

# Physics-informed Variational Autoencoders for Improved Robustness to Environmental Factors of Variation

Romain Thoreau<sup>1</sup>, Laurent Risser<sup>2</sup>, Véronique Achard<sup>3</sup>, Béatrice Berthelot<sup>4</sup>, Xavier Briottet<sup>3</sup>

<sup>1</sup>CNES - Data Campus — <sup>2</sup>Institut de Mathématiques de Toulouse

<sup>3</sup>ONERA-DOTA — <sup>4</sup>Magellium

[romain.thoreau@cnes.fr](mailto:romain.thoreau@cnes.fr)

*April 29, 2025 - Télécom Paris, séminaire IMAGES*

- ▶ This work is under review in Springer Machine Learning:
  - R. Thoreau, L. Risser, V. Achard, B. Berthelot and X. Briottet, *Physics-informed Variational Autoencoders for Improved Robustness to Environmental Factors of Variation*, arXiv preprint [arxiv.org/abs/2210.10418](https://arxiv.org/abs/2210.10418)
- ▶ A short version of this work was accepted at ML4RS Workshop, ICLR 2024:
  - <https://ml-for-rs.github.io/iclr2024/>
- ▶ Code and data: <https://github.com/Romain3Ch216/p3VAE/>

## Overview

1. Motivation: mapping the land cover in large urban areas from hyperspectral images
2.  $p^3$ VAE framework through a toy example: the damped pendulum
3. Experimental results: extrapolation of pendulum time series & hyperspectral image classification
4. Conclusions & Perspectives

1. Motivation: mapping the land cover in large urban areas from hyperspectral images
2. p<sup>3</sup>VAE framework through a toy example: the damped pendulum
3. Experimental results: extrapolation of pendulum time series & hyperspectral image classification
4. Conclusions & Perspectives

## Impact of surface sealing on urban heat island effects

- ▶ Radiative properties of artificial impermeable surfaces increase urban heat island effects [IPCC, 2021]

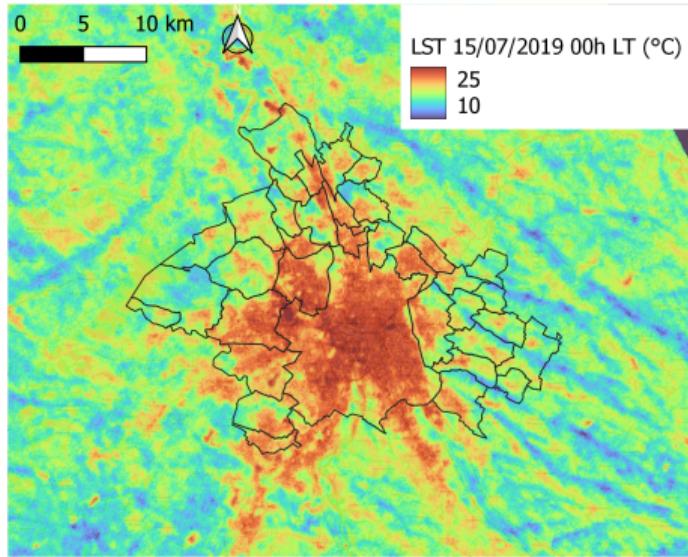


Figure: Land surface temperature measured in Toulouse from the Ecostress mission on 15 July 2019 at midnight. Source: THERMOCITY<sup>1</sup>

<sup>1</sup>Data source: <https://www.theia-land.fr/product/thermocity/>

## Impact of surface sealing on soil carbon sequestration

- ▶ Surface sealing **decrease the soil carbon storage potential**  
[Scalenghe and Marsan, 2009,  
Pereira et al., 2021,  
O'Riordan et al., 2021].
- ▶ ≈ 60,000 tonnes of estimated loss of carbon storage in Bordeaux due to surface sealing between 2012 and 2018.

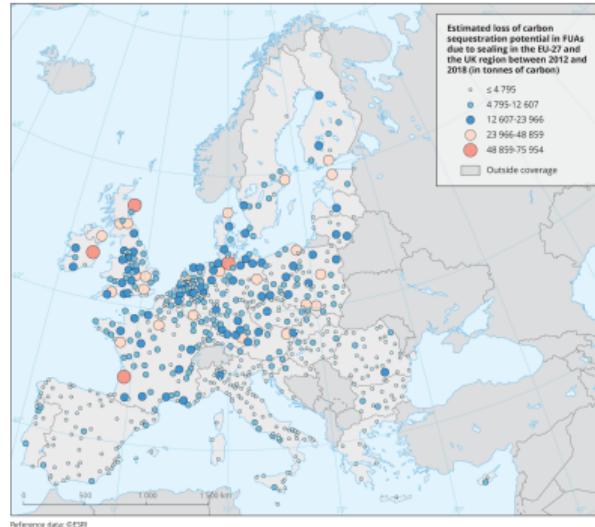


Figure: Estimated loss of potential carbon sequestration due to the estimated soil sealing increase in urban areas during 2012 and 2018. Source: European Environment Agency<sup>1</sup>

<sup>1</sup><https://www.eea.europa.eu/data-and-maps/figures/estimated-loss-of-carbon-sequestration>

# Mapping impermeable surfaces in large urban areas...

## Key issue in Earth observation

In order to mitigate the environmental impacts of surface sealing, **maps of impermeable surfaces are essential to guide and prioritize public policies.**



### Impermeable

- Asphalt
- Bitumen
- Tile
- Slate
- ...

### Permeable

- Gravel
- Bare soil
- Ground vegetation
- Tree
- ...

... using airborne hyperspectral images

■ Synthetic grass ■ Vegetation

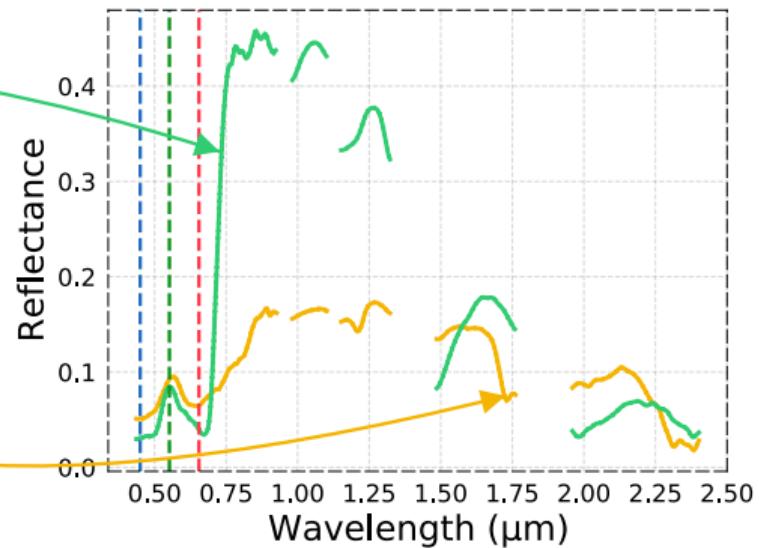
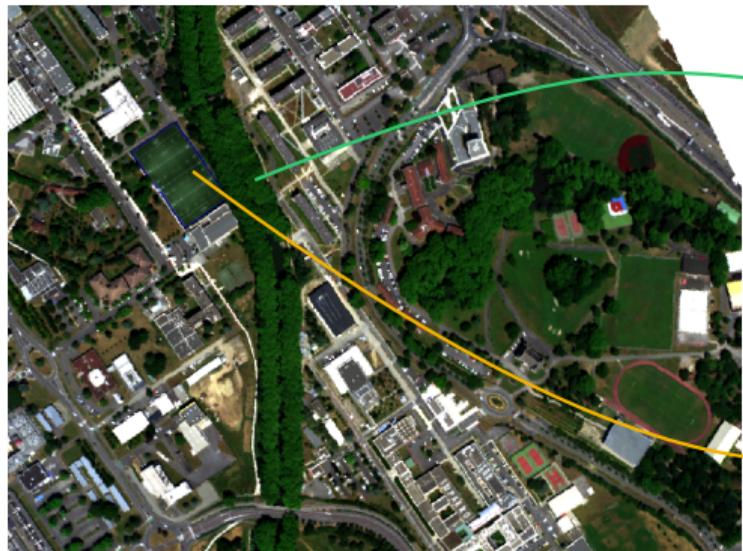


