Remote Sensing of Environment 290 (2023) 113534

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Surface Biology and Geology imaging spectrometer: A case study to optimize the mission design using intrinsic dimensionality

K. Cawse-Nicholson [a,*], A.M. Raiho [b], D.R. Thompson [a], G.C. Hulley [a], C.E. Miller [a], K.R. Miner [a], B. Poulter [b], D. Schimel [a], F.D. Schneider [a], P.A. Townsend [a,c], S.K. Zareh [a]

[a] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.
[b] NASA Goddard Space Flight Center, Biospheric Sciences Lab, Greenbelt, MD, USA.
[c] University of Wisconsin, Madison, WI, USA.

ARTICLE INFO

Edited by Jing M. Chen

Keywords:
Mission design
Surface biology and geology
Intrinsic dimensionality

ABSTRACT

The information content that can be derived from spectroscopic imagery tends to increase with finer ground sampling distance, finer spectral sampling, more frequent revisit, and higher signal-to-noise ratios (SNRs). However, these parameters are not independent, and it is thus impossible to design a space-borne imaging spectrometer to maximize all of them simultaneously. We present an instrument model and simulation environment that enable us to find the optimal combination of these four mission design parameters, using intrinsic dimensionality (ID) as the metric. ID is the size of the signal subspace – the maximum degrees of freedom when noise can be disregarded – and is a metric that is independent of any one particular algorithm or application area. This study is important for upcoming missions such as NASA's Earth System Observatory mission to study the Earth's Surface Biology and Geology (SBG), which will comprise a visible to shortwave infrared spectrometer in addition to a multi-channel thermal radiometer on a separate platform. When evaluating a desert site and a tropical forested site, we find that spectral resolution drives information content, with a significant drop in normalized ID (15–45% decrease) when simulating 15 nm spectral sampling as opposed to 10 nm spectral sampling. However, there was some variation between sites, with the forested site benefiting from 5 nm spectral sampling, whereas the desert site had poorer results at this resolution, due to the impact on noise. At 10 nm spectral sampling, ground sampling distances in the range 30–50 m provided the optimal balance between spatial resolution and SNR, although more frequent revisit, potentially by combining data from multiple missions, would maximize total information content.

1. Introduction

NASA has planned a series of Earth Observing missions to monitor changing climate and natural hazards, allowing for information-based decision making. The missions comprising this Earth System Observatory (ESO) are designed to answer key science questions laid out in the 2017 Earth Science Decadal Survey ((National Academies of Sciences, Engineering, & Medicine, 2018)). One of the missions proposed by the Decadal Survey (DS) is termed "Surface Biology and Geology (SBG)", which will study terrestrial ecosystems, agriculture, coastal and inland waters, snow and ice, surface geology, and volcanic eruptions, among other observables. The recommended mission will consist of two independent platforms: one a visible to shortwave infrared (VSWIR) spectrometer; and the other a multi-band thermal infrared (TIR) radiometer (Cawse-Nicholson et al., 2021; Stavros et al., 2022). The DS ranked all of the most critical science and applications questions and assigned each to one of three prioritization categories: "Most Important", "Very Important", and "Important". SBG will answer several science questions ranked "Most Important" by the DS (National Academies of Sciences, Engineering, & Medicine, 2018):

- *How is the water cycle changing? Are changes in evapotranspiration and precipitation accelerating, with greater rates of evapotranspiration and thereby precipitation, and how are these changes expressed in the space-time distribution of rainfall, snowfall, evapotranspiration, and the frequency and magnitude of extremes such as droughts and floods?*
- *What are the structure, function, and biodiversity of Earth's ecosystems, and how and why are they changing in time and space?*

* Corresponding author.
E-mail address: Kerry-anne.cawse-nicholson@jpl.nasa.gov (K. Cawse-Nicholson).

https://doi.org/10.1016/j.rse.2023.113534
Received 20 October 2022; Received in revised form 23 February 2023; Accepted 4 March 2023
Available online 10 March 2023

- *What are the fluxes (of carbon, water, nutrients, and energy) between ecosystems and the atmosphere, the ocean, and the solid Earth, and how and why are they changing?*
- *What are the fluxes (of carbon, water, nutrients, and energy) within ecosystems, and how and why are they changing?*
- *How can large-scale geological hazards be accurately forecast in a socially relevant time frame?*

The DS also poses a number of additional "Important" and "Very Important" science questions. It is clear that these span a wide range of applications, and the optimal mission design for each science question may differ. Here we illustrate a method for determining optimal mission design, focusing on the VSWIR spectrometer proposed to form part of SBG.

ID is a metric for information content and can be thought of as the size of the signal subspace. In other words, ID is the number of significant principal components of a spectroscopic image (Asner et al., 2012; Thompson et al., 2017; Cawse-Nicholson et al., 2013, 2019). ID has also explicitly been shown to be an application-agnostic metric for spectroscopic mission design (Cawse-Nicholson et al., 2022). When used in a sensitivity analysis for a wide range of airborne scenes, ID decreased with a lower signal-to-noise ratio (SNR), larger ground sample distance (GSD), coarser spectral sampling, and less frequent revisit (Cawse-Nicholson et al., 2022). The optimal mission will have the maximum overall ID. Consequently, the optimal instrument design may be thought to be the highest possible SNR, and highest spatial, spectral, and temporal resolutions. However, these variables are not independent. Very small pixels with narrow spectral bands result in less reflected solar energy reaching the sensor, and therefore a lower SNR. The same instrument could increase revisit by increasing swath size (through altitude or telescope), but the consequence would be larger GSD (pixel size).

In this manuscript, we will first describe a realistic instrument model that relates the above parameters; next we will detail the Hypertrace model (Brodrick et al., 2021; Raiho et al., 2022) that is used to simulate different design parameters, such as SNR, GSD, and spectral sampling; and finally, we will demonstrate the optimal results based on the appropriate instrument model.

