INTEGRATING PHYSICS DEDUCTIVE BIASES INTO LEARNING FOR HYPERSPECTRAL CLASSIFICATION

Romain Thoreau[1]*, **Laurent Risser**[2], **Véronique Achard**[3], **Béatrice Berthelot**[4] **& Xavier Briottet**[3]
[1]CNES, [2]Toulouse Mathematics Institute - CNRS, [3]ONERA-DOTA, [4]Magellium, France

ABSTRACT

Airborne hyperspectral imaging has great potential for land cover mapping with very detailed nomenclatures thanks to its spectral dimension, which is highly informative about the chemical composition of matter. In the past years, the scarcity of training data in regards to the significant spectral intra-class variability has motivated the development of inductive biases for the generalization of deep learning classification models. In contrast, we investigate in this paper an orthogonal line of research which consists in integrating deductive biases derived from a priori physical knowledge into weakly supervised learning in order to improve the robustness of classification models to changes in local irradiance conditions. Our experiments on simulated and real data demonstrate the benefits of our method in terms of classification accuracy. Code & data: `https://github.com/Romain3Ch216/p3VAE/`

1 INTRODUCTION

Deep learning techniques have gained widespread attention in airborne hyperspectral image processing, including land cover classification, on which we focus in this paper. The pixel information of hyperspectral images provided in ground-level reflectance, at very high spatial resolution and very high spectral resolution on a wide spectral range, is indeed highly informative of the chemical composition of the land surface (*e.g.* asphalt, gravel, bare soil). Nonetheless, the principal challenge of pixel-wise hyperspectral image classification lies in the scarcity of training data (prohibitively expensive to acquire through field campaigns) with respect to the diversity of ground materials and to their large spectral intra-class variability. In order to reach high generalization performance in this weakly supervised regime, most works have focused on the integration of inductive biases into learning. Inductive biases commonly refer to the assumptions made over the model architecture (*e.g.* attention mechanisms (Sun et al., 2022)), the objective function (*e.g.* contrastive learning (Zhao et al., 2022)) or the learning algorithm (*e.g.* few-shot learning (Liu et al., 2018)).

In contrast, we aim in this paper to integrate deductive biases derived from physical prior knowledge. The motivation of our work is that the variations of the local irradiance, a non-observed physical quantity that depends on the 3D geometry of the image rather than on the land cover itself, can be related to specific spectral variations through a radiative transfer model. Our methodology aims to leverage this a priori physical knowledge to improve the robustness of a classification model to irradiance variations in the context of scarce labeled data. In contrast to Aragon-Calvo & Carvajal (2020) and Zérah et al. (2023) that respectively use the framework of the autoencoder or variational autoencoder (VAE) (Kingma & Welling, 2014) to inverse a fully physical forward model, our work combines standard neural network layers with non-trainable physics layers in the decoder of a semi-supervised VAE (Kingma et al., 2014). Besides, in contrast to (Takeishi & Kalousis, 2021) that complement an imperfect physical model with a neural network, we do the other way around by complementing a neural network with physics layers that analytically compute the spectral shift of the data according to the variations of a subset of the latent space.

The outline of the paper is as follows. In section 2, we derive a physics deductive bias for hyperspectral representation learning and introduce a methodology to integrate this bias into a classification model. In section 3, we present the simulated and real data on which we validate our methodology through numerical experiments, detailed in section 4. Finally, we conclude in section 5.

*Work done while at ONERA. Correspondance to romain.thoreau@cnes.fr

2 METHODOLOGY

2.1 PHYSICS DEDUCTIVE BIASES FOR HYPERSPECTRAL REPRESENTATION LEARNING

Hyperspectral generative factors[1] In the remote sensing community, hyperspectral images are usually preprocessed by atmospheric correction codes, prior to the land cover classification, in order to estimate the ground-level reflectance from the sensor-level radiance. The reflectance is the ratio of the reflected radiant flux on the incident radiant flux, and only depends on the matter chemical composition. However, most atmospheric correction codes (*e.g.* COCHISE (Miesch et al., 2005)) assume a flat ground, hence uniform irradiance over every pixels, resulting in different reflectance estimates for the same land cover with different local irradiance (that depends on the scene 3D geometry). Therefore, we suggest in this paper to model the generative factors of hyperspectral data as follows: 1) the pixel land cover (assuming that one pixel contains only one material), 2) the local irradiance, and 3) *intrinsic* factors that reflect slight variations in the material chemical composition (*e.g.* variations in water content of different trees from the same species) and larger variations in material composition due to the fact that the nomenclature groups different materials under the same class (*e.g.* different tree species gathered in a unique class).