Figure: Hyperspectral airborne image (1m spatial resolution) of ISAE-SUPAERO / ONERA / CREPS.  
Source: ONERA

... using airborne hyperspectral images

■ Synthetic grass ■ Vegetation

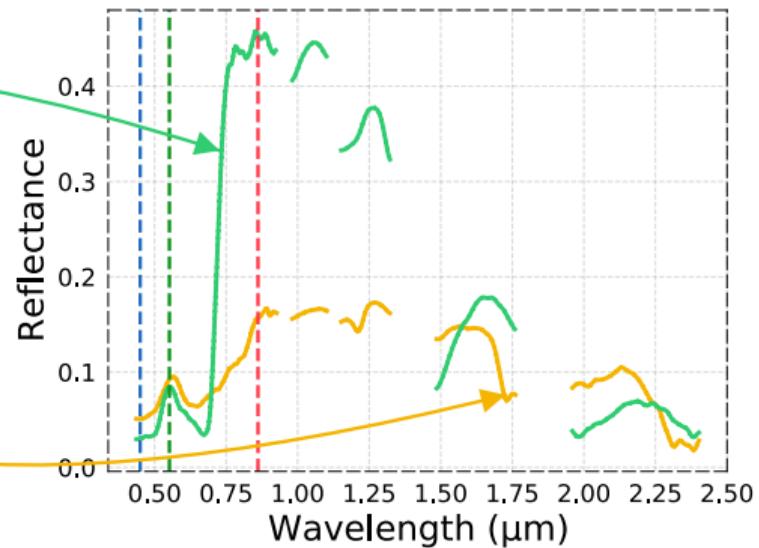
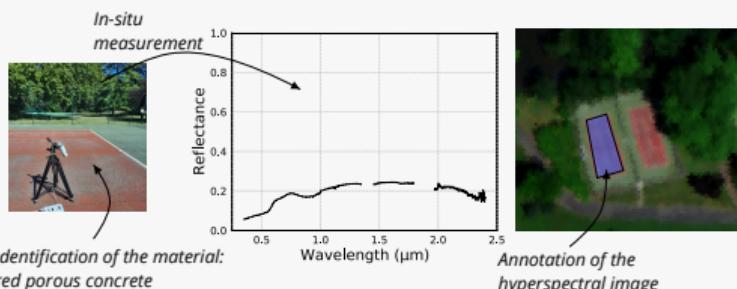


Figure: Hyperspectral airborne image (1m spatial resolution) of ISAE-SUPAERO / ONERA / CREPS.  
Source: ONERA

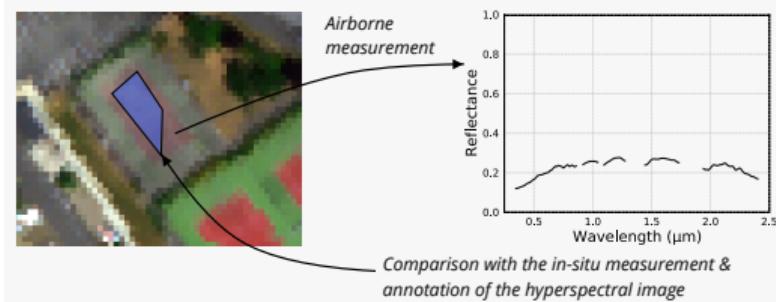
But land cover data is scarce!

The annotation of the land cover through field campaigns and photo-interpretation is *very expensive*.  
→ The proportion of labeled pixels in a hyperspectral image is usually *very small*.

## Field campaign



## Photo-interpretation

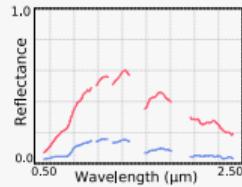
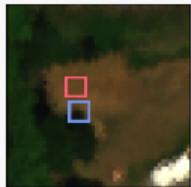


# Can we integrate prior knowledge about the spectral intra-class variability?

## Environmental intra-class variability

Caused by **different illumination conditions**

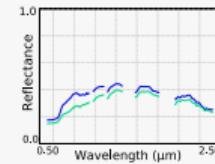
- ▶ Pixels in the shadows or in the sunlight,
- ▶ Pixels on a flat ground or on a slope.



## Intrinsic intra-class variability

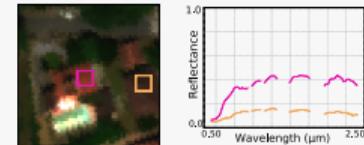
Caused by **slight variations of the chemical composition / non-lambertian effects**

- ▶ Different environmental exposures,
- ▶ Different agings (e.g., recent / old asphalt),
- ▶ For vegetation, different water / chlorophyll concentrations.



Caused by **large variations of the chemical composition**

- ▶ Different materials gathered in the same semantic class



1. Motivation: mapping the land cover in large urban areas from hyperspectral images
2.  $p^3$ VAE framework through a toy example: the damped pendulum
3. Experimental results: extrapolation of pendulum time series & hyperspectral image classification
4. Conclusions & Perspectives

## Toy example: the damped pendulum

Let's consider a training data set comprising pendulum time series  $\mathbf{x}^{(i)} := [\vartheta^{(i)}(t_1) \dots \vartheta^{(i)}(t_L)]^T$ , where  $\vartheta$  is the pendulum's angle.

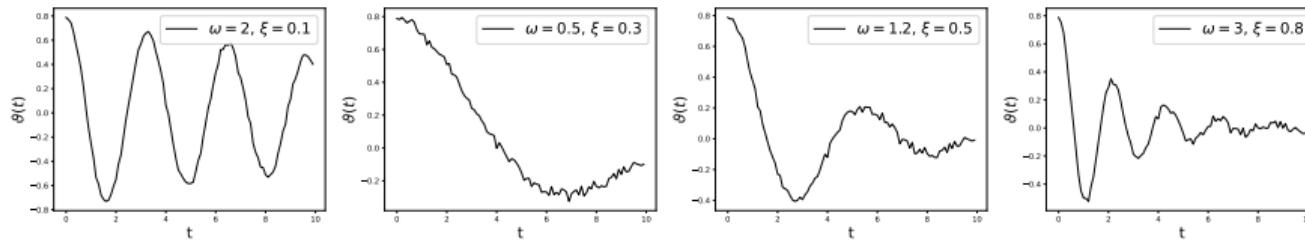


Figure: Samples of the training data set.

The pendulum trajectories are generated from *unknown* factors of variation:

- $\omega$ , the pendulum's angular frequency (*intrinsic* to the pendulum),
- $\xi$ , the fluid damping coefficient (depending on the *environmental* conditions).

### Objective of this work

Infer the generative factors of variations and the pendulum dynamics from noisy data points.

## Toy example: the damped pendulum

We assume that the pendulum dynamics, described by the following ordinary differential equation (ODE), are unknown:

$$\underbrace{\frac{d^2\vartheta}{dt^2}(t) + \xi \frac{d\vartheta}{dt}(t) + \omega^2 \sin \vartheta(t)}_{\text{assumed unknown}} = 0$$

## Toy example: the damped pendulum

We assume that the pendulum dynamics, described by the following ordinary differential equation (ODE), are unknown:

$$\underbrace{\frac{d^2\vartheta}{dt^2}(t) + \xi \frac{d\vartheta}{dt}(t) + \omega^2 \sin \vartheta(t)}_{\text{assumed unknown}} = 0$$

We only assume that physical prior knowledge  $f_E$  describes how the fluid damping coefficient  $\xi$  induces an exponential decrease of the pendulum's angle:

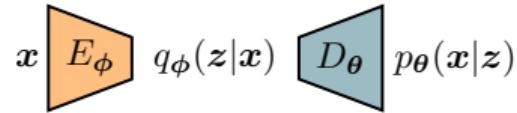
$$\begin{aligned} - f_E(\boldsymbol{x}, \xi) &= \boldsymbol{x} \exp(-\frac{1}{2}\xi \boldsymbol{t}), \boldsymbol{t} = [t_1 \dots t_L]^T \\ - \vartheta(t) &= f_E(\mathcal{F}_{\text{ODE}}[\ddot{\vartheta} + \underbrace{f_I(\vartheta; \omega)}_{\substack{\text{unknown, to be} \\ \text{learned by a NN}}}] = 0]; \xi \end{aligned}$$

where  $\mathcal{F}_{\text{ODE}}$  denotes an ODE solver with regard to  $\vartheta$ .

