Interpretable Representation Learning for High Resolution Satellite Image Time Series

Yoël Zerah

Silvia Valero, Jordi Inglada,

Centre d'Etudes Spatiales de la BIOsphere (CESBIO)

May 24th 2022

Sentinel-2 optical images

New opportunity for land monitoring

The Sentinel-2 Constellation has 2 satellites dedicated to land masses monitoring, launched in 2015 and 2017.

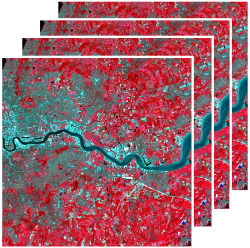


Figure 1 – Satellite Image Time Series

Figure 2 – Sentinel 2 image with clouds

Challenges of S2 data exploitation

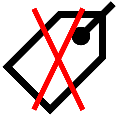
- High dimension data
- Irregular sampling (spatial & temporal)
- Labelled datasets are costly and rare
- Complex with high variability

Satellite data low dimensional representation learning

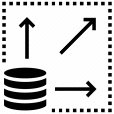
Data representation

Finding transformation $z = r(x)$ of data x that is useful for subsequent applications.

Desired representation requirements for satellite data

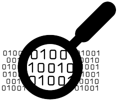


Unsupervised



Deployable at large scale

Probabilistic

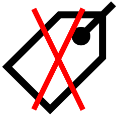


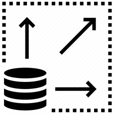
Interpretable

Data representation

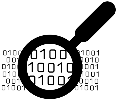
Finding transformation $z = r(x)$ of data x that is useful for subsequent applications.

Desired representation requirements for satellite data


Unsupervised


Deployable at large scale


Probabilistic


Interpretable

Existing representation learning methodologies

Deep generative methodologies have been proposed for combining **deep learning** architectures and **Bayesian** inference framework with **unsupervised training** :
$\Rightarrow$ Variational Auto Encoders

$$\mathcal{L}(q_\lambda) = -\mathbb{E}[\log p(x|z)] + \mathbb{KL}(q_\lambda(z|x)||p(z))$$

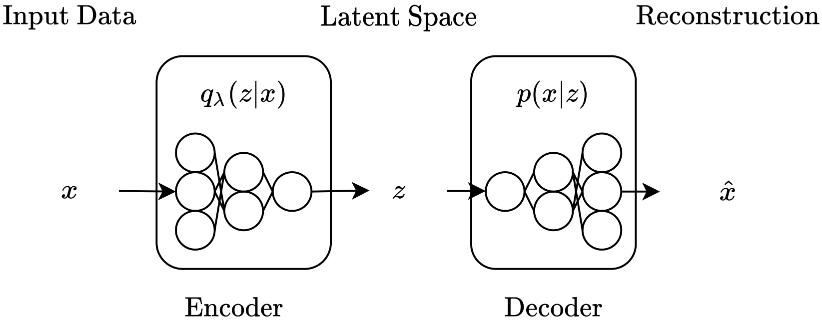


Figure 3 – Variational Auto Encoder [1]

Latent space

- Usual prior on latent space : $p(z) = \mathcal{N}(0, I)$
- Reparametrization trick :

$$z = \mu_z(x) + \varepsilon \cdot \sigma_z(x), \qquad \varepsilon \sim \mathcal{N}(0, I) \qquad \Rightarrow z \sim \mathcal{N}(\mu_z(x), \sigma_z^2(x)) \qquad (1)$$

1. Diederik P Kingma et Max Welling. *Auto-Encoding Variational Bayes*. 2014. arXiv : 1312.6114 [stat.ML].

Objectives

Limitation of using VAEs

- Not developped for SITS
- Uninterpretable latent space
- Usual $p(z) \sim \mathcal{N}(0, 1)$ is arbitrary

Objectives

- Adapt VAE for SITS
- Infer interpretable representation with VAE
- Integrate prior knowledge of data into VAE latent space

What is a good representation of physical data ?