A deductive bias derived from a radiative transfer model We aim to integrate into the learning process a priori information defined by the ratio of the estimated reflectance $\boldsymbol{x}$ by an atmospheric correction code over the true reflectance[2] $\boldsymbol{x}^*$. We approximate this quantity (see the appendix for further details), that expresses how the signal varies at a wavelength λ given variations of the direct irradiance, denoted as I_{dir}, and the diffuse irradiance (scattered by the atmosphere), denoted as I_{dif}:

$$\frac{\boldsymbol{x}}{\boldsymbol{x}^*}(\lambda) \approx \frac{I^*_{dir}(\lambda) + I^*_{dif}(\lambda)}{I^{code}_{dir}(\lambda) + I^{code}_{dif}(\lambda)} \tag{1}$$

where the superscript $*$ is given for non-observed true physical quantities and the superscript $code$ is given for physical quantities assumed by the atmospheric correction code. The true direct irradiance I^*_{dir} differs from the assumed direct irradiance I^{code}_{dir} in shadows and in slopes. We model shadows with a quantity $\delta^*_{dir} \in [0, 1]$ corresponding to the portion of pixel directly lit by sun, and slopes with the local solar zenith angle, denoted as Θ^*, in contrast to the solar zenith angle assumed by the atmospheric correction code Θ^{code}. The true diffuse irradiance I^*_{dif} mainly differs from the assumed diffuse irradiance I^{code}_{dif} for pixels from where a small portion of the sky is visible. This quantity is commonly named the sky viewing angle factor that we denote as $\Omega^* \in [0, 1]$. Overall, we model the true direct and diffuse irradiance as follows:

$$I^*_{dir}(\lambda) = \delta^*_{dir} \cdot \frac{cos\ \Theta^*}{cos\ \Theta^{code}} \cdot I^{code}_{dir}(\lambda) \qquad I^*_{dif}(\lambda) = \Omega^* \cdot p^*_\Theta \cdot I^{code}_{dif}(\lambda) \tag{2}$$

where p^*_Θ is a correction coefficient that accounts for the anisotropy of the diffuse irradiance. Fig. 1 illustrates the radiative terms used in our model and Fig. 2 illustrates the ratio we derived for varying irradiance conditions, showing how non-linear the spectral variations are with respect to direct and diffuse irradiance variations.

2.2 INTEGRATION OF DEDUCTIVE BIASES INTO A SEMI-SUPERVISED VAE

In this section, we introduce a formalism entitled p^{3}VAE that integrates the deductive bias described in section 2.1 into a semi-supervised VAE (Kingma et al., 2014).

Latent variables modeling We split the latent space into three subsets $\mathcal{Z}_A$, $\mathcal{Z}_P$ and $\mathcal{Y}$. First, the continuous latent variable $\boldsymbol{z}_A \in \mathcal{Z}_A$ is meant to encode the *intrinsic* factors of variation. Second, the continuous latent variable $z_P \in \mathcal{Z}_P$ is meant to encode the irradiance conditions, precisely the physical quantity $\delta^*_{dir} \cdot cos\ \Theta^*$ that weights the direct irradiance. Third, the discrete latent variable $y \in \mathcal{Y}$ is meant to encode the land cover class. Additional details about the choice of probability distributions for priors and posteriors are given in the appendix.

[1] The factors that individually account for a source of variability in the data (Desjardins et al., 2012).

[2] The reflectance that would be computed by an atmospheric correction code with the exact knowledge of the local irradiance conditions.

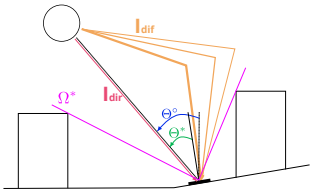
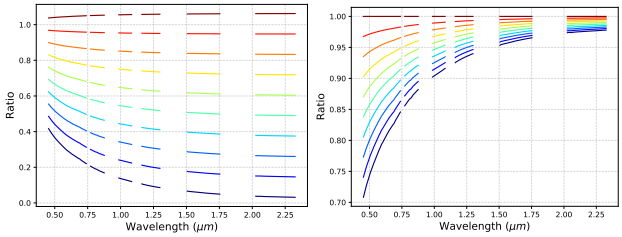


Figure 1: Illustration of irradiance terms.



$I_{dif} = I^o_{dif}, \delta_{dir} \in [0, 1]$ $\quad$ $I_{dir} = I^o_{dir}, \Omega \in [0.3, 1]$

Figure 2: Ratio of the estimated reflectance over the true reflectance for varying irradiance conditions.