## Variational Autoencoders (VAE)

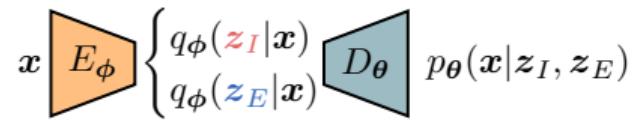
VAEs [Kingma and Welling, 2014] are latent generative models trained with stochastic gradient descent in order to maximize a lower-bound of the marginal likelihood  $\log p_{\theta}(\mathbf{x}^{(i)})$ :

- $\mathcal{L}(\boldsymbol{\theta}, \phi; \mathbf{x}^{(i)}) = \mathbb{E}_{q_{\phi}(z|\mathbf{x})} [\log p_{\theta}(\mathbf{x}^{(i)}|z)] - D_{KL}(q_{\phi}(z|\mathbf{x}^{(i)}) \| p(z)) \leq \log p_{\theta}(\mathbf{x}^{(i)})$
- The encoder  $E_{\phi}$  computes the variational approximation  $q_{\phi}(z|\mathbf{x})$  to the intractable posterior  $p_{\theta}(z|\mathbf{x})$ ,
- Statistical independance of latent variables  $z$  is encouraged by the isotropic Gaussian prior  $p(z)$ .



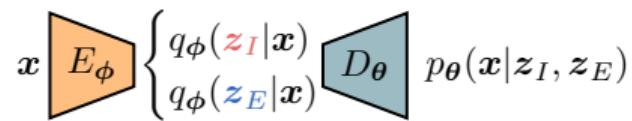
## Disentanglement in latent generative models

**Disentangled latent representations**  $z$  independently capture the true underlying factors, denoted as  $v$ , that explain the data [Carbonneau et al., 2022], e.g. a change of  $\omega$  (resp.  $\xi$ ) implies a change in  $z_I$  only (resp.  $z_E$ ).



## Disentanglement in latent generative models

**Disentangled latent representations**  $z$  independently capture the true underlying factors, denoted as  $v$ , that explain the data [Carboneau et al., 2022], e.g. a change of  $\omega$  (resp.  $\xi$ ) implies a change in  $z_I$  only (resp.  $z_E$ ).



Impossibility of unsupervised disentanglement learning, Theorem 1 in [Locatello et al., 2019]

Informal result: unsupervised disentanglement learning is fundamentally impossible for arbitrary generative models *without inductive biases both on the models and on the data.*

# $p^3$ VAE: a physics-informed VAE that integrates partial prior knowledge

$p^3$ VAE leverages four fundamental elements:

- ▶ Prior physical knowledge, encoded into  $f_E$ , of how data  $x$  relates to [environmental](#) factors of variation.

## p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge, encoded into  $f_E$ , of how data  $\mathbf{x}$  relates to environmental factors of variation. Here,  $f_E$  describes how the fluid damping coefficient  $\xi$  induces an exponential decrease of the pendulum's angle:

$$\begin{aligned} &— f_E(\mathbf{x}, \xi) = \mathbf{x} \exp\left(-\frac{1}{2}\xi \mathbf{t}\right), \mathbf{t} = [t_1 \dots t_L]^T \\ &— \vartheta(t) = f_E(\mathcal{F}_{\text{ODE}}[\ddot{\vartheta} + \underbrace{f_I(\vartheta; \omega)}_{\substack{\text{unknown, to be} \\ \text{learned by a NN}}}] = 0); \xi \end{aligned}$$

where  $\mathcal{F}_{\text{ODE}}$  denotes an ODE solver with regard to  $\vartheta$ .

## p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge, encoded into  $f_E$ , of how data  $\mathbf{x}$  relates to environmental factors of variation. Here,  $f_E$  describes how the fluid damping coefficient  $\xi$  induces an exponential decrease of the pendulum's angle:

$$\begin{aligned} &— f_E(\mathbf{x}, \xi) = \mathbf{x} \exp\left(-\frac{1}{2}\xi \mathbf{t}\right), \mathbf{t} = [t_1 \dots t_L]^T \\ &— \vartheta(t) = f_E(\mathcal{F}_{\text{ODE}}[\ddot{\vartheta} + \underbrace{f_I(\vartheta; \omega)}_{\substack{\text{unknown, to be} \\ \text{learned by a NN}}}] = 0); \xi \end{aligned}$$

where  $\mathcal{F}_{\text{ODE}}$  denotes an ODE solver with regard to  $\vartheta$ .  $f_E$  is integrated into the decoder of the VAE:

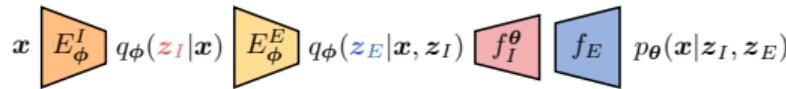


where  $p_\theta(\mathbf{x}|\mathbf{z}_I, \mathbf{z}_E) = \mathcal{N}(\mathbf{x}|f_E(\mathcal{F}_{\text{ODE}}[\ddot{\vartheta} + f_I^\theta(\vartheta; \mathbf{z}_I) = 0]; \mathbf{z}_E), \sigma^2 \mathbf{I})$

## p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

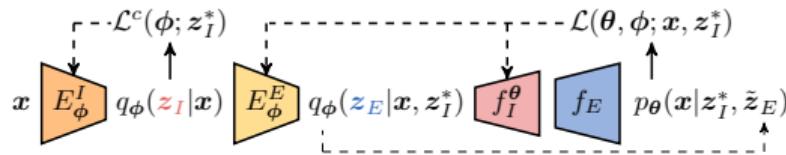
- ▶ Prior physical knowledge of how data  $\mathbf{x}$  relates to **environmental** factors of variation.
- ▶ Partial supervision of the **intrinsic** factors of variations.



# p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge of how data  $\mathbf{x}$  relates to environmental factors of variation.
- ▶ Partial supervision of the intrinsic factors of variations.



- Supervised part The supervised loss function  $\mathcal{L}^{sup}$  is a standard lower-bound of the log-likelihood plus a cross-entropy loss  $\mathcal{L}^c$  [Kingma et al., 2014]:

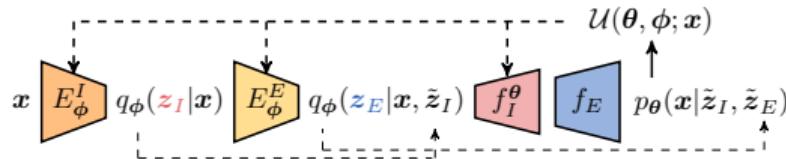
$$\mathcal{L}^{sup} = \mathcal{L}(\boldsymbol{\theta}, \phi; \mathbf{x}^{(i)}, \mathbf{z}_I^{(i)}) + \lambda \mathcal{L}^c(\phi; \mathbf{z}_I^{(i)})$$

$$\begin{aligned} \text{where } -\mathcal{L}(\boldsymbol{\theta}, \phi; \mathbf{x}^{(i)}, \mathbf{z}_I^{(i)}) &= \mathbb{E}_{q_\phi(\mathbf{z}_E|\mathbf{x}^{(i)}, \mathbf{z}_I^{(i)})} [\log p_\theta(\mathbf{x}^{(i)}|\mathbf{z}_I^{(i)}, \mathbf{z}_E)] \\ &\quad - D_{KL}(q_\phi(\mathbf{z}_E|\mathbf{x}^{(i)}, \mathbf{z}_I^{(i)}) \| p_\theta(\mathbf{z}_E)) \leq \log p_\theta(\mathbf{x}^{(i)}, \mathbf{z}_I^{(i)}) \end{aligned}$$

# p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge of how data  $x$  relates to **environmental** factors of variation.
- ▶ Partial supervision of the **intrinsic** factors of variations.