2. Methods

Hypertrace is an open-source tool that accepts user-specified reflectances, atmospheric conditions, solar and observation geometries, and instrument parameters, and uses these to simulate the radiance received by the sensor and subsequently retrieve the surface reflectance and atmospheric parameters (Brodrick et al., 2021; Raiho et al., 2022). The radiative transfer calculations to simulate the top of atmosphere radiance are approximated by a radiative transfer emulator that is included with the Hypertrace package, and the retrieval of surface reflectance uses the Imaging Spectrometer Optimal FITting codebase (ISOFIT; Thompson et al., 2018). Many atmospheric correction algorithms first estimate atmospheric parameters using known spectral intervals, and then invert a radiative transfer model to solve for surface reflectance. ISOFIT solves for both the atmospheric and surface parameters simultaneously, which captures spectrally-broad atmospheric perturbations as well as the well-constrained narrow-band features, without making assumptions on surface or atmospheric conditions. By posing the problem as a probabilistic formulation, ISOFIT accounts for uncertainties in all inputs, and provides a closed account of observed errors. ISOFIT uses the measured radiance y to estimate a state vector x, which is a concatenation of surface reflectance in all bands, aerosol optical depth, and atmospheric water content. In optimal estimation formulation, the forward simulation f (in this case a radiative transfer model) is written such that $y = f(x) + \epsilon$, where ϵ represents random noise. Maximum A Posteriori inversion of this forward radiative transfer model is used to solve for surface reflectance as well as aerosol optical depth and atmospheric water content. This amounts to a Bayesian-based approach that determines the Gaussian distribution of a state vector that best predicts the observations, given a prior probability (Thompson et al., 2018).

By simultaneously solving for surface and atmospheric parameters, the model is able to fully optimize the reflectance estimate. This enables us to determine the impact of various atmospheric and instrument parameters on retrieval accuracy. The use of Hypertrace for SBG-like simulations is fully described in previous work (Cawse-Nicholson et al., 2022; Raiho et al., 2022). For this study, our Hypertrace parameters assumed: standard MODTRAN-defined "mid-latitude summer" atmospheric profile (Anderson et al., 1986); 0.1 aerosol optical depth; 1 cm atmospheric water vapor content; nadir viewing angle; 1000 local solar time (37.21° solar zenith angle, and 53.82° solar azimuth angle). These are reasonable atmospheric parameters for the United States, however, we have repeated this experiment with various atmospheric parameters and found our results unchanged. All simulation parameters are captured in the Appendix.

For the reflectance input into Hypertrace, two airborne images were selected: one acquired over the Cuprite desert site in Nevada, USA, and the other acquired over a tropical forest in Kerala, India. Both were acquired by the Airborne Visible InfraRed Imaging Spectrometer - Next Generation (AVIRIS-NG), which has a spectral range of 380–2500 nm and 5 nm spectral sampling. The images had <5% cloud cover, and the image properties are described in Table 1, with the images shown in Fig. 1. These two images were processed using ISOFIT and manually evaluated by an expert to ensure the optimal estimation of the reflectance truth (in particular, the expert ensured that no visible anomalies or artifacts remained that could be propagated through our experiment). For our experiments, when spectral resampling is necessary, the spectral bands are linearly interpolated for each pixel. When spatial resampling is necessary, each band is independently resampled using bilinear interpolation.

The desert site near Cuprite, Nevada has been used in numerous remote sensing geological studies due to the lack of vegetation and the exposed minerology (e.g. Ashley and Abrams, 1980; Kruse et al., 1993; Swayze et al., 2014). The site is near the Stonewall Mountain caldera, and the surface geology is primarily tertiary volcanic rock and quarternary deposits, including tertiary rhyolitics, conglomerates and basalt in the East, and Cambrian metasedimentary rocks, tertiary conglomerates, and intrusives in the West (Swayze et al., 2014). The topical forest site is located in the Sholayar Reserve Forest in Kerala province, India. This site is considered a biodiversity hotspot within India, and contains many tree species, such as the Black Varnish Tree (*Holigarna arnottiana*), Silk Oak (*Grevillea robusta*), Dhaman (*Grewia tiliifolia*), Java Plum (*Syzygium cumini*), and many others, as well as plantations of tea, coffee, rubber, and teak (Ahmad et al., 2020).

The AVIRIS-NG data are sufficient for the analysis of GSD, spectral sampling, and SNR, but do not enable a temporal analysis. In a previous sensitivity study (Cawse-Nicholson et al., 2022), VIIRS daily reflectance data over 11 biodiversity hotspots and 2 agricultural regions were used to evaluate the impact of revisit. Here we will consider those sensitivity results in the context of our instrument model, by tying the revisit and GSD (i.e. a wider telescope with the same instrument could enable more frequent observations, but at the cost of coarser GSD). If we normalize

Table 1
Properties of the AVIRIS-NG images used as Hypertrace inputs. The two images have a different number of bands due to changes made to the internal processing pipeline.

Biome	Location	# bands	GSD	Image size (pixels)	Flying altitude	Date of acquisition
Desert	Nevada, USA	432	3.7 m	1200 × 1000	18,300 ft	2014-06-25
Tropical Forest	Kerala, India	425	3.6 m	501 × 209	13,300 ft	2016-01-07