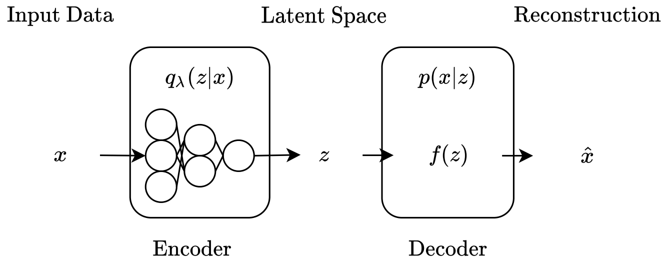
$\Rightarrow$ parameters of physical models

Summary

1. Integrating physics into VAE

2. NDVI time series model

3. Experimental results

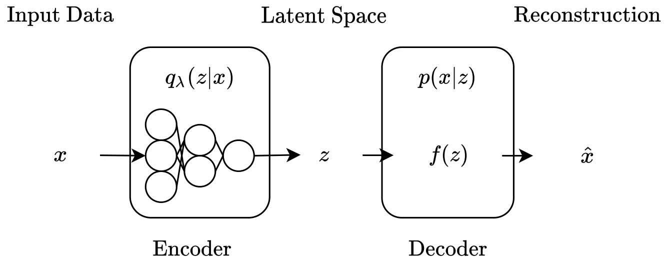


VAE with Decoder-Simulator (VAE-DS)

- Decoder's neural network is replaced by a user-defined model [2]
- Latent variables are **semantically bound to model's parameters**

2. Miguel A. Aragon-Calvo. "Self-supervised learning with physics-aware neural networks – I. Galaxy model fitting". In : *Monthly Notices of the Royal Astronomical Society* (2020).

Loss of VAE with decoder-simulator

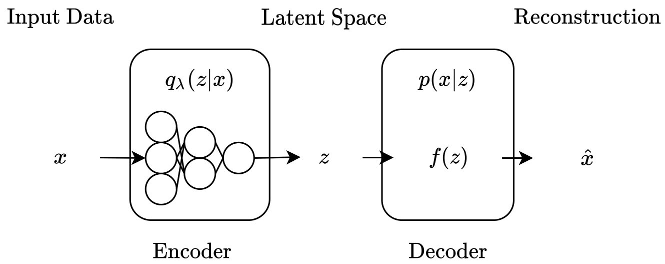


Loss of VAE-DS

Without specific prior $p(z)$:

$$\mathcal{L}(q_\lambda) = -\mathbb{E}[\log p(x|z)] + \cancel{\mathbb{KL}(q_\lambda(z|x) \| p(z))}$$

Loss of VAE with decoder-simulator



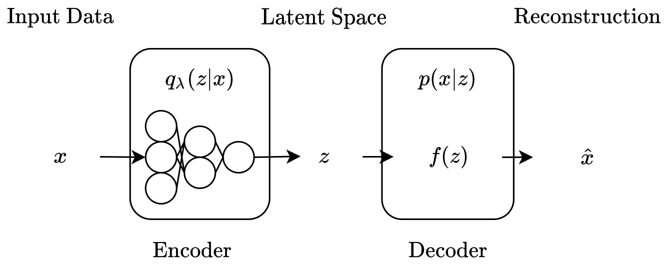
Loss of VAE-DS

Without specific prior $p(z)$:

$$\mathcal{L}(q_\lambda) = -\mathbb{E}\left[\log p(\mathrm{x}|\mathrm{z})\right] + \xcancel{\mathbb{KL}\left(q_\lambda(\mathrm{z}|\mathrm{x}) \| p(\mathrm{z})\right)}$$

$$\approx \frac{1}{L}\sum_{i=1}^{L} \log 2\pi\sigma_x(z_i) + \frac{(x - \mu_x(z_i))^2}{\sigma_x^2(z_i)}$$

Loss of VAE with decoder-simulator



Decoder's distribution's output parameters

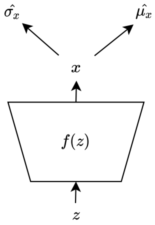


Figure 4 – **Simulator-decoder**

Monte Carlo sampling of latent space :

- $\mu_x(z) \approx \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i$
- $\sigma_x(z) \approx \frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i - \mu_x(z))^2$

with $\hat{x}_i = f(z_i)$.