Integrating physics into the decoder The decoder is the combination of standard neural network layers $f_A^{\boldsymbol{\theta}}$ with a non-trainable physical part f_P. First, $f_A^{\boldsymbol{\theta}}$ estimates the true reflectance $\boldsymbol{x}^* = f_A^{\boldsymbol{\theta}}(y, z_A)$ given the land cover and the latent representation of the *intrinsic* intra-class variability. Then, f_P, based on eq. 1, computes the spectral shift induced by the variation of the local irradiance, encoded by z_P:

$$f_P(z_P, \boldsymbol{x}^*) = \frac{z_P \frac{I^{code}_{dir}}{cos\ \Theta^{code}} + g(z_P) I^{code}_{dif}}{I^{code}_{dir} + I^{code}_{dif}} \boldsymbol{x}^* \tag{3}$$

where g is an affine function that empirically approximates the local diffuse irradiance (weight and bias of g are hyperparameters). All in all, the output of the decoder is $f_P \circ f_A(y, z_A, z_P)$. Additional details on the likelihood are given in the appendix.

Optimization We optimize the objective function derived by Kingma et al. (2014), composed of a supervised part and an unsupervised part (when y is not observed). We introduce a slight but fundamental variation to the algorithm: we do not back-propagate the gradients of the unsupervised part of the objective with respect to the parameters of $f_A^{\boldsymbol{\theta}}$. As a matter of fact, the high flexibility of $f_A^{\boldsymbol{\theta}}$ could lead to very low reconstruction errors on the training data with inconsistent values of $(\boldsymbol{z}_A, z_P, y)$, *i.e.* an entanglement of the local irradiance, of the *intrinsic* factors of variations and of the land cover.

Inference At inference, Kingma et al. (2014) use the approximate predictive distribution $q_\phi(y|\boldsymbol{x})$ to make predictions. However, although the true predictive distribution $p_{\boldsymbol{\theta}}(y|\boldsymbol{x})$ is intractable, we perform Monte Carlo sampling to estimate $\arg\max_y p_{\boldsymbol{\theta}}(y|\boldsymbol{x})$. More precisely, we use the two recognition networks of the encoder to perform importance sampling:

$$p_{\boldsymbol{\theta}}(y|\boldsymbol{x}) \propto \mathbb{E}_{y^* \sim q_{\boldsymbol{\phi}}(y|\boldsymbol{x})} \mathbb{E}_{\boldsymbol{z}_A, z_P \sim q_{\boldsymbol{\phi}}(\boldsymbol{z}_A, z_P|\boldsymbol{x}, y^*)} \frac{p(\boldsymbol{z}_A, z_P)}{q_\phi(\boldsymbol{z}_A, z_P|\boldsymbol{x})} p_\theta(\boldsymbol{x}|y, \boldsymbol{z}_A, z_P) \tag{4}$$

The major consequence of this inference scheme is to explicitly use the physical part f_P.

3 DATA

Simulated data We simulated an airborne hyperspectral image with the radiative transfer software DART (Gastellu-Etchegorry et al., 2012) composed of five land cover classes (#1 vegetation, #2 sheet metal, #3 sandy loam, #4 tile and #5 asphalt). Some classes group different materials (*i.e.* subclasses) for more realistic *intrinsic* intra-class variability. The training data set only partially represents the diversity of irradiance conditions, especially for class #3 whose labeled pixels are uniformly exposed to the sun. On the contrary, the unlabeled data set and the test set include data points with a greater variety of irradiance conditions. Additional details are given in the appendix.

Real data We used subsets of the airborne hyperspectral image acquired during the CAMCATT-AI4GEO experiment (Roupioz et al., 2023), publicly available at `https://camcatt.sedoo.fr/`. We labeled approximately 20,000 pixels through a field campaign, divided into eight land cover classes (#1 Tile, #2 Asphalt, #3 Vegetation, #4 Painted sheet metal, #5 Water, #6 Gravels, #7 Metal and #8 Fiber cement). Additional details are given in the appendix.

4 EXPERIMENTS

Experimental plan We compared our hybrid model, p^{3}VAE, to a semi-supervised VAE (Kingma et al., 2014) and to a common spectral CNN architecture optimized in a supervised way. We tested two training options for the CNN: training #1 over the labeled set only and training #2 over the labeled and unlabeled sets. Models are optimized with the Adam algorithm (Kingma & Ba, 2014) and hyperparameters, including penalty coefficients of the weight decay, weights of the KL divergence term and parameters of g, are selected on a validation set. 10 experiments with different random seeds were done.

Table 1: Mean F1 score over 10 runs on the simulated (left) and real (right) data sets .