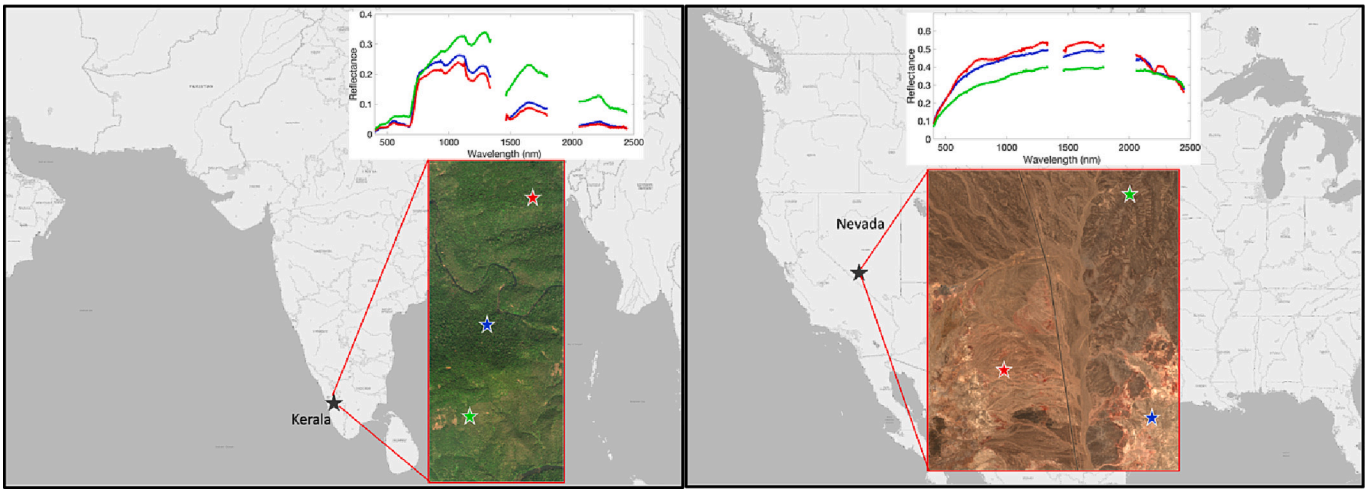

Fig. 1. Red (630 nm), green (530 nm), and blue (465 nm) composites of the tropical forest scene (left) and desert scene (right). Black stars show the spatial location of each image, overlaid on the ESRI basemap in QGIS. Red, green and blue stars show the locations of the sample reflectance spectra for each image. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the VIIRS ID (VNID) produced by Cawse-Nicholson et al. (2022) by the value at an 8-day revisit, then:

Revisit (GSD)	8-day (60 m)	10-day (50 m)	12-day (40 m)	16-day (30 m)
VNID	1	0.86	0.73	0.59

The instrument model used in this study is heavily influenced by the Earth Mineral dust source InvesTigation (EMIT) instrument, which was launched to the International Space Station (ISS) in 2022. EMIT is an imaging spectrometer with a spectral range of 380–2500 nm, a bandwidth of 7.4 nm, and 288 spectral channels. Due to its position on the ISS it has an orbiting altitude of ~400 km, which results in a ground sample distance of ~60 m at nadir (Connelly et al., 2021). The SBG VSWIR spectrometer is still in the design phase, but for the purposes of this experiment, we have assumed that the SBG spectrometer will be a Dyson optical layout like EMIT. EMIT utilizes a Dyson imaging spectrometer with an optical throughput of F/1.8. Reflected solar energy enters the slit, passing through the refractive element to the grating where light is spectrally dispersed. The light returns through the refractive element to a CHROMA-D (Teledyne Imaging Sensors, Inc.) HgCdTe detector array with a range of 380–2500 nm (Green et al., 2020). If two CHROMA-D spectrometers were flown side-by-side at an altitude of 705 km, the instrument could be configured to achieve a 30 m GSD and a 16-day revisit. In our instrument model, we modified the four parameters under consideration (signal-to-noise ratio, ground sampling distance, spectral sampling, and revisit). We treated the orbital altitude, system f-number and detector size as constants, and the spectral sampling and telescope focal length as free parameters to be optimized. The spectrometer is assumed to have a detector pitch of 0.0018 cm and a spectrometer slit width of 18 μm (note that this may not be the final SBG design). All simulation parameters are captured in the Appendix.

The noise was simulated using the model described in Thompson et al. (2018, 2020) which accounts for noise sources including electronic readout, dark current, grating throughput, and internal instrument reflections. The noise is drawn from a multivariate Gaussian distribution such that the standard deviation of change in radiance (i.e. noise) $\sigma_\lambda(\boldsymbol{L}) = \boldsymbol{a}_\lambda \sqrt{\boldsymbol{b}_\lambda \bullet \boldsymbol{L}_\lambda} + \boldsymbol{c}_\lambda$, where $\mathbf{L}$ is the observed radiance at the center wavelength of each spectral band, λ (Thompson et al., 2018). This noise representation is based on historical convention (Thompson et al., 2018), and a represents efficiency (such as throughput loss through gratings, quantum efficiency, and dark current), b represents photon noise, which typically dominates the error, and c represents additional noise due to photon counting and electronic readout errors (this is typically close to zero). In this work, the parameters $\boldsymbol{a,b,c}$, are modelled based on EMIT. The model is dependent on the optical system's f-number, detector pitch, spectral band centers and widths, integration time, and various efficiency parameters. In our experiments, we varied the spectral bands as described in each experiment, and we varied the integration time to account for ground sample distance and SNR. All other parameters remained as per the EMIT model (Thompson et al., 2018, 2020). In the instrument model, a 30 m GSD corresponds to an integration time of 0.0044 s and a 16-day revisit. Coupled with 10 nm band spacing, our noise model produced an SNR curve shown in Fig. 2, which has a median SNR of 456 in the visible to near infrared (VNIR; < 1000 nm) and a median SNR of 309 in the shortwave infrared (SWIR; > 1000 nm). For the purpose of the illustration in Fig. 2, the reference observation is a nadir-view acquisition with a clear sky (rural aerosols, 50 km visibility), for a surface with constant 25% reflectance at sea-level with a 23.4° solar zenith angle. These conditions were chosen for the illustration as they represent a very clear atmosphere, so Fig. 2 shows