What priors can be brought to latent space ?

- Model parameters can be bounded
- Model parameters can be ordered

Bounding latent distributions

Bounded distributions can be directly sampled with **Inverse Transform Method** :

$$z = F_{\mathcal{A}}^{-1}(u), \quad u \sim \mathcal{U}(0,1) \quad \Rightarrow \quad z \sim \mathcal{A} \tag{2}$$

with $F_{\mathcal{A}}^{-1}$ a tractable inverse CDF of distribution $\mathcal{A}$.

Ordering Latent Distributions

ordered samples $z_0 < z_1$ are rectified :

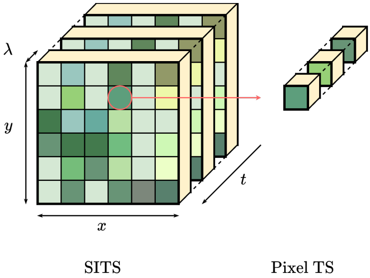
$$\begin{aligned} z_0 &\leftarrow z_0 \\ z_1 &\leftarrow \max(z_0, z_1) \end{aligned} \tag{3}$$

Summary

Normalized Difference Vegetation Index (NDVI)

$$\text{NDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{R}}}{\rho_{\text{NIR}} + \rho_{\text{R}}} \in [-1, 1]$$

NDVI characterizes photsynthetic vegetation vigor and activity.



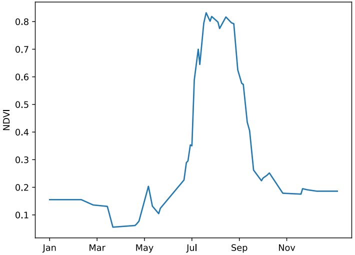
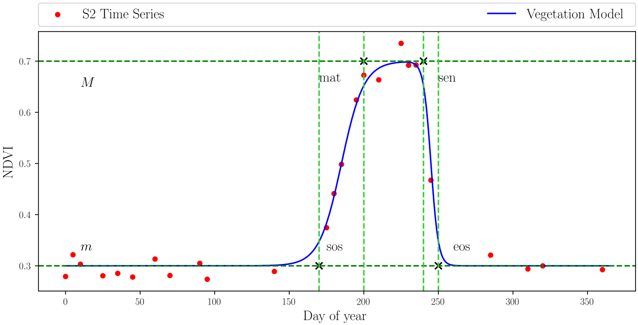


Figure 5 – **Sunflower NDVI time series**

Phenological model

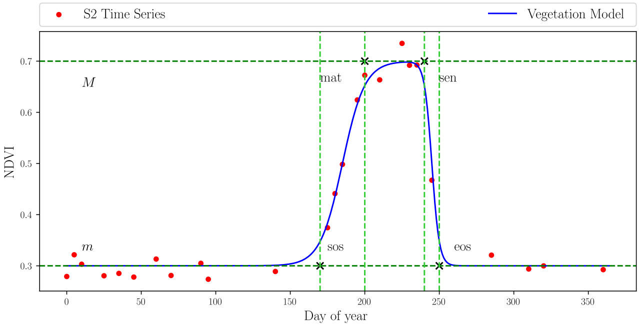


Variable	Description
M	Maximum of double logistic
m	minimum of double logistic
sos	date of *Start Of Season*, the start of NDVI growth
mat	date of *Maturity*, the end of the NDVI growth
sen	date of *Senescence*, the start of NDVI decay
eos	date of *End Of Season*, end of NDVI decay

Figure 6 – **Phenological model** [3]

3. Xiaoyang Zhang et al. "Monitoring vegetation phenology using MODIS". In : *Remote Sensing of Environment* 84.3 (2003), p. 471-475.

