget each left the network output inconsistent, whereas the correction of Section V makes it exact. Without noise, consistent linear reconstruction therefore preserves colour even though the full spectrum remains underdetermined.

V. WHAT THE MEASUREMENT DETERMINES

For any reconstruction $\hat{s}$, the error $e = \hat{s} - s$ has a measured component and an invisible one. With the observable residual $r = y - R\hat{s}$, the identity $Re = \epsilon - r$ gives

$$e \;=\; R^{+}(\epsilon - r) \;+\; (I - \Pi_S)\, e, \tag{3}$$

an orthogonal split into a measured component, fixed by the residual and the noise, and an unmeasured component in $\ker R$, which the sensor cannot see.

Users often need a derived quantity rather than the spectrum: a colour, a band value, or the two bands of an index. For a linear map F to such a quantity, let $\kappa_F(R) = \|F(I-\Pi_S)\|_2$, the spectral norm, and define the relative visibility of the null space in F,

$$\hat{\kappa}_F(R) = \frac{\|F(I-\Pi_S)\|_2}{\|F\|_2} \in [0,1]. \tag{4}$$

Then

$$\|Fe\| \;\le\; \|FR^{+}\|\,\|\epsilon - r\| \;+\; \kappa_F(R)\,\|(I-\Pi_S)e\|. \tag{5}$$

Here $\hat{\kappa}_F = 0$ holds exactly when the rows of F lie in S, the classical estimability condition [18]; then $FR^{+}y$ recovers Fs without any reconstruction when $\epsilon = 0$. Equivalently, $\kappa_F = \max\{\|Fh\| : \|h\| = 1, h \in \ker R\}$ is the largest effect of a unit invisible spectral change. Small $\hat{\kappa}_F$ means weak sensitivity to bounded null-space error; values near one mean some change in the quantity is invisible.

The undamped consistency correction $\hat{s}' = \hat{s} + R^{+}r$, the minimum-norm data-consistency update of linear inverse theory [19], [20], gives $e' = R^{+}\epsilon + (I - \Pi_S)e$: it replaces the measured error by projected noise and leaves the unmeasured error unchanged. Without noise it removes exactly the measured component, $\|e'\|^2 = \|e\|^2 - \|\Pi_S e\|^2$, and $Fe' = F(I - \Pi_S)e$ vanishes for every quantity with $\hat{\kappa}_F = 0$. The spectral statement is the Matrix-R result that substituting the measured fundamental metamer never increases error [6], [21]; the noisy case and derived quantities are what we add. With noise it helps only when the measured error exceeds the noise it reintroduces. For quantities with $\hat{\kappa}_F > 0$ there is no guarantee: the correction can increase their error when the measured component was offsetting part of the unmeasured one. For the three-channel XYZ sensor, $R = D\,\mathrm{XYZ}$ with D diagonal, so the colour rows lie in S and $\mathrm{XYZ}\,e = D^{-1}(\epsilon - r)$: without noise, XYZ error is exactly the weighted residual $\|D^{-1}r\|$, and every consistent reconstruction returns the true colour.

Instrument values. Sentinel-2 and Landsat do not determine colour: $\hat{\kappa}_{\mathrm{XYZ}}$ is 0.53 for OLI/OLI-2 and 0.62 for MSI, and higher still for single bands (median 0.98 and 1.00), so zero misfit alone does not bound their colour error. Across 368 procedural sensors the median is 0.40 (10th–90th percentile 0.10–0.77); in the lower tail, a bounded null-space error has little effect on XYZ. Vegetation and water indices are the opposite extreme: for the two grid bands each index uses, $\hat{\kappa}_F$ is 0.98–1.00 on every satellite instrument, so the measurement leaves them almost entirely undetermined.