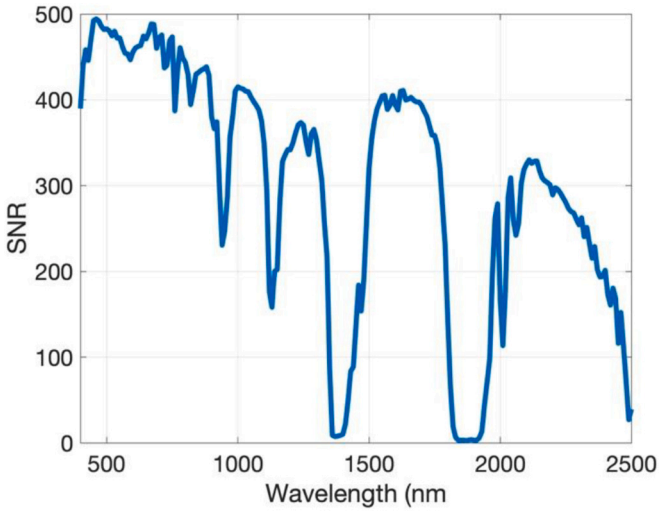

Fig. 2. The signal-to-noise ratio at different wavelengths for a 30 m GSD and 10 nm spectral sampling, assuming the EMIT-based instrument model. Note that atmospheric water absorptions in the atmosphere (SNR close to 0) near 1400 and 1900 nm effectively block all reflected energy from reaching the instrument. Energy in the rest of the spectrum reaches the instrument through so-called "atmospheric windows".

only instrument noise, without any atmospheric effects. The relationship between GSD and revisit, and GSD and median SNR (across all wavelengths) is shown in Fig. 3.

In our experiments, we laid out a grid of GSD values (10, 20, 30, 40, 50, 60 m) and spectral sampling intervals (5, 10, 15, 20, 25, 30 nm). For each combination of GSD and spectral sampling interval we computed the integration time using the instrument model, and used the noise model to predict the SNR based on the integration. These three variables (GSD, spectral sampling, and SNR) are used as Hypertrace inputs to derive representative reflectance images. Revisit is assumed to be linearly related to GSD (e.g. 30 m GSD implies 16-day revisit, and 60 m GSD implies 8-day revisit).

In summary, a reflectance image, spectral sampling, and GSD are used as inputs. An instrument model uses spectral sampling and GSD to determine SNR, and uses GSD to determine revisit. Hypertrace uses all these parameters to simulate a reflectance image we might retrieve from the chosen instrument design parameters. ID is the metric we use to determine the science benefit of each design choice. If the Hypertrace simulated reflectance image is scaled and mean-centered for each pixel, it can be written as the n x b matrix X, where n is the number of pixels, and b is the number of bands, then ID computes the number of significant eigenvalues of the image covariance matrix X^TX. Here we use the Random Matrix Theory method to determine the ID, which is fully described in previous work (Cawse-Nicholson et al., 2019; Cawse-Nicholson et al., 2022). The full experiment pipeline is summarized in Fig. 4, and a sample of the input reflectance, estimated surface reflectance, and top of atmosphere radiance is shown in Fig. 5.

3. Results

Both images were used to generate simulated reflectance and resulting ID for GSD values (10, 20, 30, 40, 50, 60 m) and spectral sampling intervals (5, 10, 15, 20, 25, 30 nm). For each image, the ID is normalized (NID) by the highest ID value across all experiments (max ID = 43 in the forested scene, and max ID = 70 in the desert scene).

Fig. 6 shows the NID for both scenes. The tropical forest in Kerala, India has its maximum information content (using NID as a proxy) at 60 m GSD and 5 nm spectral sampling. At 10 nm spectral sampling, 30 m GSD has the maximum NID value. At GSD < 30 m, the SNR is low (not enough photons are captured when both pixel size and bandwidth are small), resulting in a lower NID for 10 nm spectral sampling, and even more substantially for 5 nm spectral sampling with small GSD. The NID remains relatively high for all GSD ≥ 30 m at 5 nm and 10 nm spectral resolution. There is a significant (~15%) drop in ID for a spectral sampling of 15 nm compared to 10 nm sampling, highlighting the importance of detecting fine spectral features for discriminating species and other environmental conditions.

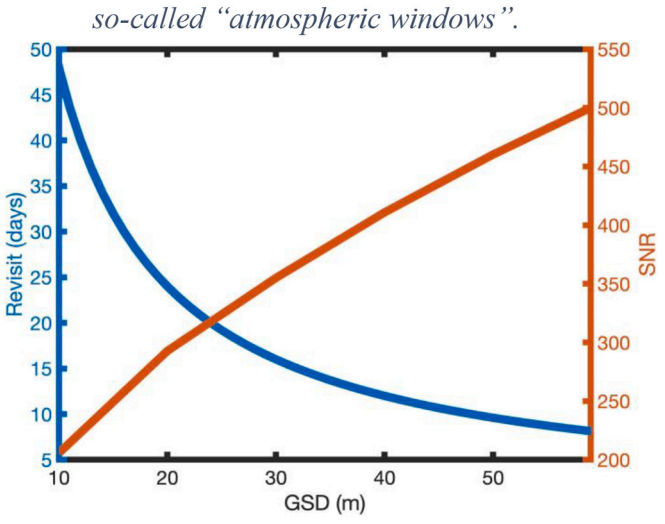

Fig. 3. GSD is related to both revisit and SNR by the instrument model. Here, a 30 m GSD is associated with a 16-day revisit and a median SNR of 355, assuming 10 nm spectral resolution.

The desert site in Cuprite, Nevada, USA, has its maximum information content (using NID as a proxy) at 50 m GSD and 10 nm spectral sampling. In particular, 10 nm was clearly the optimum spectral sampling, with the NID remaining relatively high for a GSD of 30–50 m. In comparison with the forested scene, there is an even more significant drop in NID as spectral bands widen from 10 nm, although at 5 nm spectral sampling the NID is monotonically decreasing for progressively smaller GSD. This is likely due to the fact that minerals are often differentiated by very narrow spectral features (i.e. requiring finer than 10 nm sampling), but these may be difficult to detect when noise levels are high (i.e. 5 nm spectral sampling does not enable enough photons within that wavelength range to reach the instrument).