- Supervised part

$$\mathcal{L}^{sup} = \mathcal{L}(\theta, \phi; \mathbf{x}^{(i)}, \mathbf{z}_I^{(i)}) + \lambda \mathcal{L}^c(\phi; \mathbf{z}_I^{(i)})$$

- Unsupervised part

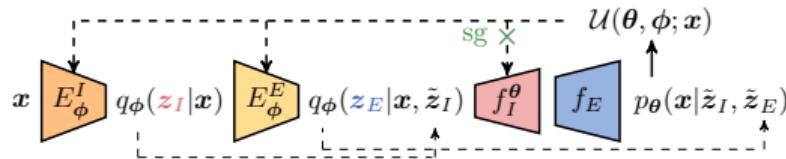
$$-\mathcal{U}(\theta, \phi; \mathbf{x}^{(i)}) = -\mathbb{E}_{q_\phi(\mathbf{z}_I|\mathbf{x}^{(i)})}[\mathcal{L}(\theta, \phi; \mathbf{x}^{(i)}, \mathbf{z}_I)] + H[q_\phi(\mathbf{z}_I|\mathbf{x})] \leq \log p_\theta(\mathbf{x}^{(i)})$$

## p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge of how data  $\mathbf{x}$  relates to environmental factors of variation.
- ▶ Partial supervision of the intrinsic factors of variations
- ▶ Balance between  $f_I^\theta$  and  $f_E$  through a stop-gradient operator applied on the unsupervised loss function:

$$-\mathcal{U}(\boldsymbol{\theta}, \phi; \mathbf{x}^{(i)}) = -\text{sg}[\mathbb{E}_{q_\phi(\mathbf{z}_I|\mathbf{x}^{(i)})}[\mathcal{L}(\boldsymbol{\theta}, \phi; \mathbf{x}^{(i)}, \mathbf{z}_I)]] + H[q_\phi(\mathbf{z}_I|\mathbf{x})] \leq \log p_{\boldsymbol{\theta}}(\mathbf{x}^{(i)})$$



## p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge of how data  $\mathbf{x}$  relates to **environmental** factors of variation.
- ▶ Partial supervision of the **intrinsic** factors of variations
- ▶ Balance between  $f_I^\theta$  and  $f_E$  through a **stop-gradient operator** applied on the unsupervised loss function, yielding a two-step algorithm in an expectation-maximization fashion:
  - Supervised part = step 1:

$$\max_{\boldsymbol{\theta}, \boldsymbol{\phi}} \mathcal{L}^{sup}(\boldsymbol{\theta}, \boldsymbol{\phi}; \mathbf{x}, \mathbf{z}_I)$$

- Unsupervised part = step 2:

$$\max_{\boldsymbol{\phi}} \mathcal{U}(\boldsymbol{\phi}; \boldsymbol{\theta}, \mathbf{x})$$

The second step only decreases the tightness of the bound, *i.e.* the gap between the lower bound  $\mathcal{U}(\boldsymbol{\theta}, \boldsymbol{\phi}; \mathbf{x})$  and the marginal log-likelihood  $\log p_\theta(\mathbf{x})$ .

## p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge of how data  $\mathbf{x}$  relates to **environmental** factors of variation.
- ▶ Partial supervision of the **intrinsic** factors of variations
- ▶ Balance between  $f_I^\theta$  and  $f_E$  through a *stop-gradient operator*
- ▶ Estimation of  $\arg \max_{\mathbf{z}_I} p_\theta(\mathbf{z}_I | \mathbf{x})$  at inference through importance sampling, when  $\mathbf{z}_I$  is discrete

# p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge of how data  $\mathbf{x}$  relates to **environmental** factors of variation.
- ▶ Partial supervision of the **intrinsic** factors of variations
- ▶ Balance between  $f_I^\theta$  and  $f_E$  through a ***stop-gradient operator***
- ▶ Estimation of  $\arg \max_{\mathbf{z}_I} p_\theta(\mathbf{z}_I | \mathbf{x})$  at inference through importance sampling, when  $\mathbf{z}_I$  is discrete :

$$p_\theta(\mathbf{z}_I | \mathbf{x}) \propto \mathbb{E}_{\tilde{\mathbf{z}}_I \sim q_\phi(\mathbf{z}_I | \mathbf{x})} \mathbb{E}_{q_\phi(\mathbf{z}_E | \mathbf{x}, \tilde{\mathbf{z}}_I)} \frac{p_\theta(\mathbf{z}_E)}{q_\phi(\mathbf{z}_E | \mathbf{x})} p_\theta(\mathbf{x} | \mathbf{z}_I, \mathbf{z}_E)$$

→ Inference with an MC estimate of  $p_\theta(\mathbf{z}_I | \mathbf{x})$  rather than the neural approximation  $q_\phi(\mathbf{z}_I | \mathbf{x})$  makes an explicit use of the prior knowledge  $f_E$ !

## p<sup>3</sup>VAE: a physics-informed VAE that integrates partial prior knowledge

p<sup>3</sup>VAE leverages four fundamental elements:

- ▶ Prior physical knowledge of how data  $\mathbf{x}$  relates to environmental factors of variation.
- ▶ Partial supervision of the intrinsic factors of variations
- ▶ Balance between  $f_I^\theta$  and  $f_E$  through a stop-gradient operator
- ▶ Estimation of  $\arg \max_{\mathbf{z}_I} p_\theta(\mathbf{z}_I | \mathbf{x})$  at inference through importance sampling, when  $\mathbf{z}_I$  is discrete

1. Motivation: mapping the land cover in large urban areas from hyperspectral images
2.  $p^3$ VAE framework through a toy example: the damped pendulum
3. Experimental results: extrapolation of pendulum time series & hyperspectral image classification
4. Conclusions & Perspectives

## Extrapolation of pendulum time series: qualitative comparison

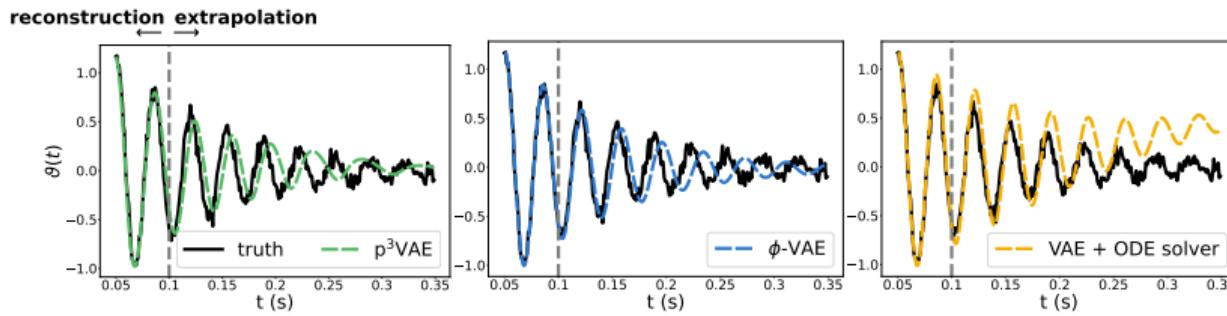


Figure: Predicted dynamics of  $p^3$ VAE (ours),  $\phi$ -VAE [Takeishi and Kalousis, 2021] and VAE + ODE solver [Yildiz et al., 2019, Toth et al., 2020] for the damped pendulum vs. ground truth trajectory.

## Extrapolation of pendulum time series: qualitative comparison

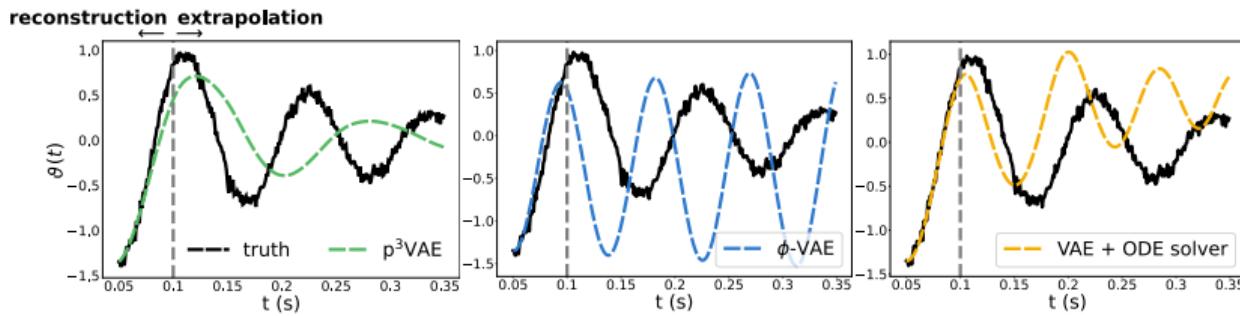


Figure: Predicted dynamics of  $p^3$ VAE (ours),  $\phi$ -VAE [Takeishi and Kalousis, 2021] and VAE + ODE solver [Yildiz et al., 2019, Toth et al., 2020] for the damped pendulum vs. ground truth trajectory.