Inference model		Classes: Veg.	S.M.	S.L.	Tile	Asp.	Avg	Classes: Tile	Asp.	Veg.	S.M.	Wat.	Gra.	Met.	F.C.	Avg
CNN	training #1	0.90	0.81	0.77	0.79	0.75	0.80	0.92	0.41	0.91	0.92	0.83	0.89	0.87	0.56	0.79
	training #2	0.92	0.79	*0.90*	0.87	0.86	0.86	*0.96*	*0.53*	*0.98*	*1.00*	*0.99*	0.98	*0.99*	*0.87*	*0.91*
VAE	$q_\phi(y\|\boldsymbol{x})$	0.93	0.80	0.87	0.86	0.74	0.84	0.94	0.27	0.88	0.96	0.83	0.92	**0.96**	0.66	0.80
	$p_\theta(y\|\boldsymbol{x})$	0.94	0.88	**0.90**	**0.92**	0.85	0.90	0.82	0.40	0.86	0.73	0.84	0.60	0.79	0.63	0.71
p^{3}VAE	$q_\phi(y\|\boldsymbol{x})$	0.92	0.82	0.88	0.87	0.81	0.86	**0.95**	**0.51**	**0.98**	0.99	0.98	**0.99**	**0.96**	0.81	**0.90**
	$p_\theta(y\|\boldsymbol{x})$	**0.96**	**0.97**	**0.90**	0.90	**0.93**	**0.93***	0.93	0.49	0.97	0.91	0.99	0.79	0.90	0.78	0.84

* *A statistical hypothesis test shows that the difference of average F1 score between the vanilla VAE and p^{3}VAE is significant (0.3% p-value).*

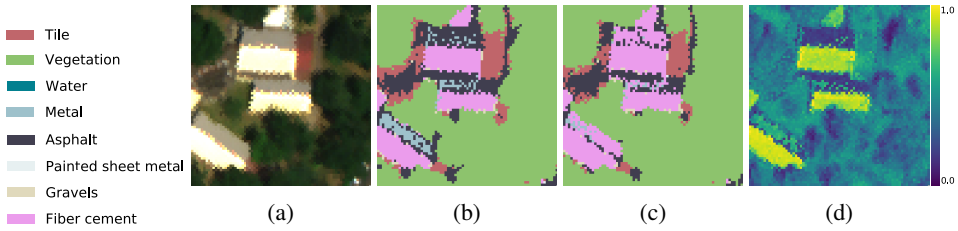


Figure 3: (a) RGB composition of a real test hyperspectral scene composed of roofs in fiber cement (b) VAE land cover predictions (c) p^{3}VAE land cover predictions (d) p^{3}VAE z_P predictions.

Classification accuracy Tab. 1 reports the mean F1 score over 10 runs on both data sets. On simulated data, p^{3}VAE globally outperformed the VAE and the CNN by large margins (in regards to the small standard deviations on the order of 0.01 for every models) when the approximation of $\arg\max_y p_{\boldsymbol{\theta}}(y|\boldsymbol{x})$ is used, even when the CNN is trained with more labeled data (training #2). In particular, a very large increase was reached compared to the VAE and the CNN, as well as to p^{3}VAE with the use of $q_{\boldsymbol{\phi}}(y|\boldsymbol{x})$, for class #3 that includes test spectra under very different irradiance conditions than training spectra. We argue that this is due to the explicit use of the physical part that naturally extrapolates during inference to unseen irradiance conditions during training. On real data, p^{3}VAE reached higher accuracy than the VAE and the CNN, though the estimation of $\arg\max_y p_{\boldsymbol{\theta}}(y|\boldsymbol{x})$ significantly led to less accurate results than the direct use of $q_{\boldsymbol{\phi}}(y|\boldsymbol{x})$. We believe that this difference of performance can be explained by the larger *intrinsic* intra-class variability in the real data set. Qualitatively, Fig. 6 (see additional land cover maps in the appendix) shows that the latent variable z_P can indeed be interpreted as the factor $\delta_{dir}^{\cdot} cos\,\Theta$ that weights the direct irradiance (that is low in shadows and high on roofs inclined towards the sun), and that p^{3}VAE improved the classification of ambiguous fiber cement pixels due to low irradiance.

5 CONCLUSION

In this paper, we introduced a methodology to integrate physics deductive biases into a variational autoencoder for hyperspectral image classification. Numerical experiments demonstrated the higher extrapolation capacity of our method compared to conventional machine learning models in a weakly supervised regime for pixels under out-of-distribution irradiance conditions. In future work, we would like to extend our methodology to handle land cover classes without training labels.

REFERENCES

Miguel A Aragon-Calvo and JC Carvajal. Self-supervised learning with physics-aware neural networks–i. galaxy model fitting. *Monthly Notices of the Royal Astronomical Society*, 498(3): 3713–3719, 2020.

Guillaume Desjardins, Aaron Courville, and Yoshua Bengio. Disentangling factors of variation via generative entangling, 2012.

Jean-Philippe Gastellu-Etchegorry, Eloi Grau, and Nicolas Lauret. Dart: A 3d model for remote sensing images and radiative budget of earth surfaces. *Modeling and simulation in engineering*, (2), 2012.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *stat*, 1050:1, 2014.

Durk P Kingma, Shakir Mohamed, Danilo Jimenez Rezende, and Max Welling. Semi-supervised learning with deep generative models. *Advances in neural information processing systems*, 27, 2014.