When averaging the NID values for both scenes, the highest NID_{avg} = 0.93 at 10 nm spectral sampling and 50 nm GSD, followed by NID_{avg} = 0.92 at 10 nm spectral sampling and 40 m GSD, NID_{avg} = 0.91 at 10 nm spectral sampling and 30 m GSD, and NID_{avg} = 0.90 at 5 nm spectral sampling and 60 m GSD.

However, the Hypertrace runs alone do not independently show the impact of revisit. In other words, in Fig. 6, the 30 m GSD is associated with a 16-day revisit, whereas the 60 m GSD is associated with an 8-day revisit. But there is additional information to be gained through more frequent revisit. In a previous study (Cawse-Nicholson et al., 2022), daily VIIRS data was used to investigate the impact of revisit on ID across a wide range of biomes. Using the VIIRS normalized ID (VNID) as a weighting, we can compare the impact of revisit. This is emphasized in Fig. 7, where the ID from the forested image in Fig. 6 is multiplied by the VNID vector and re-normalized. This shows that while a single 30 m image may contain more information than a 60 m image (NID = 0.88 vs. NID = 0.84 at 10 nm spectral sampling), there is significant information gain in more frequent revisit (VNID = 0.59 vs. NID = 1 for 16-day vs. 8-day revisit). However, when considering design decisions, at 10 nm spectral sampling, a 60-m/16-day instrument would yield only ~half the information content compared to a 30-m/8-day instrument when considering both spatial and temporal dimensions (VNID x NID = 0.5 vs. VNID x NID = 0.88).

4. Discussion

In prior work (Cawse-Nicholson et al., 2022), we showed that ID is a good metric for assessing instrument design parameters. While specific algorithms use particular parts or features of the spectrum, ID captures the full size of the signal-subspace, within which all the algorithm solutions reside. ID is also an agnostic way of comparing parameters that have different implications for different algorithms. For example, if one design change results in improvements in one application but decreases performance in another, then how do we weigh the value of that design decision overall? ID enables a fair comparison, and does not require the creation of any higher-level products. In addition, because ID is a mathematical property of the image, there is no need to unmix the image or map to species or classes in order to determine the ID.

However, our results are dependent on our inputs. Here we chose a desert scene and a forested scene, and the impact of design decisions was slightly different. Most notably, the desert scene was more strongly impacted by changes in spectral sampling, although 5 nm sampling was an advantage for the forested scene but a disadvantage for the desert scene. The desert scene also had a sudden drop in NID at 60 m spatial resolution, for 10 nm and 15 nm spectral sampling. Small mineral outcrops were no longer detected when the spatial resolution became coarse, but the more homogenous forested site was not impacted as significantly. Both scenes had potential solutions around 30–50 m GSD for 10 nm spectral sampling. Future studies over a wider range of biomes

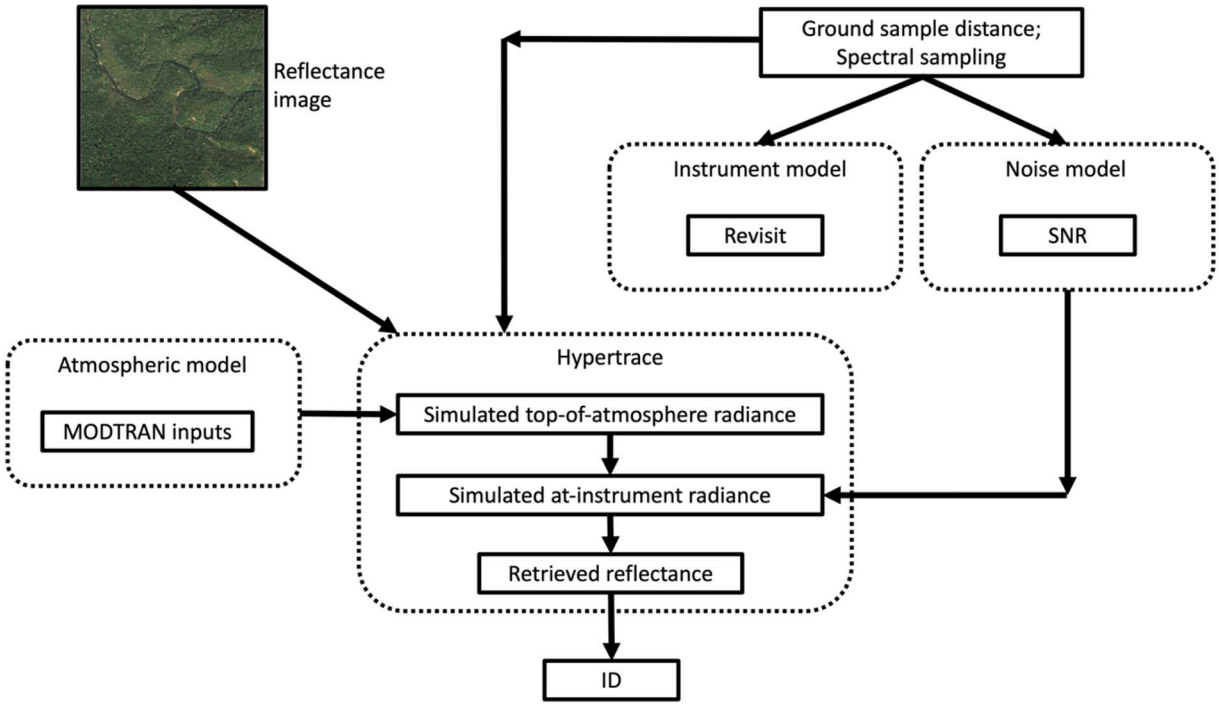

Fig. 4. A flow diagram showing the model inputs and outputs. GSD and spectral are used to calculate revisit and SNR in the instrument and noise models, respectively. The "true" reflectance image, alongside spectral sampling, SNR, and GSD are used in Hypertrace to simulate retrieved reflectances, and ID is used as a metric to determine the impact of each design choice.

