Physics-aware Machine Learning in the Earth sciences

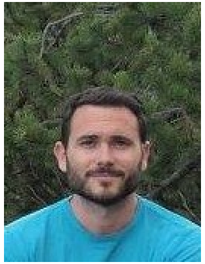
Gustau Camps-Valls
Image Processing Laboratory
Universitat de València

http://isp.uv.es @isp_uv_es gustau.camps@uv.es

With a lot of help from my friends ...

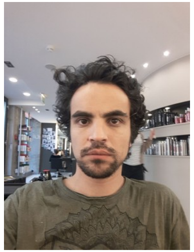




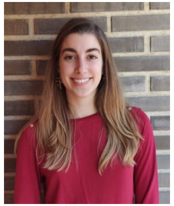














With a lot of help from my $$$ friends ...

esa

NASA

erc
European Research Council
Established by the European Commission

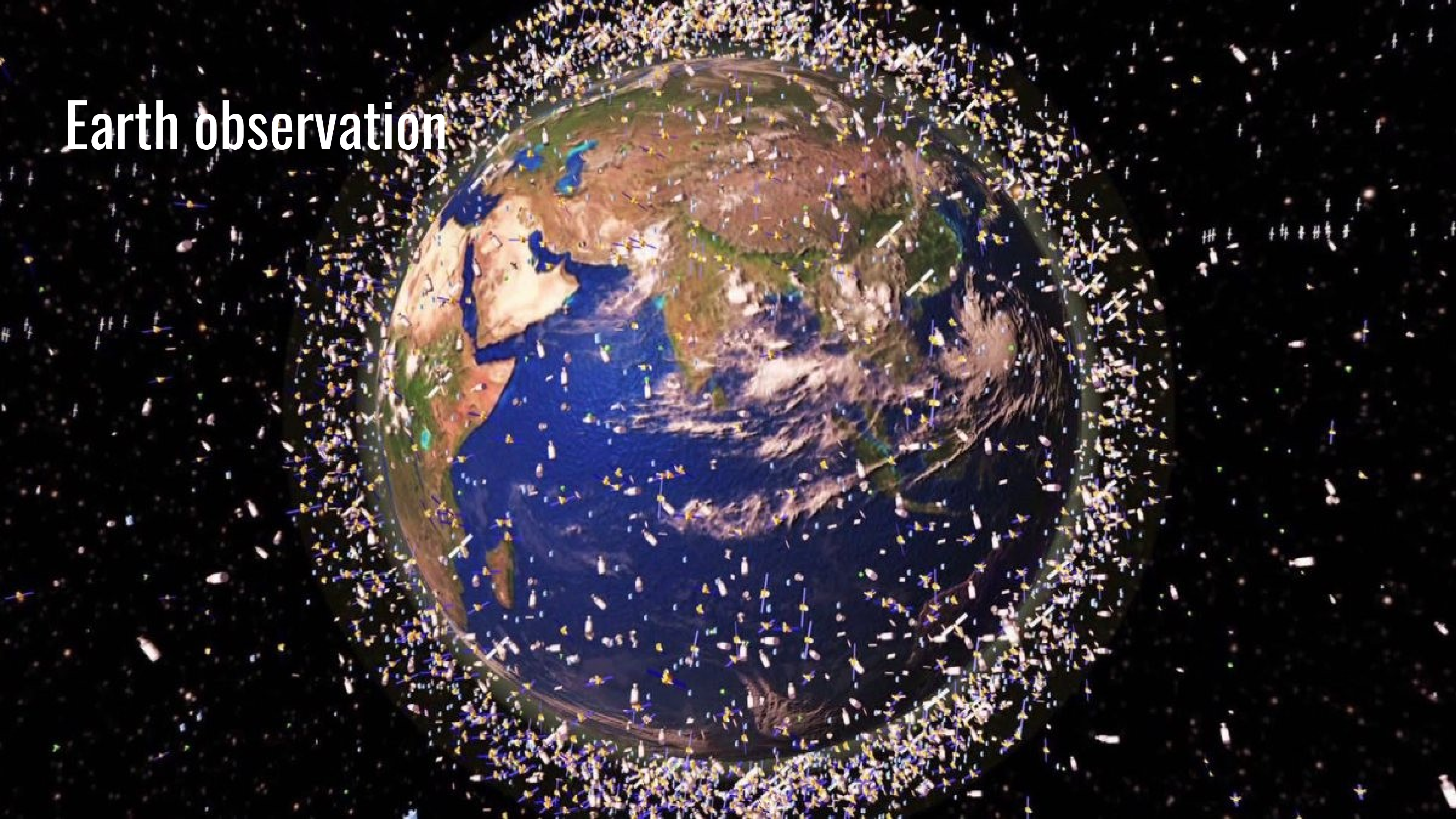
EUMETSAT

UN environment programme