Bing Liu, Xuchu Yu, Anzhu Yu, Pengqiang Zhang, Gang Wan, and Ruirui Wang. Deep few-shot learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(4):2290–2304, 2018.

C. Miesch, L. Poutier, V. Achard, X. Briottet, X. Lenot, and Y. Boucher. Direct and inverse radiative transfer solutions for visible and near-infrared hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 43(7):1552–1562, 2005. doi: 10.1109/TGRS.2005.847793.

L. Roupioz, X. Briottet, K. Adeline, A. Al Bitar, D. Barbon-Dubosc, R. Barda-Chatain, P. Barillot, S. Bridier, E. Carroll, C. Cassante, A. Cerbelaud, P. Déliot, P. Doublet, P.E. Dupouy, S. Gadal, S. Guernouti, A. De Guilhem De Lataillade, A. Lemonsu, R. Llorens, R. Luhahe, A. Michel, A. Moussous, M. Musy, F. Nerry, L. Poutier, A. Rodler, N. Riviere, T. Riviere, J.L. Roujean, A. Roy, A. Schilling, D. Skokovic, and J. Sobrino. Multi-source datasets acquired over toulouse (france) in 2021 for urban microclimate studies during the camcatt/ai4geo field campaign. *Data in Brief*, 48:109109, 2023. ISSN 2352-3409. doi: https://doi.org/10.1016/j.dib.2023.109109. URL https://www.sciencedirect.com/science/article/pii/S2352340923002287.

Le Sun, Guangrui Zhao, Yuhui Zheng, and Zebin Wu. Spectral–spatial feature tokenization transformer for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2022.

Naoya Takeishi and Alexandros Kalousis. Physics-integrated variational autoencoders for robust and interpretable generative modeling. *Advances in Neural Information Processing Systems*, 34: 14809–14821, 2021.

Yoël Zérah, Silvia Valero, and Jordi Inglada. Physics-driven probabilistic deep learning for the inversion of physical models with application to phenological parameter retrieval from satellite times series. *IEEE Transactions on Geoscience and Remote Sensing*, 2023.

Lin Zhao, Wenqiang Luo, Qiming Liao, Siyuan Chen, and Jianhui Wu. Hyperspectral image classification with contrastive self-supervised learning under limited labeled samples. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2022.

A APPENDIX

A.1 APPROXIMATION OF THE RATIO x/x^*

We assume, as most atmospheric codes do, that land surfaces are lambertian, *i.e.* that they reflect radiation isotropically. The reflectance of a pixel at wavelength λ is commonly defined in the literature as follows:

$$\boldsymbol{x}(\lambda) = \frac{\pi R_{dir}(\lambda)}{I_{tot}(\lambda)\tau_{dir}(\lambda)} \quad \text{with} \begin{cases} R_{dir}(\lambda) & = R_{tot}(\lambda) - R_{env}(\lambda) - R_{atm}(\lambda) \\ I_{tot}(\lambda) & = I_{dir}(\lambda) + I_{dif}(\lambda) + I_{coup}(\lambda) + I_{refl}(\lambda) \end{cases} \tag{5}$$

where:

- R_{tot} is the radiance measured by the sensor,
- R_{atm} is the portion of R_{tot} that is scattered by the atmosphere without any interactions with the ground,
- R_{env} is the portion of R_{tot} that comes from the neighbourhood of the pixel,
- R_{dir} is the portion of R_{tot} that comes from the pixel,
- I_{dir} is the irradiance directly coming from the sun,
- I_{dif} is the irradiance scattered by the atmosphere,
- I_{coup} is the irradiance coming from the coupling between the ground and the atmosphere,
- I_{refl} is the irradiance coming from neighbouring 3D structures.

In section 2.1, we neglect the contributions of I_{coup} and I_{refl} to derive eq. 1.

A.2 LATENT VARIABLES MODELING

An intuitive interpretation of $\boldsymbol{z}_A$ is the abundance of different subclasses within a land cover class. Therefore, we model $\boldsymbol{z}_A$ prior and posterior with Dirichlet distributions. We model z_P prior and posterior with Beta distributions for their flexibility and domain bounded between 0 and 1. Specifically, we define the priors as follows:

$$p(z_P) := Beta(z_P|\alpha^o, \beta^o); \qquad p(\boldsymbol{z}_A) := Dir(\boldsymbol{z}_A|\gamma^o) \tag{6}$$

We empirically set α^o to 1 and $\beta^o = \frac{1-\cos\ \Theta^{code}}{\cos\ \Theta^{code}+\epsilon}\alpha^o$ where ϵ is a small constant to avoid division by zero. This prior distribution favors high values of z_P while $\mathbb{E}_{p(z_P)} z_P = \cos\ \Theta^{code}$, which is what we expect on a flat ground. We set $\gamma^o = [1 \dots 1]_{1\times|\boldsymbol{z}_A|}$, which is equivalent to having a uniform prior. We assume that y has a uniform prior and that y, $\boldsymbol{z}_A$ and z_P are a priori independent.