## Extrapolation of pendulum time series: qualitative comparison

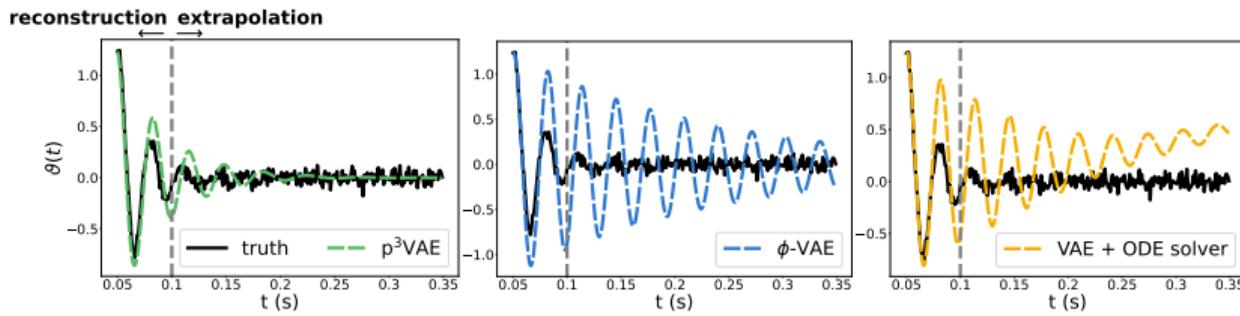


Figure: Predicted dynamics of  $p^3$ VAE (ours),  $\phi$ -VAE [Takeishi and Kalousis, 2021] and VAE + ODE solver [Yildiz et al., 2019, Toth et al., 2020] for the damped pendulum vs. ground truth trajectory.

## Extrapolation of pendulum time series: quantitative comparison

**Mutual Information Gap** [Chen et al., 2018], a disentanglement metric

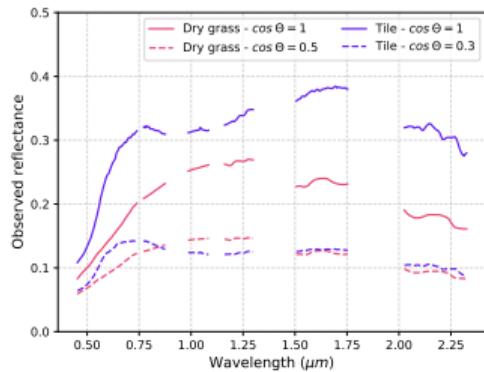
$$\text{MIG}(v_i) = \frac{I(v_i, z_*) - I(v_i, z_\circ)}{H(v_i)}$$

where  $I$  denotes the mutual information,  $H$  denotes the entropy,  $z_*$  is the dimension of the latent code with the highest mutual information and  $z_\circ$  the second highest.

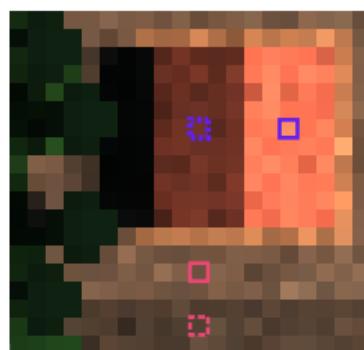
Table: Mean absolute error of reconstruction (in-MAE), extrapolation (out-MAE), and MIG on the test data, averaged over 5 runs.

Model	in-MAE ↓	out-MAE ↓	MIG ↑	
			$\omega$	$\xi$
VAE	<b>5.2</b>	1.4E2	4.8E-3	4.8E-3
$\phi$ -VAE	8.4	9.2E1	1.1E-1	<b>4.5E-2</b>
p <sup>3</sup> VAE	5.3	<b>3.9E1</b>	<b>3.3E-1</b>	2.9E-2

# $p^3$ VAE for hyperspectral image classification: environmental spectral variability



(a) Reflectance spectra



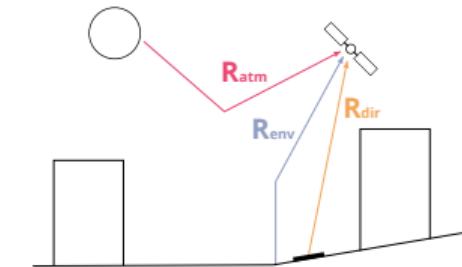
(b) RGB view



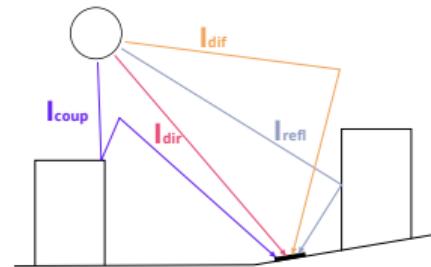
(c) 3D view

Figure: Illustration of the spectral variability induced by variations in local irradiance.

## Atmospheric correction codes: flat ground assumption



(a) Radiance terms in [0.4  $\mu\text{m}$  - 2.5  $\mu\text{m}$ ]



(b) Irradiance terms in [0.4  $\mu\text{m}$  - 2.5  $\mu\text{m}$ ]

Atmospheric correction codes aim to estimate ground-level pixel reflectance  $x$  from radiance at aircraft level:

$$x(\lambda) = \frac{\pi R_{dir}(\lambda)}{I(\lambda)\tau_{dir}}$$

where  $R_{dir}$  is the portion of radiance directly coming from the pixel,  $I = I_{dir} + I_{dif} + I_{coup} + I_{refl}$  is the sun irradiance, and  $\tau_{dir}$  is the direct upward transmission.

Most atmospheric correction codes (e.g. COCHISE [Miesch et al., 2005]) assume that the ground is flat, i.e. that irradiance  $I$  is uniform over the image.

## Derivation of the physical prior

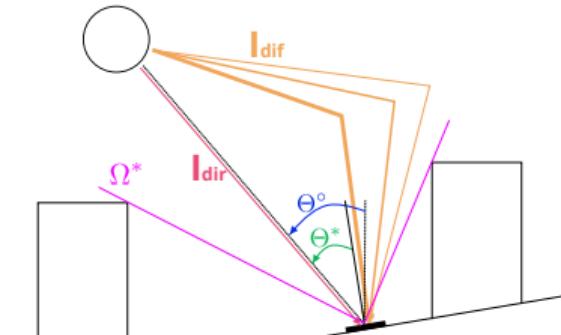
The shift between the assumed irradiance  $I^{code}$  and the actual local irradiance  $I^*$  induce spectral variations that we approximate as follows:

$$\frac{x}{x^*}(\lambda) = \frac{I^*(\lambda)}{I^{code}(\lambda)} \approx \frac{I_{dir}^*(\lambda) + I_{dif}^*(\lambda)}{I_{dir}^{code}(\lambda) + I_{dif}^{code}(\lambda)}$$

where:

$$I_{dir}^*(\lambda) = \delta_{dir}^* \cdot \cos \Theta^* \frac{I_{dir}^{code}(\lambda)}{\cos \Theta^{code}}$$