Measurement consistency. In the three-channel experiment the unweighted relative misfit $\|R(\hat{s} - s)\|/\|Rs\|$ tracks the exact noiseless XYZ error $\|D^{-1}r\|$ closely: across the fifteen noiseless colour runs (all three estimators across 5 seeds), its Pearson correlation with CIEDE2000 is 0.94, and 0.81 under noise. In the response-description comparison, by contrast, spectral RMSE is the better predictor of colour and index error, as the visibility values above imply.

Repair by consistency projection. In practice the correction is damped, $\hat{s}' = \hat{s} + R^{\top}(RR^{\top} + \lambda I)^{-1}(y - R\hat{s})$, which leaves the null-space component untouched and shrinks the residual [22]; the identities above are its undamped limit. On the noiseless colour runs, post-hoc correction reduces the learned reconstructor's mean CIEDE2000 from 8.65 to 0.00 and lowers its spectral error, without retraining. Under noise, $\lambda = 256\sigma^2$ worked best among the damping values tested; with it, correction after training gives a mean CIEDE2000 of 4.51, against 4.80 when training through the projection, at the same spectral error.

On band-limited instruments the correction repairs only what the bands determine. On the satellite responses colour is not determined, so correction lowers the linear estimators' colour error only partly, for BJTU-UVA and Cuprite spectra alike (Fig. 2c). Indices built from an instrument's own bands are reproduced exactly without noise, whereas indices built from grid bands are not determined: correction lowers their mean absolute error from 0.028 to 0.024 on average, with no guarantee: for ridge with centre-only descriptions the error rises from 0.0401 to 0.0414.

When to correct. Does excess misfit rank the effect of correction on colour? It is the relative measurement misfit $\|R(\hat{s} - s)\|/\|Rs\|$ minus the noise level $\|y - Rs\|/\|Rs\|$. It needs the reference spectrum, so it ranks cases after the fact rather than at inference time; a case is one combination of estimator, description and procedural response bank. On the reserved banks, excess misfit was negatively associated with relative colour change (Spearman $\rho = -0.54$, meeting the prespecified threshold of -0.40; Fig. 2d): the larger the excess, the more correction lowered colour error. Case by case it is a weaker guide: 6 of 9 cases with excess misfit of at least 0.05 improved in CIEDE2000, and among 55 cases with excess misfit of at most 0.02, where correction should leave colour nearly unchanged, colour changed by up to 16% against a 5% limit. Excess misfit is therefore a useful indicator of where correction matters, but its size alone does not support a per-case decision.

VI. Conclusion

Splitting reconstruction error into measured and unmeasured parts answers both questions. For Q1, on Sentinel-2 and Landsat almost all error lies in the null space and a moment-based Gaussian performs almost as well as the full curve, whereas for sensors with wide, overlapping bands a coarse description misrepresents what is measured and costs accuracy. For Q2, without noise, undamped correction removes exactly the measured error: it guarantees quantities the sensor determines, such as colour from a colour sensor, and not those it leaves open, such as narrow-band indices from Sentinel-2 or Landsat. In practice this suggests three rules. Moment-based Gaussians suffice for narrow, well-separated bands such as those of Sentinel-2 and Landsat, while full curves are worth carrying for wide, overlapping bands or when the sensor population is not known in advance. Before relying on consistency to protect a quantity, compute its $\hat{\kappa}_F$. Apply the correction after training, damped when noise is present. Because a description's cost depends on the population and the prior as much as on the description itself, the value of a response encoding in sensor-conditioned models is best judged, as here, with the estimator and prior held fixed.

The decomposition, the visibility measure and the correction hold for any reconstruction, linear or learned. The costs of each description, by contrast, were measured for linear estimators with a 16-dimensional PCA prior, on BJTU-UVA spectra sampled at eight pixels per scene, for finite response families and one simulated noise level; the correction can also move an estimate outside that prior subspace. The learned reconstructor was trained on full responses only, so how compressing response metadata affects sensor-conditioned networks remains to be measured.