Phenological model



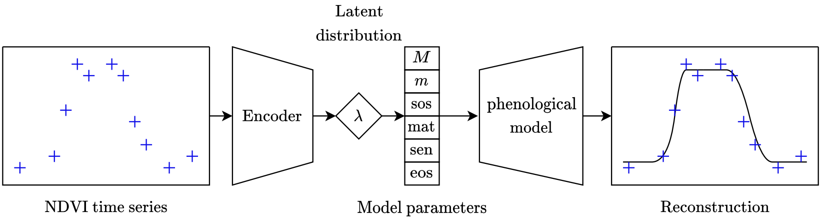
Double-logistic model

$$f_\phi(t) = \left(\max_{\text{NDVI}} - \min_{\text{NDVI}}\right)\left(S_1(t) - S_2(t)\right) + \min_{\text{NDVI}}$$

$$S_1(t) = \left(1 + \exp\left(2\frac{\text{sos} + \text{mat} - 2t}{\text{mat} - \text{sos}}\right)\right)^{-1} \qquad S_2(t) = \left(1 + \exp\left(2\frac{\text{sen} + \text{eos} - 2t}{\text{eos} - \text{sen}}\right)\right)^{-1}$$

Proposed approach : integrating phenological model into VAE

Latent variables are semantically bound to phenological parameters.



Summary

Datasets

S2 dataset

- 10^6 S2 time series
- 2018 time series from 31TCJ tile
- 20 land cover classes
- Time series interpolated to 5-day sampling.

Simulated dataset

Using the phenological model, we simulate NDVI Time series for validation purposes.

Experimental setup

Training Setup

- Number of Monte Carlo samples of latent space per time series : 10
- Learning rate : 10^{-4}
- Latent distribution : Truncated gaussians

Validation experiences

- Evaluation of the quality of reconstruction
- Evaluation of the quality of inferred phenological parameters.

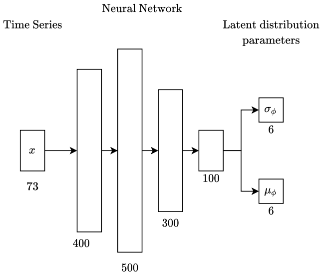


Figure 6 – **Simple encoder architecture with 4 hidden layers and ReLU activation**

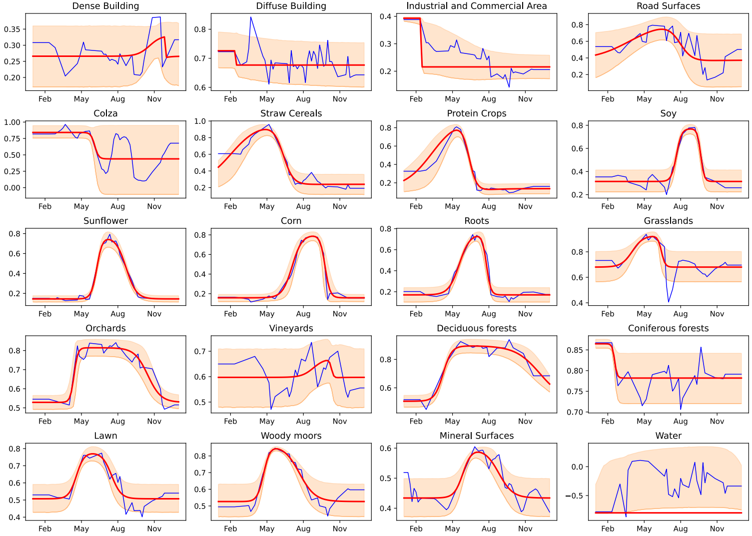


Figure 7 – Time series reconstructions for 20 class samples. Blue : Original time serie - Red : Reconstruction of phenological mode - Orange 90% confidence interval