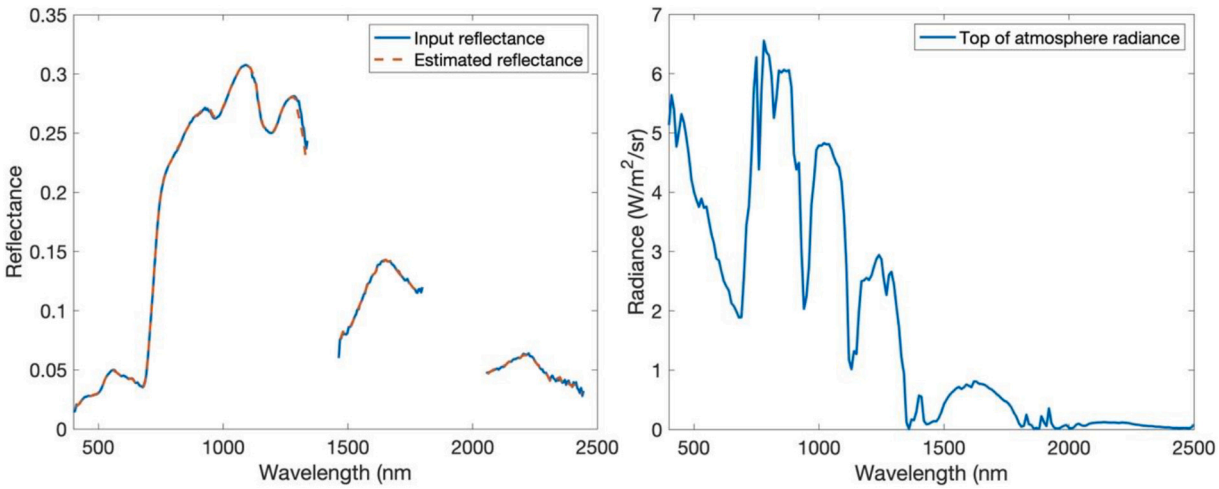

Fig. 5. For a single sample pixel within the forested scene at 30 m GSD, we show (left) the input surface reflectance alongside the estimated reflectance following the experiment shown in Fig. 4, and (right) the top of atmosphere radiance from which the estimated reflectance is derived.

would enable a more detailed study of the global impact.

In our results, we also demonstrated the trade-off between revisit and GSD. However, it should be noted that it is not yet clear if the scales of temporal and spatial ID are comparable, and this should be investigated in high-frequency spectroscopic data. Until recently, this data has not been available, which is why VIIRS was used as a proxy in the preliminary study. But the SBG HIgh Frequency Time series (SHIFT) airborne campaign is currently underway and will provide weekly AVIRIS-NG acquisitions over parts of Southern California over a 5-month period, which will be evaluated in future analysis.

Aquatic scenes should also be considered in the future in order to fully understand the impact to SBG science goals. Such scenes were not considered in this analysis because small artifacts in the "truth image" that may remain after standard atmospheric correction (including the impact of glint) are enhanced during these simulations, and these artifacts tend to dominate the results, especially over dark (water) targets. A simulated scene may provide a better candidate for such analysis (e.g. the simulated environment described by Kravitz et al., 2021). In contrast, scenes containing snow should also be considered in future studies, since these are bright targets that can identify other problems such as saturation. Simulated scenes over all surface types are especially beneficial for these studies (including forward modeling for simulating canopy reflectance), since the truth is perfectly known, and the user has control over many variables which could impact the findings, such as scene complexity, seasonal changes, etc. Future studies could also consider the impact of variable spectral sampling in different parts of the spectrum, enabling a tradeoff between spectral sampling and SNR.

The Decadal Survey suggested different parameters for different

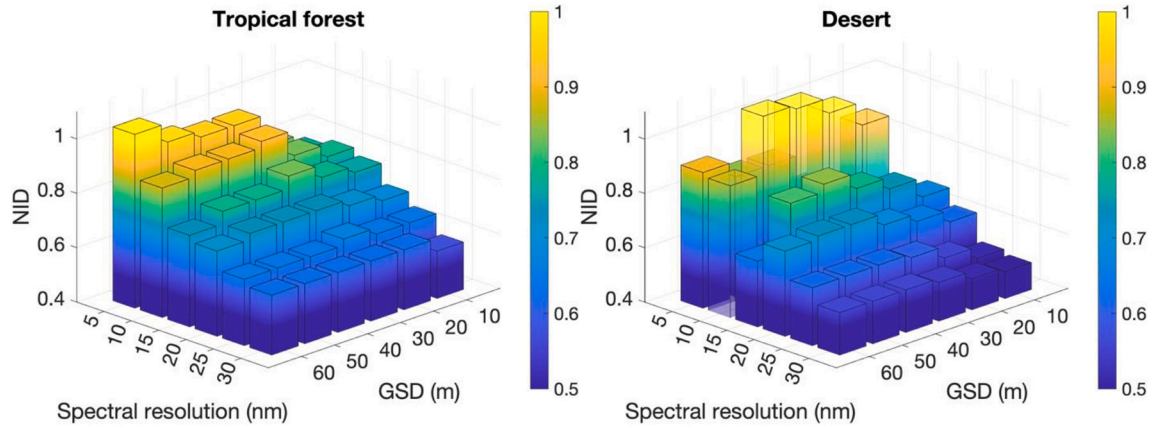

Fig. 6. NID for the forested scene (left) and desert scene (right) is shown for all combinations of GSD and spectral sampling. These parameters also drive SNR and revisit. Both images have high NID at 10 nm spectral sampling and 30–50 m spatial resolution, although the forested scene had the maximum NID at 5 nm spectral sampling.