A.3 LIKELIHOOD OF P^{3}VAE

In contrast to a standard VAE, we do not model the likelihood as a multivariate normal. Denoting $f_P(f_A(y, \boldsymbol{z}_A), z_P)$ as $\boldsymbol{\mu}$, we instead define the likelihood as follows:

$$p_\theta(\boldsymbol{x}|y, z_P, \boldsymbol{z}_A) := \frac{1}{Z} \underbrace{\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \sigma^2 I)}_{\mathcal{A}(\boldsymbol{x})} \cdot \underbrace{exp\big(-\alpha\, arccos(\frac{\boldsymbol{x}^T\boldsymbol{\mu}}{\|\boldsymbol{x}\|\|\boldsymbol{\mu}\|})\big)}_{\mathcal{B}(\boldsymbol{x})} \tag{7}$$

where σ and α are hyperparameters and $Z = \int \mathcal{A}(\boldsymbol{x}) \cdot \mathcal{B}(\boldsymbol{x}) d\boldsymbol{x}$ is a finite constant such that the density integrates to one. In order to prove that Z is finite, we can first easily show that $A : \mathcal{X} \longrightarrow \mathbb{R}$ is square-integrable, as well as $B : \mathcal{X} \longrightarrow \mathbb{R}$:

$$\int |\mathcal{B}(\boldsymbol{x})|^2 dx = \int_{[0,1]^B} exp\big(-2\alpha\, arccos(\frac{\boldsymbol{x}^T\boldsymbol{\mu}}{\|\boldsymbol{x}\|\|\boldsymbol{\mu}\|})\big) d\boldsymbol{x} \leq \int_{[0,1]^B} d\boldsymbol{x} = 1 \tag{8}$$

Moreover, A and B being continous over $\mathcal{X}$, the Cauchy-Schwarz inequality implies that:

$$\left| \int \mathcal{A}(\boldsymbol{x}) \cdot \mathcal{B}(\boldsymbol{x}) dx \right| \leq \left(\int \mathcal{A}(\boldsymbol{x})^2 d\boldsymbol{x} \right)^{\frac{1}{2}} \left(\int \mathcal{B}(\boldsymbol{x})^2 dx \right)^{\frac{1}{2}} = C \in \mathbb{R} \tag{9}$$

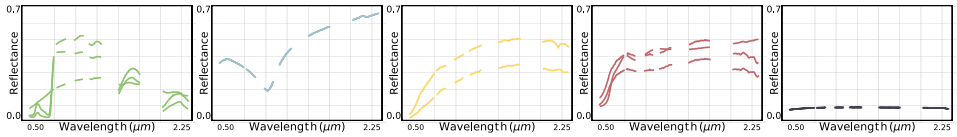
Thus, $Z \in \mathbb{R}$ and $p_{\boldsymbol{\theta}}(\boldsymbol{x}|y, z_P, \boldsymbol{z}_A)$ properly defines a probability density function.

A.4 SIMULATED DATA

300 spectral bands with a 6.5 nm resolution and a 0.5 m ground sampling distance were simulated, without the Earth-atmosphere coupling, using reference reflectance spectra for 5 land cover classes shown in Fig. 5. A false-color image of the training data with its ground truth is shown on Fig. 4. Approximately $10,000$ test spectra were analytically generated from the reference spectra with exhaustive irradiance conditions.

(a) (b) (c)

Figure 4: (a) False color composition of the simulated hyperspectral image and its (b) labeled training and (c) unlabeled training ground truths.



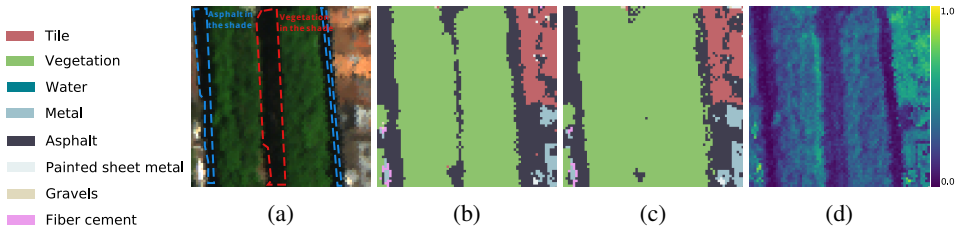
(a) Vegetation (b) Sheet Metal (c) Sandy Loam (d) Tile (e) Asphalt