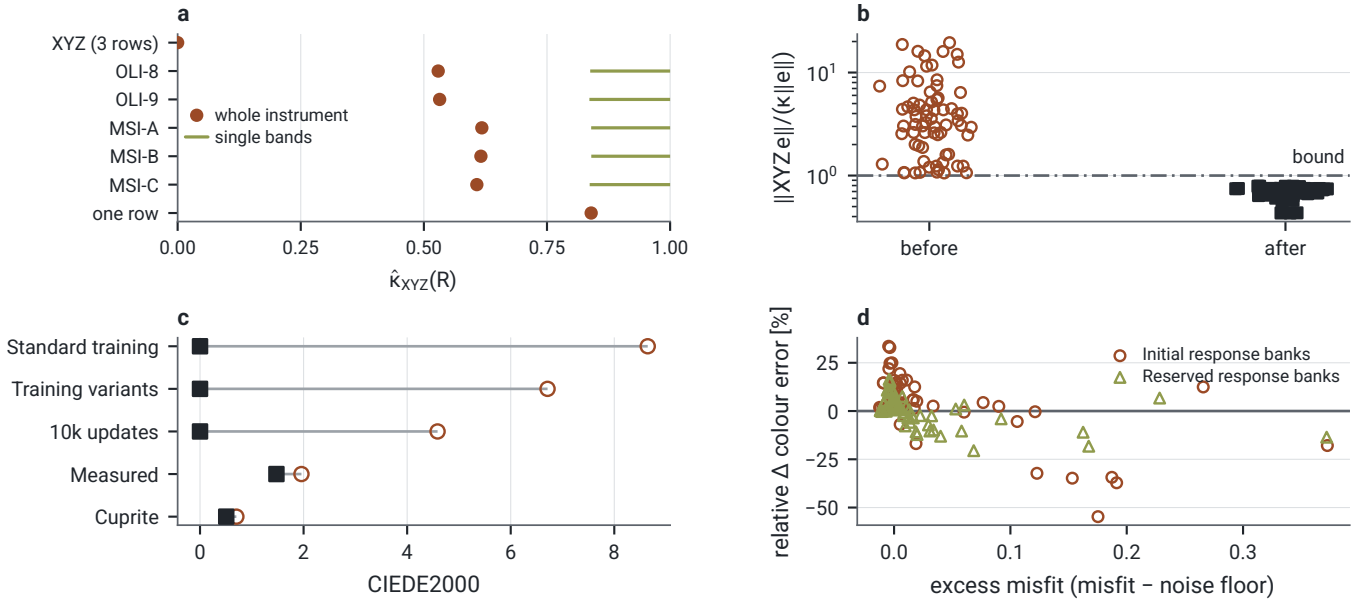


Fig. 2. What consistency correction repairs. (a) Relative colour visibility $\hat{\kappa}_{\mathrm{XYZ}}(R)$ of the null space for the XYZ sensor, Landsat 8/9 OLI (OLI-8/9), Sentinel-2A/B/C MSI (MSI-A/B/C) and a single band (one row); circles are whole instruments and bars the range over each instrument's single bands. (b) Ratio $\|\mathrm{XYZ}\, e\|/(\kappa_{\mathrm{XYZ}}\|e\|)$ before and after undamped noiseless correction on the procedural sensors (log scale). After correction Eq. (5) bounds it by one; before, it can exceed one because the bound also contains the measured error. (c) Noiseless mean CIEDE2000 before (open circles) and after (filled squares) correction: the learned reconstructor on the XYZ sensor under standard training, training variants and five-times-longer training (10k updates), and ridge and LMMSE on the satellite responses for BJTU-UVA (Measured) and Cuprite spectra. (d) Excess misfit against relative change in colour error, on the initial (circles) and reserved (triangles) procedural response banks.

References

[1] E. King, J. Rodriguez, D. Llanes, T. J. Doster, T. Emerson, and J. Koch, "STARS: Sensor-agnostic transformer architecture for remote sensing," in *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2024.

[2] J. Prexl and M. Schmitt, "SenPa-MAE: Sensor parameter aware masked autoencoder for multi-satellite self-supervised pretraining," in *DAGM German Conference on Pattern Recognition (GCPR)*, 2024.