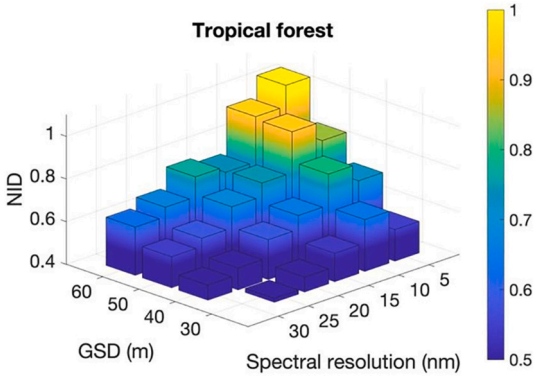

Fig. 7. The forested image from Fig. 6 has been normalized by temporal ID to illustrate the impact of frequent revisit on overall information content.

applications, but a VSWIR mission with 10 nm spectral sampling, 30 m GSD, SNR $\geq$ 400 in the VSIR and SNR $\geq$ 250 in the SWIR, and 8-day revisit, would meet most of the suggested design parameters. While this guidance serves as a baseline for the mission concept, we have used a mathematical solution to independently verify the optimum mission design. Here, we see a high-NID solution at 10 nm spectral sampling, and 30–50 m GSD (associated with median SNR = 456 in the VNIR and SNR = 309 in the SWIR for 30 m GSD, and SNR = 648 in the VNIR and SNR = 429 in the SWIR for 50 m GSD). However, the instrument model used here assumed two co-mounted CHROMA-D instruments, which would allow 16-day revisit at 30 m GSD, which is longer than that recommended by the DS and may decrease our ability to detect short term events (Raiho et al., 2022).

We could reconfigure the same instrument to enable 8-day revisit by increasing the GSD to 60 m. That would enable an increase in information content in the temporal dimension for vegetated sites such as the tropical forest seen here, but that GSD is too coarse to optimally detect small mineral outcrops in the desert site. A more beneficial solution may be to consider combining data with other proposed spectroscopic missions, such as the European Space Agency's CHIME (Copernicus Hyperspectral Hyperspectral Imaging Mission for the Environment). If the two instruments have compatible orbits and spectral and radiometric properties, then a data harmonization could result in 30 m data every 8-days.

It is encouraging to find that a mathematical evaluation produced similar results to the parameters recommended by experts in the DS.

5. Conclusion

Historically, many missions have been designed based on expert opinion, often driven by experience with airborne data. Here we have proposed a mathematical method to optimize mission design by relating design parameters (GSD, spectral sampling, SNR and revisit) using an instrument model, and simulating what images retrieved by such an instrument design might look like. ID is the algorithm-independent metric we chose to weigh the benefit of each combination of design parameters. When evaluating two different images – one over a desert and the other over a tropical forest – we found that the optimal instrument design was very sensitive to spectral sampling, with 10 nm resolution significantly better than 15 nm in both images, and 5 nm spectral sampling advantageous for the forested site but not the desert site, due to the consequent impact on SNR. For 10 nm spectral sampling, we found that the optimal spatial resolution was between 30 and 50 m (which resulted in a median SNR of 355 and 460, respectively). Temporal resolution was also significant, as shown by a VIIRS analysis, although a shorter revisit may be better achieved through harmonization with similar proposed missions. Our mathematical evaluation produced similar results to the parameter recommended by experts in the DS and provide a path forward for making quantitatively formed mission design decisions.

Open research

AVIRIS-NG reflectance data are available fromhttps://avirisng.jpl.nasa.gov/dataportal/ The code for both ISOFIT and Hypertrace is available via Zenodo: https://doi.org/10.5281/zenodo.4614337.

CRediT authorship contribution statement

K. Cawse-Nicholson: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **A.M. Raiho:** Writing – review & editing. **D.R. Thompson:** Methodology, Software, Writing – review & editing. **G.C. Hulley:** Writing – review & editing. **C.E. Miller:** Conceptualization, Project administration, Writing – review & editing. **K.R. Miner:** Project administration, Writing – review & editing. **B. Poulter:** Writing – review & editing. **D. Schimel:** Conceptualization, Methodology, Project administration, Writing – review & editing. **F.D. Schneider:** Writing – review & editing. **P.A. Townsend:** Writing – review & editing. **S.K. Zareh:** Methodology, Software, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

AVIRIS-NG reflectance data are available from https://avirisng.jpl.nasa.gov/dataportal/

Acknowledgments

We thank three anonymous reviewers who provided valuable feedback that significantly improved this manuscript. Support to all authors was provided by NASA and the Jet Propulsion Laboratory, California Institute of Technology and to PAT by NSF Macrosystems Biology and NEON Early Science award 1638720. This study was also supported by the Space-based Imaging Spectroscopy and Thermal (SISTER) pathfinder, and the Surface Biology and Geology (SBG) project, a NASA Earth Science Designated Observable. The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Appendix A. Appendix

Table A1
All simulation parameters are shown here for the atmospheric model (within Hypertrace) and the instrument and noise models.

Hypertrace parameters		
Atmospheric model	MODTRAN definition	Mid-latitude summer
	Aerosol optical depth	0.1
	Atmospheric water vapor	1 cm
	Viewing angle	Nadir (0°)
	Local solar time	1000
	Solar zenith angle	37.21°
	Solar azimuth angle	53.82°
Instrument/noise model	Instrument type	CHROMA-D
	Orbiting altitude	705 km
	Detector pitch	0.0018 cm
	Spectrometer slit width	18 μm
	Ground sample distance	{10, 20, 30, 40, 50, 60 m}
	Spectral sampling interval	{5, 10, 15, 20, 25, 30 nm}
	F number	1.8

References

Ahmad, S., Pandey, A.C., Kumar, A., Lele, N.V., Bhattacharya, B.K., 2020. Forest health estimation in sholayar reserve Forest, Kerala using AVIRIS-NG hyperspectral data. Spat. Inf. Res. 28 (1), 25–38. https://doi.org/10.1007/s41324-019-00260-6.

Anderson, G.P., Clough, S.A., Kneizys, F.X., Chetwynd, J.H., Shettle, E.P., 1986. AFGL atmospheric constituent profiles (0.120 km). AIR FORCE GEOPHYSICS LAB HANSCOM AFB MA. https://apps.dtic.mil/sti/pdfs/ADA175173.pdf.