$$I_{dif}^*(\lambda) = \Omega^* \cdot p_\Theta^* \cdot I_{dif}^{code}(\lambda)$$



$\delta_{dir}^*$ : portion of pixel directly lit by the sun

$\Theta$ : solar zenith angle

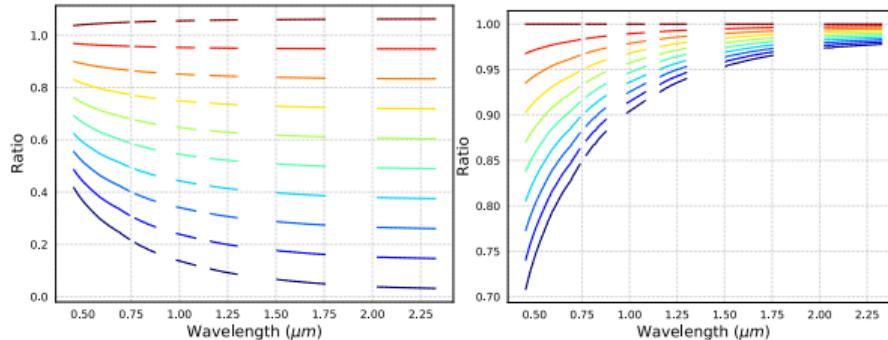
$\Omega^*$ : sky viewing factor

$p_{dif}(\Theta^*)$ : an anisotropy coefficient

## Derivation of the physical prior

The shift between the assumed irradiance  $I^{code}$  and the actual local irradiance  $I^*$  induce spectral variations that we approximate as follows:

$$\frac{x}{x^*}(\lambda) = \frac{I^*(\lambda)}{I^{code}(\lambda)} \approx \frac{I_{dir}^*(\lambda) + I_{dif}^*(\lambda)}{I_{dir}^{code}(\lambda) + I_{dif}^{code}(\lambda)} \quad (1)$$

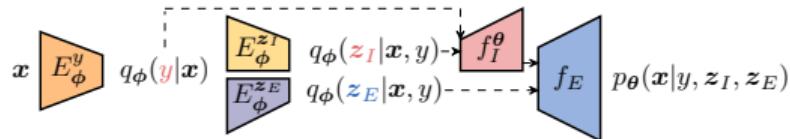


(a)  $I_{dif} = I_{dif}^{code}$ ,  $\delta_{dir} \in [0, 1]$

(b)  $I_{dir} = I_{dir}^{code}$ ,  $\Omega \in [0.3, 1]$

Figure: Illustration of the ratio  $x/x^*$  for some irradiance conditions

## p<sup>3</sup>VAE formulation for hyperspectral classification



$\delta_{dir}^* \cos \Theta^* \leftarrow z_E$  ▷  $z_E$  should be grounded to the direct irradiance

$\Omega^* p_{dif}(\Theta^*) \leftarrow g(z_E)$  ▷ Diffuse irradiance is estimated  
by an empirical function of  $z_E$

$x^* = f_I^\theta(z_I, y)$  ▷ The true reflectance is estimated  
by the neural network

$\mathbb{E}[x] = f_E(z_E, x^*) = \frac{z_E \cdot I_{dir}^{code} + g(z_E) \cdot I_{dif}^{code}}{\cos \Theta^{code} \cdot I_{dir}^{code} + I_{dif}^{code}} x^*$  ▷ Given the irradiance conditions,  $f_E$  computes  
how the spectra would be observed

## Simulated hyperspectral data: quantitative evaluation of accuracy & disentanglement

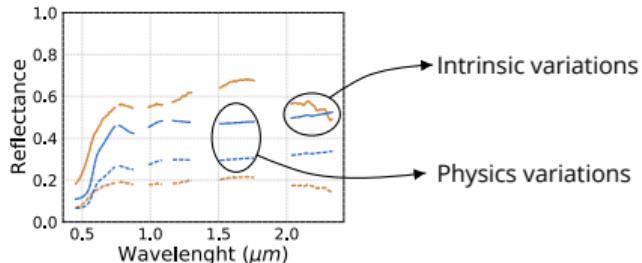
Table: Mean F1 score per class over 10 runs

Inference model		Classes					Average
		Vegetation	Sheet metal	Sandy loam	Tile	Asphalt	
Supervised	CNN [Hu et al., 2015]	0.90	0.81	0.77	0.79	0.75	0.80
	CNN - full annotations	0.92	0.79	<b>0.90</b>	0.87	0.86	0.86
Semi-supervised	ssInfoGAN [Spurr et al., 2017]	0.86	0.79	0.75	0.75	0.69	0.77
	Gaussian VAE [Kingma et al., 2014]	0.93	0.80	0.87	0.86	0.74	0.84
	p <sup>3</sup> VAE w/o $f_E$	0.93	0.81	0.86	0.86	0.76	0.84
	p <sup>3</sup> VAE	<b>0.96</b>	<b>0.97</b>	<b>0.90</b>	<b>0.90</b>	<b>0.93</b>	<b>0.93</b>

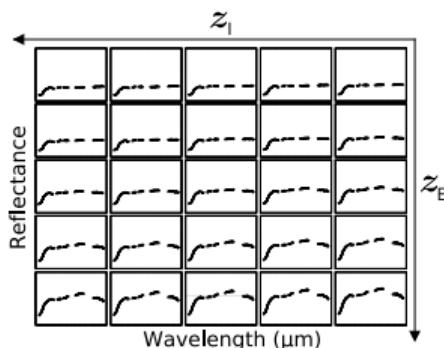
Table: Mean Mutual Information Gap over 10 runs

Model	Factors		
	$y$	$I_{dir}$	$I_{dif}$
ssInfoGAN [Spurr et al., 2017]	4.0E-1	3.7E-2	1.2E-2
Gaussian VAE [Kingma et al., 2014]	2.7E-1	4.2E-2	<b>2.9E-2</b>
p <sup>3</sup> VAE w/o $f_E$	4.3E-1	1.6E-2	6.3E-3
p <sup>3</sup> VAE	<b>5.1E-1</b>	<b>2.0E-1</b>	3.7E-3

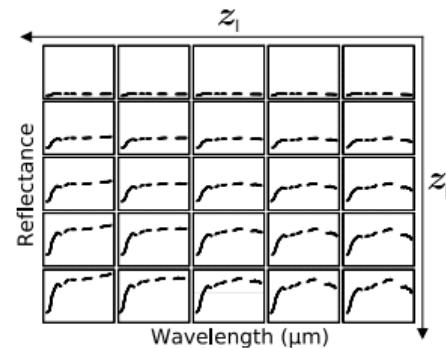
## Simulated hyperspectral data: qualitative evaluation of disentanglement