Figure 5: Reference spectra used for the DART simulation

A.5 REAL DATA

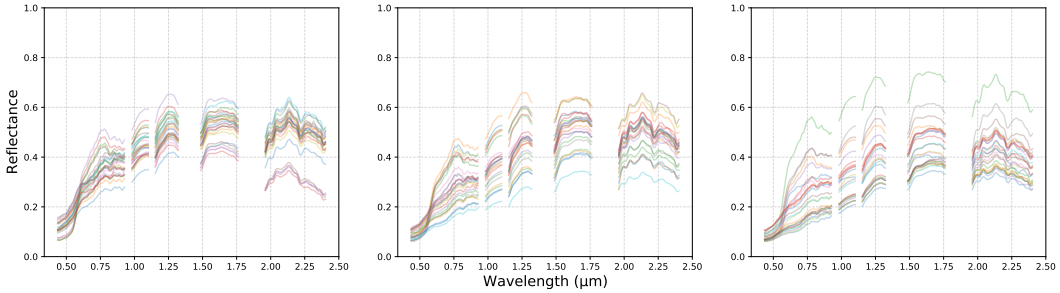
Real hyperspectral data was acquired with a AisaFENIX 1K camera, which covers the spectral range from 0.4 µm to 2.5 µm with a 1 m ground sampling distance. Data was converted in radiance at aircraft level through radiometric and geometric corrections (Roupioz et al., 2023). Then, the radiance image was converted to surface reflectance with the atmospheric correction algorithm COCHISE (Miesch et al., 2005). Random labeled spectra are shown in Fig. 7.

A.6 ADDITIONAL LAND COVER MAPS

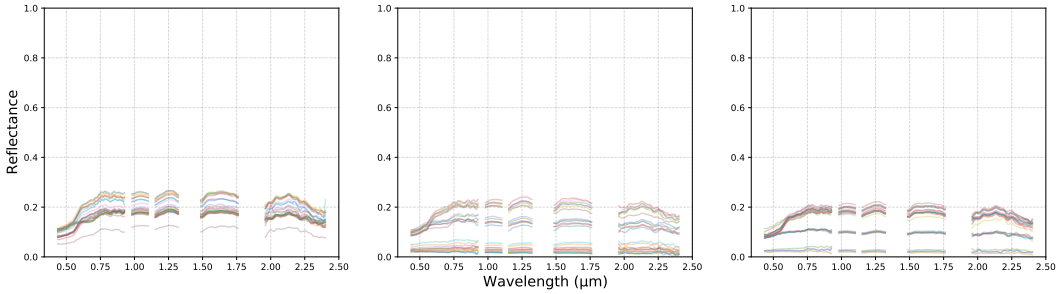


(a) (b) (c) (d)

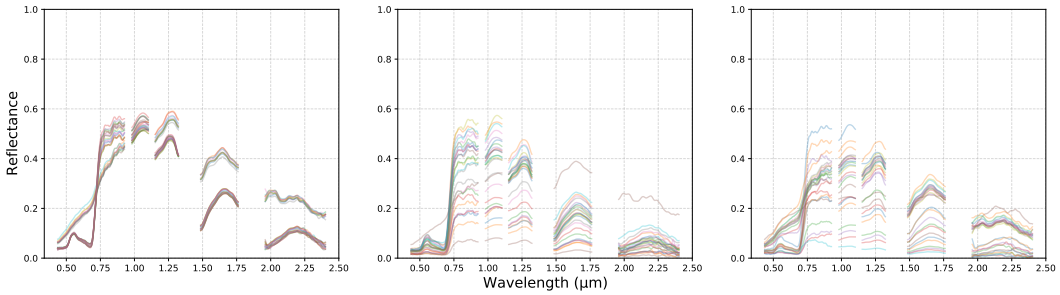
Figure 6: (a) RGB composition of a real test hyperspectral scene composed of ground vegetation in the shadows of trees (b) VAE land cover predictions (c) p^{3}VAE land cover predictions (d) p^{3}VAE z_P predictions.