Ashley, R.P., Abrams, M.J., 1980. Alteration mapping using multispectral images; Cuprite mining district, Esmeralda County, Nevada. https://doi.org/10.3133/ofr80367. US Geological Survey Open File Report (No. 80-367).

Asner, G.P., Knapp, D.E., Boardman, J., Green, R.O., Kennedy-Bowdoin, T., Eastwood, M., Field, C.B., 2012. Carnegie airborne Observatory-2: increasing science data dimensionality via high-fidelity multi-sensor fusion. Remote Sens. Environ. 124, 454–465. https://doi.org/10.1016/j.rse.2012.06.012.

Brodrick, P.G., Thompson, D.R., Fahlen, J.E., Eastwood, M.L., Sarture, C.M., Lundeen, S. R., Olson-Duvall, W., Carmon, N., Green, R.O., 2021. Generalized radiative transfer emulation for imaging spectroscopy reflectance retrievals. Remote Sens. Environ. 261, 11247. https://doi.org/10.1016/j.rse.2021.112476.

Cawse-Nicholson, K., Damelin, S.B., Robin, A., Sears, M., 2013. Determining the intrinsic dimension of a hyperspectral image using random matrix theory. IEEE Trans. Image Process. 22 (4), 1301–1310. https://doi.org/10.1109/TIP.2012.2227765.

Cawse-Nicholson, K., Hook, S.J., Miller, C.E., Thompson, D.R., 2019. Intrinsic dimensionality in combined visible to thermal infrared imagery. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12 (12), 4977–4984. https://doi.org/10.1109/JSTARS.2019.2938883.

Cawse-Nicholson, K., Townsend, P.A., Schimel, D., Assiri, A.M., Blake, P.L., Buongiorno, M.F., et al., 2021. NASA's surface biology and geology designated observable: a perspective on surface imaging algorithms. Remote Sens. Environ. 257, 112349 https://doi.org/10.1016/j.rse.2021.112349.

Cawse-Nicholson, K., Raiho, A., Thompson, D.R., Hulley, G., Miller, C.E., Miner, K., Zareh, S.K., 2022. Intrinsic dimensionality as a metric for the impact of mission design parameters. Journal of geophysical research. Biogeosciences 127 (8), e2022JG006876. https://doi.org/10.1029/2022JG006876.

Connelly, D.S., Thompson, D.R., Mahowald, N.M., Li, L., Carmon, N., Okin, G.S., Green, R.O., 2021. The EMIT mission information yield for mineral dust radiative forcing. Remote Sens. Environ. 258, 112380 https://doi.org/10.1016/j.rse.2021.112380.

Green, R.O., Mahowald, N., Ung, C., Thompson, D.R., Bator, L., Bennet, M., Zan, J., 2020. The earth surface mineral dust source investigation: An earth science imaging spectroscopy mission. In: 2020 IEEE Aerospace Conference, pp. 1–15. https://doi.org/10.1109/AERO47225.2020.9172731.

Kravitz, J., Matthews, M., Lain, L., Fawcett, S., Bernard, S., 2021. Potential for high fidelity global mapping of common inland water quality products at high spatial and temporal resolutions based on a synthetic data and machine learning approach. Front. Environ. Sci. 9, 587660 https://doi.org/10.3389/fenvs.2021.587660.

Kruse, F.A., Lefkoff, A.B., Dietz, J.B., 1993. Expert system-based mineral mapping in northern Death Valley, California/Nevada, using the airborne visible/infrared imaging spectrometer (AVIRIS). Remote Sens. Environ. 44 (2–3), 309–336. https://doi.org/10.1016/0034-4257(93)90024-R.

National Academies of Sciences, Engineering, & Medicine, 2018. Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space. National Academies Press, Washington, DC. https://doi.org/10.17226/24938.

Raiho, A.M., Cawse-Nicholson, K., Chlus, A., Dozier, J., Gierach, M., Miner, K., Schneider, F., Schimel, D., Serbin, S., Shiklomanov, A.N., Thompson, D.R., Townsend, P.A., Zareh, S., Skiles, M., Poulter, B., 2022. Exploring mission design for imaging spectroscopy retrievals for land and aquatic ecosystems. J. Geophys. Res. Biogeosci. https://doi.org/10.1002/essoar.10510949.1.

Stavros, E.N., Chrone, J., Cawse-Nicholson, K., Freeman, A., Glenn, N., et al., 2022. Designing an observing system to study the surface biology and geology (SBG) of the earth in the 2020s. JGR. Biogeosciences 127, e2021JG006471. https://doi.org/10.1029/2021JG006471.

Swayze, G.A., Clark, R.N., Goetz, A.F., Livo, K.E., Breit, G.N., Kruse, F.A., Ashley, R.P., 2014. Mapping advanced argillic alteration at cuprite, Nevada, using imaging spectroscopy. Econ. Geol. 109 (5), 1179–1221. https://doi.org/10.2113/econgeo.109.5.1179.

Thompson, D.R., Boardman, J.W., Eastwood, M.L., Green, R.O., 2017. A large airborne survey of Earth's visible-infrared spectral dimensionality. Opt. Express 25 (8), 9186–9195. https://doi.org/10.1364/OE.25.009186.

Thompson, D.R., Natraj, V., Green, R.O., Helmlinger, M.C., Gao, B.C., Eastwood, M.L., 2018. Optimal estimation for imaging spectrometer atmospheric correction. Remote Sens. Environ. 216, 355–373. https://doi.org/10.1016/j.rse.2018.07.003.

Thompson, D.R., Braverman, A., Brodrick, P.G., Candela, A., Carmon, N., Clark, R.N., Connelly, D., Green, R.O., Kokaly, R.F., Li, L., Mahowald, N., 2020. Quantifying uncertainty for remote spectroscopy of surface composition. Remote Sens. Environ. 247, 111898 https://doi.org/10.1016/j.rse.2020.111898.