(a) Spectra of tile from the training data set

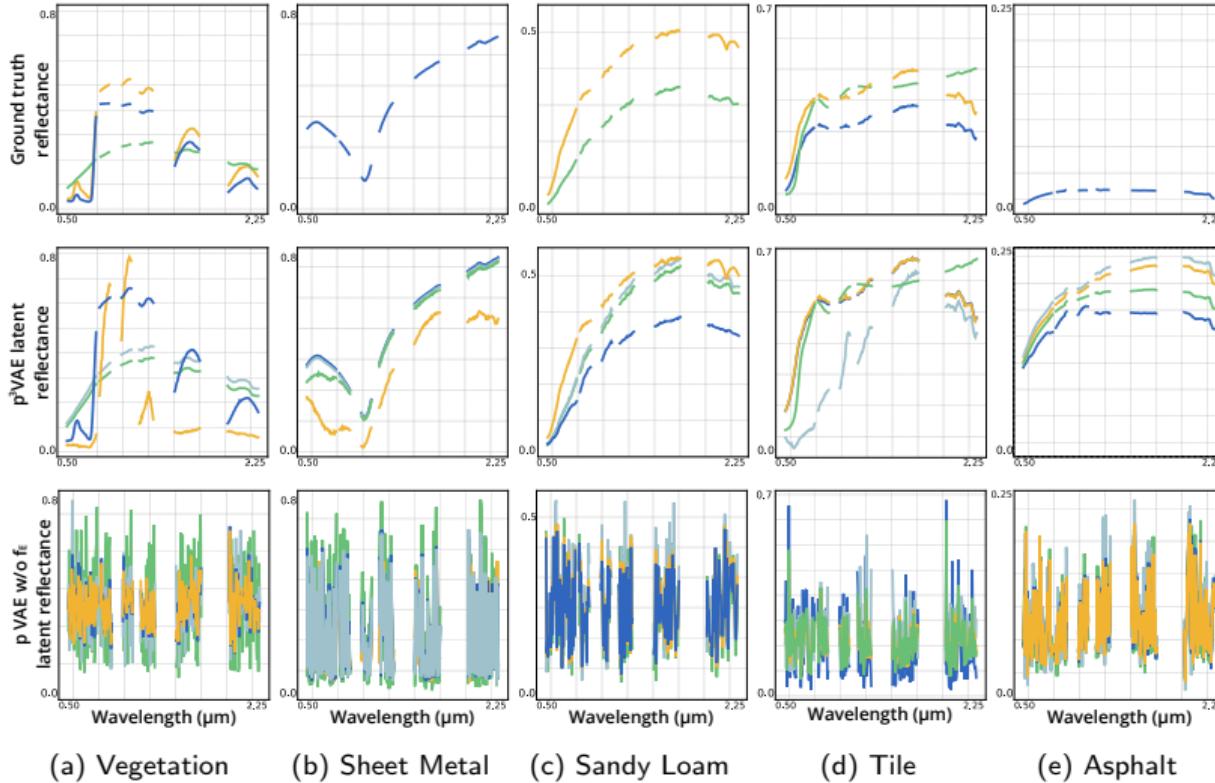


(b) Generated spectra with the gaussian VAE conditioned on the class tile



(c) Generated spectra with  $p^3$ VAE conditioned on the class tile

## Simulated hyperspectral data: interpretability



## Simulated hyperspectral data: ablation study about the gradient-stopping

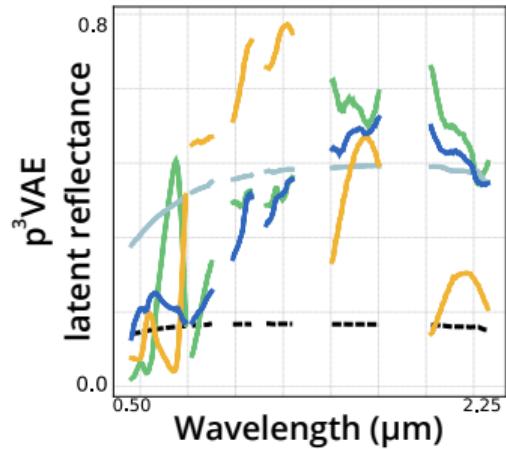


Figure: Estimated spectra of asphalt by  $p^3$ VAE  
w/o gradient-stopping

Table:  $p^3$ VAE average F1 score

Inference technique	w/o GS	w/ GS
Simulated data set		
$q_\phi(y \mathbf{x})$	0.86	0.86
$p_\theta(y \mathbf{x})$	0.92	0.93
Real data set		
$q_\phi(y \mathbf{x})$	0.88	0.90
$p_\theta(y \mathbf{x})$	0.73	0.84

## Real hyperspectral data: qualitative comparisons

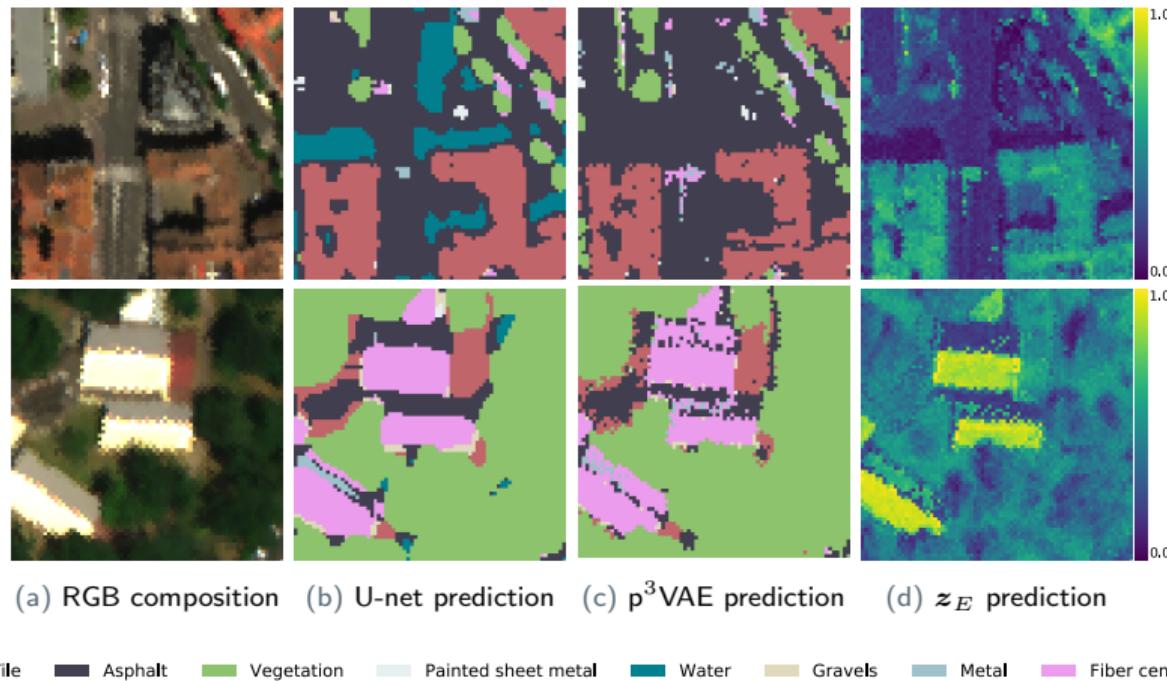


Figure: Land cover maps computed by a U-net like model [Stoian et al., 2019] and  $p^3$ VAE from CAMCATT-AI4GEO hyperspectral images over Toulouse.

1. Motivation: mapping the land cover in large urban areas from hyperspectral images
2.  $p^3$ VAE framework through a toy example: the damped pendulum
3. Experimental results: extrapolation of pendulum time series & hyperspectral image classification
4. Conclusions & Perspectives

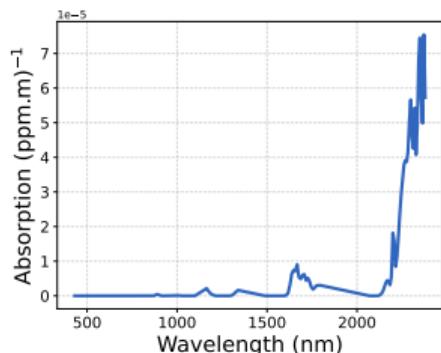
## Conclusions

We have introduced  $p^3$ VAE, a framework that integrates *a priori* physical knowledge about the environmental factors of variation in a VAE. Empirically,

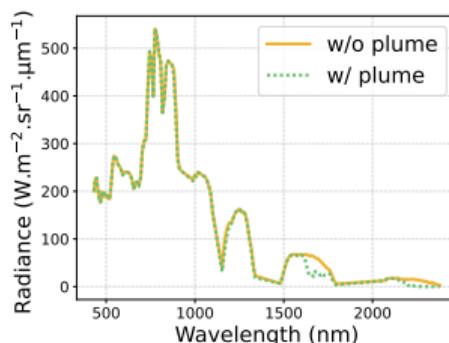
- ▶  $p^3$ VAE improves the disentanglement of the latent space in a physically meaningful way with the partial supervision of the intrinsic factors of variation,
- ▶  $p^3$ VAE increases the extrapolation capacities of the VAE on data generated from out-of-distribution environmental factors of variation.

## Perspectives: application of $p^3$ VAE to other problems

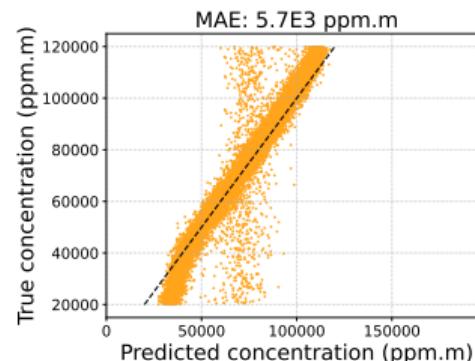
- We already have promising results for the inversion of methane plumes from hyperspectral satellite images on simulated data → application on real images.



(a) Methane monochromatic absorption



(b) Radiance spectra w/ & w/o methane plume



(c)  $p^3$ VAE methane concentration predictions VS ground truth

- $p^3$ VAE could be applied to the (more complex) inversion of industrial stack plumes.