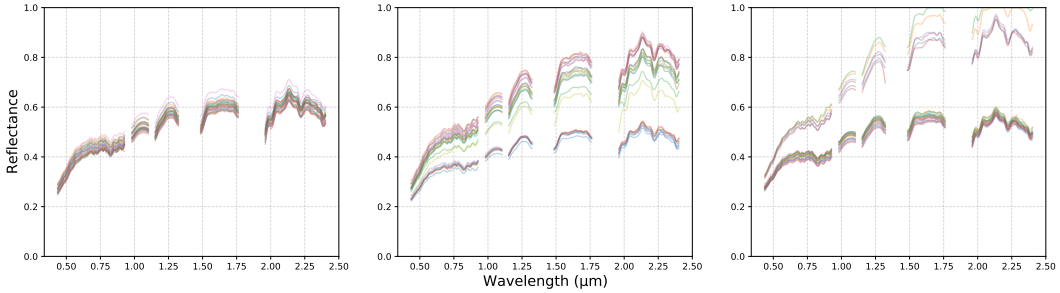
(a) Tile



(b) Asphalt

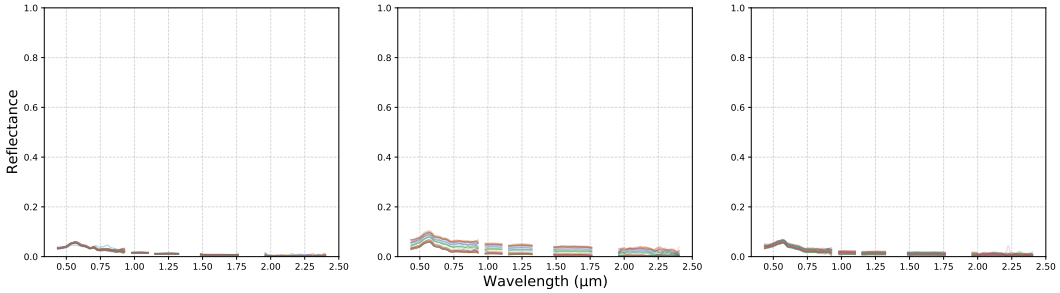


(c) Vegetation

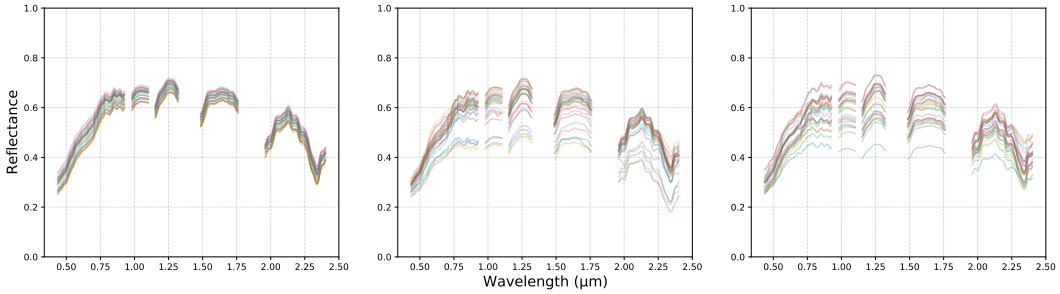


(d) Painted sheet metal

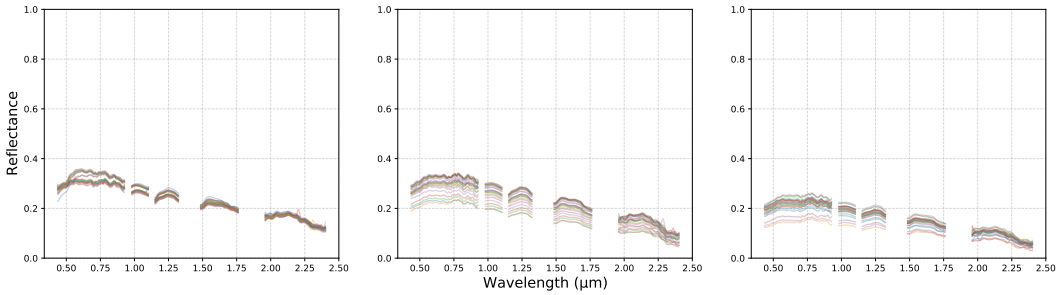
Figure 7: Random spectra from the (left) labeled, (middle) unlabeled and (right) test real data sets



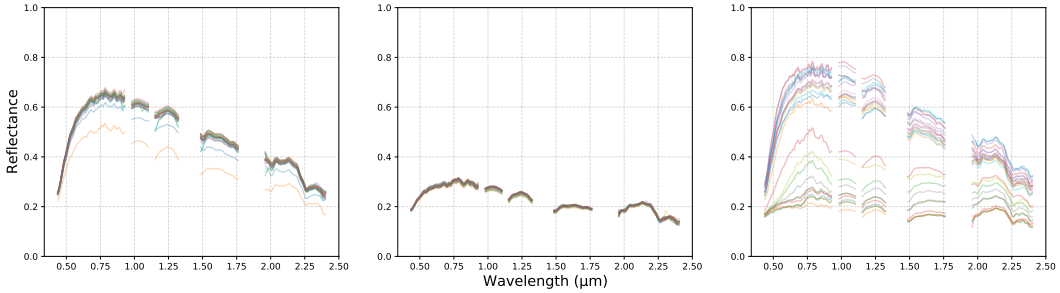
(e) Water



(f) Gravels



(g) Metal



(h) Fiber cement

Figure 7: Random spectra from the (left) labeled, (middle) unlabeled and (right) test real data sets