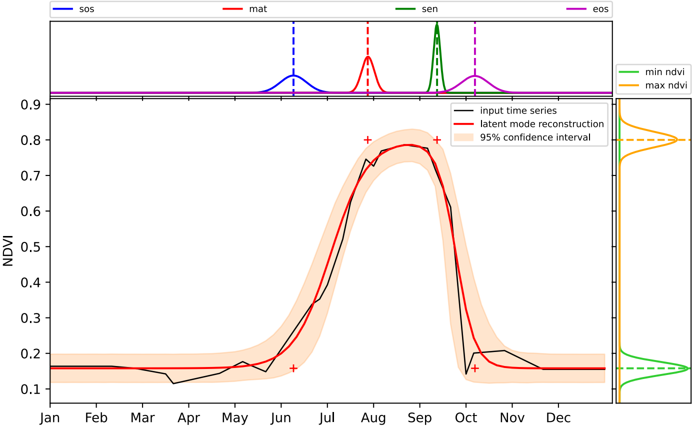


Figure 7 – **Reconstruction and latent distributions of corn ndvi time series**

Compartive evaluation of phenological parameters inference

We use our method to solve the phenological model inverse problem. We compare it to 2 other classical methods.

Characteristics	MCMC [3]	NN Regression	VAE-DS
Unsupervised	✓		✓
Probabilistic	✓	✓	✓
Large Scale		✓	✓
Training Dataset	None	Simulated	S2
Inferred distribution	Full approximate distribution	Truncated Gaussian parameters	Truncated Gaussian parameters

Table 1 – **List of experiences for inverting the phenological model**

Note :

- NN regression has the same NN structure than encoder of VAE-DS

3. Markov Chain Monte Carlo

Evaluation of the quality of phenological parameters estimation

Inference error of the three methods :

- NN regression has the best MAE
- MCMC & VAE-DS are comparable.

Method	MCMC	NN Regression	VAE-DS
Point estimate	Median	Mode	Mode
M	0.04	0.04	0.06
m	0.02	0.01	0.02
sos	8.58	6.83	10.34
mat	11.78	7.84	10.44
sen	11.54	7.22	11.96
eos	12.04	6.83	14.36

Table 2 – **Mean Absolute Error**

Evaluation of the quality of inferred latent distributions

Quality of 5-95th centile confidence intervals :

- MCMC & NN regression very close confidence interval belonging rate
- VAE-DS underestimates uncertainty
- MCMC has smaller (more precise) confidence intervals

	Confidence Interval Belonging Rate			Mean Confidence Interval Width		
	MCMC	NN Regression	VAE-DS	MCMC	NN Regression	VAE-DS
M	88.90	91.62	62.63	0.08	0.16	0.13
m	85.20	89.47	96.80	0.03	0.06	0.13
sos	85.40	91,19	36.88	10.12	29.90	14.53
mat	84.40	90,58	22.18	10.46	33.29	7.54
sen	84.00	90,54	34.37	6.41	30.14	16.03
eos	83.30	90,24	52.09	7.12	28.89	27.54

Table 3 – **5-95th centiles confidence interval metrics**

Inference performance : inference time

Deep learning methods have much faster inferrence.

Characteristics	MCMC	NN Regression	VAE-DS
Unsupervised	✓		✓
Probabilistic	✓	✓	✓
Large Scale		✓	✓
Training Dataset	None	Simulated	S2
Inferred distribution	Full approximate distribution	Truncated Gaussian parameters	Truncated Gaussian parameters
Inference Time per time series	≈ 10 s	$\approx 10^{-5}$ s	$\approx 10^{-5}$ s

Table 4 – **Inference time of phenological distributions (Laptop i5 CPU)**

Conclusion

Contributions

- New physics-guided unsupervised representation learning methodology.
- Methodology can be used to solve inverse problems.

Envisonned developments

- Encoder complexification
- Apply methodology to more complex physics models, for different data
- Use auxillary neural network to model residuals [4]

4. Naoya Takeishi et Alexandros Kalousis. "Physics-Integrated Variational Autoencoders for Robust and Interpretable Generative Modeling". In : *CoRR* (2021).

Thank you for your attention !

yoel.zerah@univ-toulouse.fr