## Perspectives: enhancement of p<sup>3</sup>VAE

- ▶ Improvement of p<sup>3</sup>VAE disentanglement capacity with disentanglement losses?
- ▶ Disentangled encodings need to satisfy local isometry [Horan et al., 2021]. How does the integration of physics foster the local isometry property?
- ▶ Recent works relax the assumption of independent factors of variation [Oublal et al., 2024]: could this explain how physics (without auxiliary loss) helps disentanglement?
- ▶ Could we adapt a pretrained instance of p<sup>3</sup>VAE to a new data distribution (e.g. new land cover classes) while keeping the disentanglement of the latent space?

Thank you for your attention! Questions?

-  Carbonneau, M.-A., Zaidi, J., Boilard, J., and Gagnon, G. (2022).  
**Measuring disentanglement: A review of metrics.**  
*IEEE transactions on neural networks and learning systems*, 35(7):8747–8761.
-  Chen, R. T. Q., Li, X., Grosse, R. B., and Duvenaud, D. K. (2018).  
**Isolating sources of disentanglement in variational autoencoders.**  
In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc.
-  Gastellu-Etchegorry, J.-P., Grau, E., and Lauret, N. (2012).  
**Dart: A 3d model for remote sensing images and radiative budget of earth surfaces.**  
*Modeling and simulation in engineering*, (2).
-  Horan, D., Richardson, E., and Weiss, Y. (2021).  
**When is unsupervised disentanglement possible?**  
*Advances in Neural Information Processing Systems*, 34:5150–5161.
-  Hu, W., Huang, Y., Wei, L., Zhang, F., and Li, H. (2015).  
**Deep convolutional neural networks for hyperspectral image classification.**  
*Journal of Sensors*, 2015(1):258619.
-  IPCC (2021).  
**Weather and Climate Extreme Events in a Changing Climate**, page 1513–1766.  
Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
-  Kingma, D. P., Mohamed, S., Jimenez Rezende, D., and Welling, M. (2014).  
**Semi-supervised learning with deep generative models.**  
*Advances in neural information processing systems*, 27.

-  Kingma, D. P. and Welling, M. (2014).  
**Auto-encoding variational bayes.**  
In Bengio, Y. and LeCun, Y., editors, *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*.
-  Locatello, F., Bauer, S., Lucic, M., Raetsch, G., Gelly, S., Schölkopf, B., and Bachem, O. (2019).  
**Challenging common assumptions in the unsupervised learning of disentangled representations.**  
In *international conference on machine learning*, pages 4114–4124. PMLR.
-  Miesch, C., Poutier, L., Achard, V., Briottet, X., Lenot, X., and Boucher, Y. (2005).  
**Direct and inverse radiative transfer solutions for visible and near-infrared hyperspectral imagery.**  
*IEEE Transactions on Geoscience and Remote Sensing*, 43(7):1552–1562.
-  O'Riordan, R., Davies, J., Stevens, C., and Quinton, J. N. (2021).  
**The effects of sealing on urban soil carbon and nutrients.**  
*SOIL*, 7(2):661–675.
-  Oublal, K., Ladjal, S., Benhaim, D., LE BORGNE, E., and Roueff, F. (2024).  
**Disentangling time series representations via contrastive independence-of-support on l-variational inference.**  
In *The Twelfth International Conference on Learning Representations*.
-  Pereira, M. C., O'Riordan, R., and Stevens, C. (2021).  
**Urban soil microbial community and microbial-related carbon storage are severely limited by sealing.**  
*Journal of Soils and Sediments*, 21:1455–1465.
-  Scalenghe, R. and Marsan, F. A. (2009).  
**The anthropogenic sealing of soils in urban areas.**  
*Landscape and urban planning*, 90(1-2):1–10.

-  Spurr, A., Aksan, E., and Hilliges, O. (2017).  
**Guiding infogan with semi-supervision.**  
In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 119–134. Springer.
-  Stoian, A., Poulain, V., Inglada, J., Poughon, V., and Derksen, D. (2019).  
**Land cover maps production with high resolution satellite image time series and convolutional neural networks: Adaptations and limits for operational systems.**  
*Remote Sensing*, 11(17):1986.
-  Takeishi, N. and Kalousis, A. (2021).  
**Physics-integrated variational autoencoders for robust and interpretable generative modeling.**  
*Advances in Neural Information Processing Systems*, 34:14809–14821.
-  Toth, P., Rezende, D. J., Jaegle, A., Racanière, S., Botev, A., and Higgins, I. (2020).  
**Hamiltonian generative networks.**  
In *International Conference on Learning Representations*.
-  Yildiz, C., Heinonen, M., and Lahdesmaki, H. (2019).  
**Ode2vae: Deep generative second order odes with bayesian neural networks.**  
*Advances in Neural Information Processing Systems*, 32.
-  Yin, Y., Vincent, L., Jérémie, D., de Bezenac, E., Ayed, I., THOME, N., et al. (2020).  
**Augmenting physical models with deep networks for complex dynamics forecasting.**  
In *International Conference on Learning Representations*.

## Balancing between $f_I^\theta$ and $f_E$ : related work

- APHYNITY [Yin et al., 2020] complements a physical prior  $F_p$ , that belongs to a known family  $\mathcal{F}_p$ , with a neural network  $F_a \in \mathcal{F}_a$  *with minimal norm*:

$$\min_{F_p \in \mathcal{F}_p, F_a \in \mathcal{F}} \|F_a\| \quad \text{subject to} \quad \forall X \in \mathcal{D}, \forall t, \frac{dX_t}{dt} = (F_p + F_a)(X_t)$$

where  $\mathcal{D}$  is the training data set.

- Physics-integrated VAEs [Takeishi and Kalousis, 2021] combine a neural network  $f_A$  with physical prior knowledge  $f_P$  into the decoder of a VAE. *They balance the capacity of  $f_A$  with a regularization term based on "physics-only" decoder.*
- Both methods rely on the fact that the physical prior knowledge, by itself, can provide a "good enough" solution. The neural network, in both cases, only complements known models.  
 → **not adapted to the category of problems we are targeting.**

## Simulation of hyperspectral data with DART [Gastellu-Etchegorry et al., 2012]

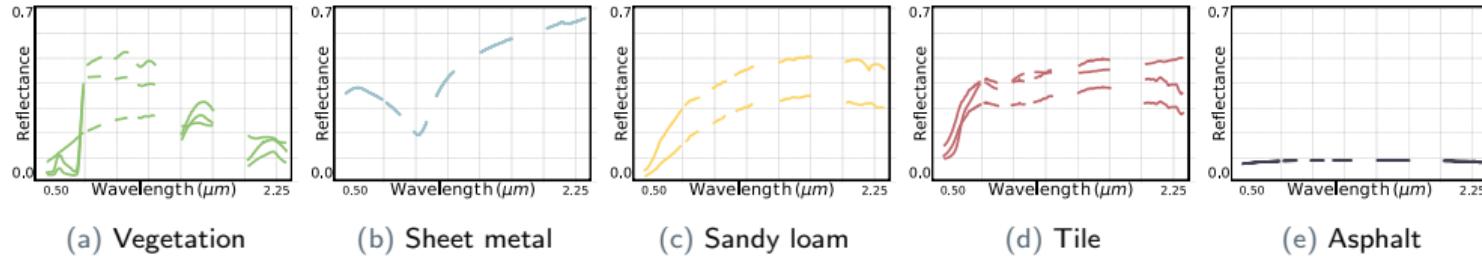


Figure: Reflectance spectra used for the simulation

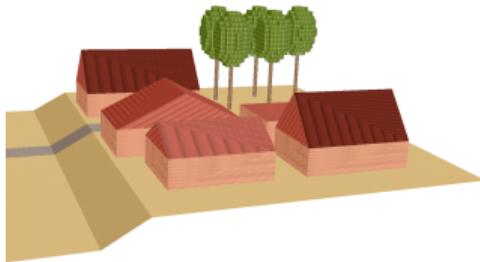


Figure: 3D scene used to simulate the hyperspectral data for training

Table: Representativity of generative factors in the training and test sets

Set	Generative factors			
	$y$	$I_{dir}$	$I_{dif}$	Intra-class mixing
Labeled training set	+++	+	+	++
Unlabeled training set	+++	++	++	++
Test set	+++	+++	+++	+++