[3] Y. Wang, Z. Xiong, C. Liu, A. J. Stewart, T. Dujardin, N.-I. Bountos, A. Zavras, F. Gerken, I. Papoutsis, L. Leal-Taixé, and X. X. Zhu, "Towards a unified Copernicus foundation model for Earth vision," 2025, arXiv preprint arXiv:2503.11849. [Online]. Available: https://arxiv.org/abs/2503.11849

[4] P. Morovič and H. Haneishi, "Estimating reflectances from multi-spectral camera responses," in *14th Color Imaging Conference*, 2006, pp. 131–137.

[5] F. Sippel, J. Seiler, and A. Kaup, "Hyperspectral image reconstruction from multispectral images using non-local filtering," in *2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP)*, 2021, pp. 1–6.

[6] G. D. Finlayson, Y.-T. Lin, and A. Kucuk, "Matrix-R theory: A simple generic method to improve RGB-guided spectral recovery algorithms," *Sensors*, vol. 25, no. 24, p. 7662, 2025.

[7] L. Chalcroft, E. Camarillo Abad, P. Christopher, T. Albrow-Owen, and O. Burton, "AnySpec: Response-conditioned spectral reconstruction across sensors," in *Proc. IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Glasgow, UK, 2026.

[8] R. Penrose, "A generalized inverse for matrices," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, no. 3, pp. 406–413, 1955.

[9] G. D. Finlayson and P. Morovič, "Metamer sets," *Journal of the Optical Society of America A*, vol. 22, no. 5, pp. 810–819, 2005.

[10] Y.-T. Lin and G. D. Finlayson, "Physically plausible spectral reconstruction," *Sensors*, vol. 20, no. 21, p. 6399, 2020.

[11] J. B. Cohen and W. E. Kappauf, "Metameric color stimuli, fundamental metamers, and Wyszecki's metameric blacks," *The American Journal of Psychology*, vol. 95, no. 4, pp. 537–564, 1982.

[12] D. Blackwell, "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, vol. 24, no. 2, pp. 265–272, 1953.

[13] Z. Du, S. You, C. Cheng, and S. Wei, "Automatic spectral calibration of hyperspectral images: Method, dataset and benchmark," in *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025, pp. 28 081–28 090.

[14] European Space Agency, "Sentinel-2 spectral response functions," https://sentiwiki.copernicus.eu/web/document-library, 2024, COPE-GSEG-EOPG-TN-15-0007, issue 4.0.

[15] U.S. Geological Survey, "Landsat spectral characteristics viewer: OLI and OLI-2 relative spectral responses," https://landsat.usgs.gov/spectral-characteristics-viewer, 2026, accessed September 8, 2026.

[16] S. K. McFeeters, "The use of the normalized difference water index (NDWI) in the delineation of open water features," *International Journal of Remote Sensing*, vol. 17, no. 7, pp. 1425–1432, 1996.

[17] T. Mansencal *et al.*, "Colour science for Python," Version 0.4.7, 2025. [Online]. Available: https://pypi.org/project/colour-science/

[18] S. R. Searle, *Linear Models*. New York: Wiley, 1971.

[19] M. M. Deal and G. Nolet, "Nullspace shuttles," *Geophysical Journal International*, vol. 124, no. 2, pp. 372–380, 1996.

[20] J. Schwab, S. Antholzer, and M. Haltmeier, "Deep null space learning for inverse problems: convergence analysis and rates," *Inverse Problems*, vol. 35, no. 2, p. 025008, 2019.

[21] Y.-T. Lin, G. D. Finlayson, and A. Kucuk, "An optimality property of matrix-R theorem, its extension, and the application to hyperspectral pan-sharpening," in *Color and Imaging Conference (CIC)*, vol. 31, 2023, pp. 144–149.

[22] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *Journal of the ACM*, vol. 9, no. 1, pp. 84–97, 1962.