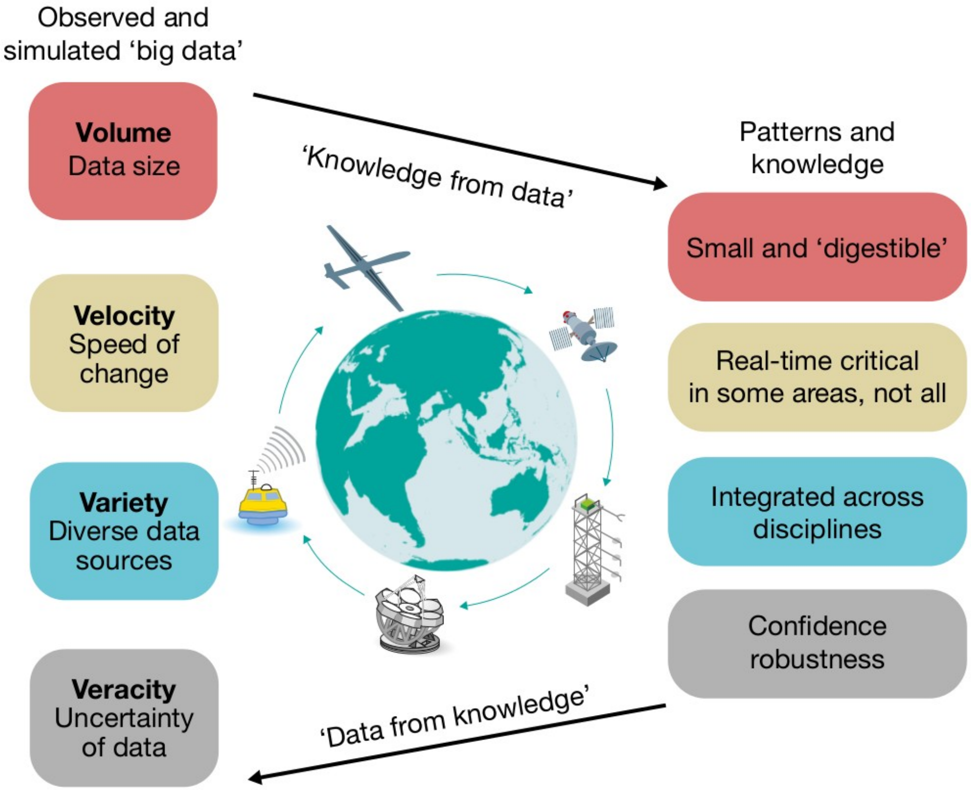
MARIE CURIE ACTIONS

Earth science

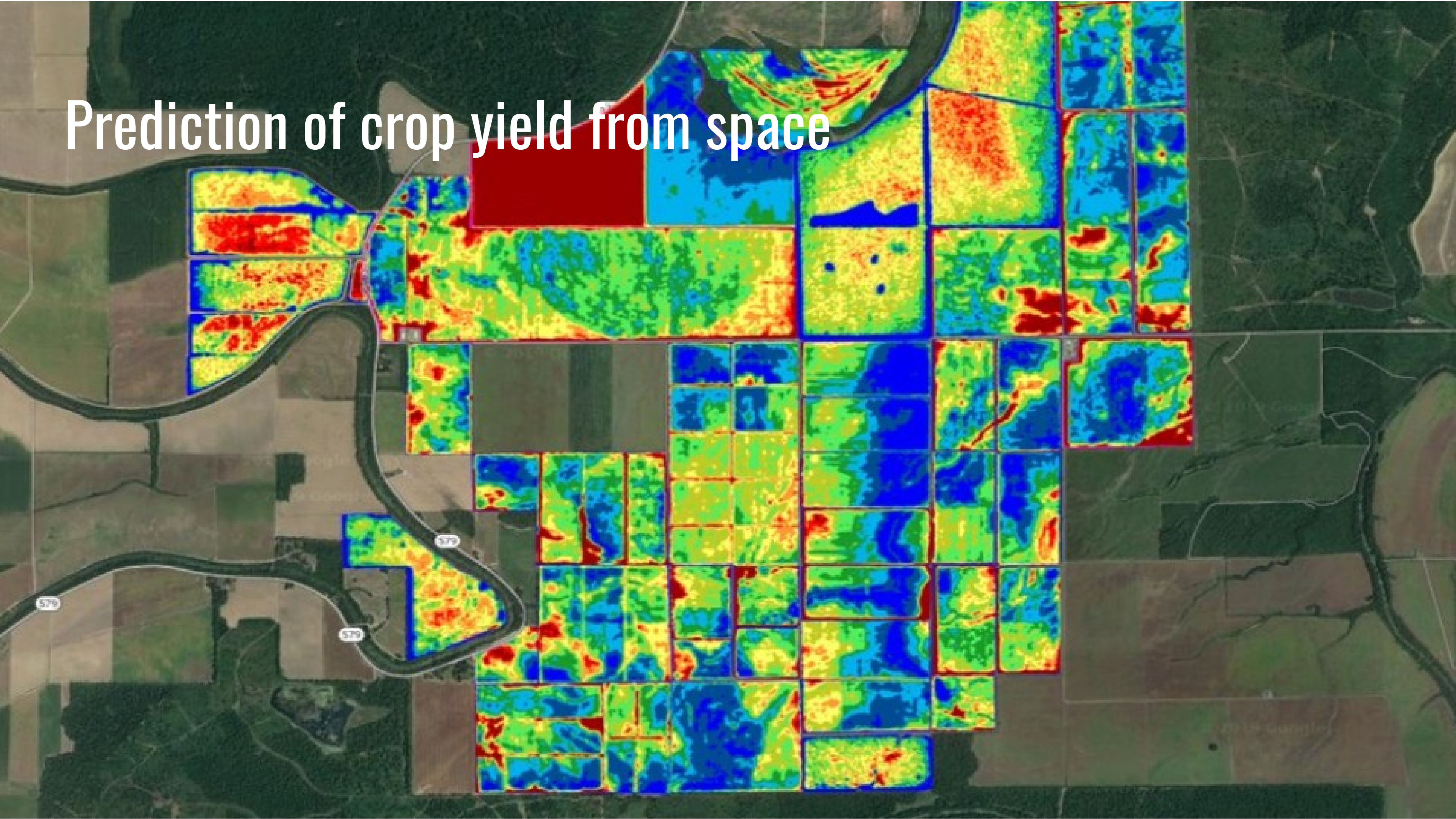
Earth observation



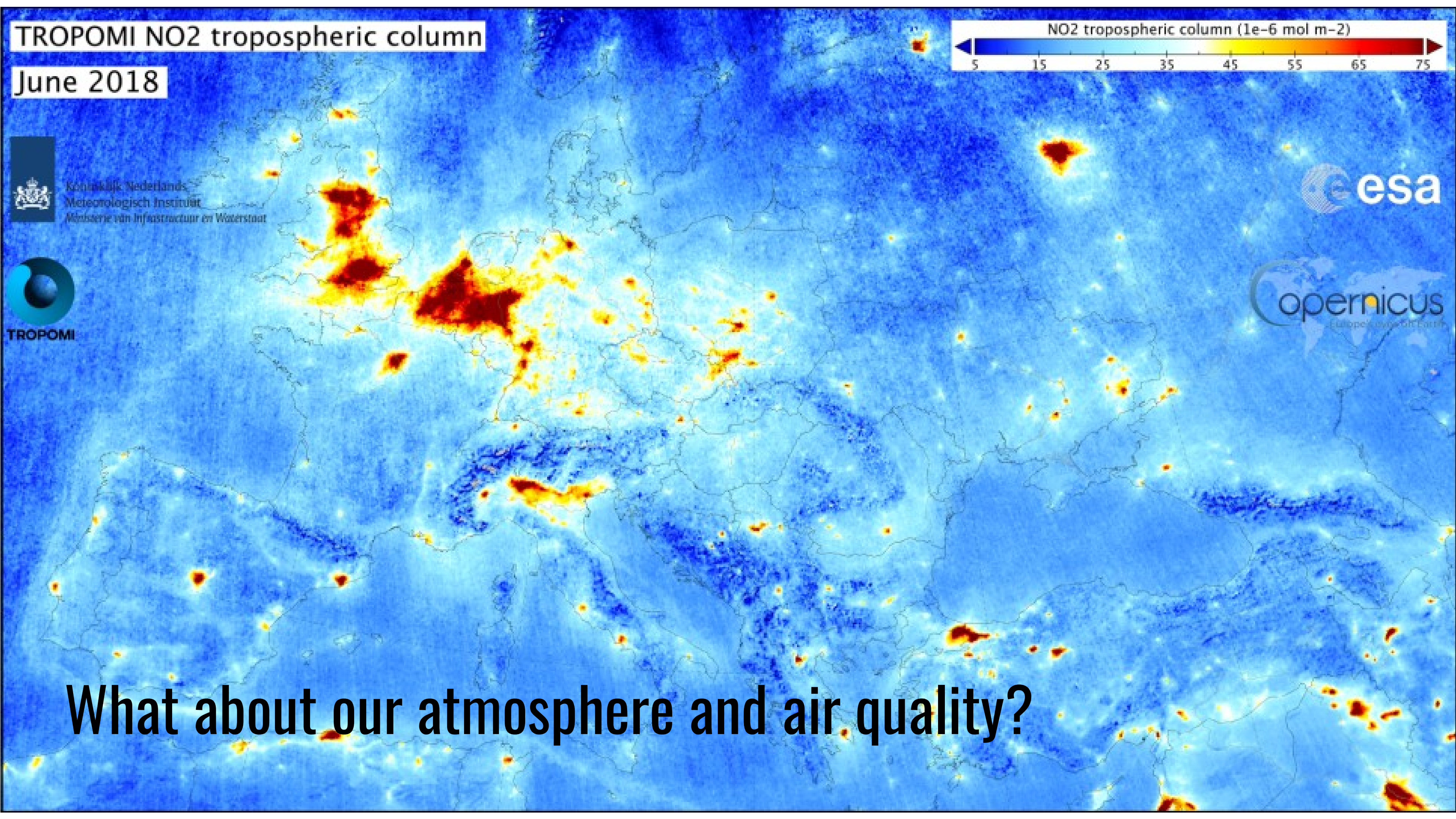
The Earth data deluge



Prediction of crop yield from space




How is our coastline and ocean?

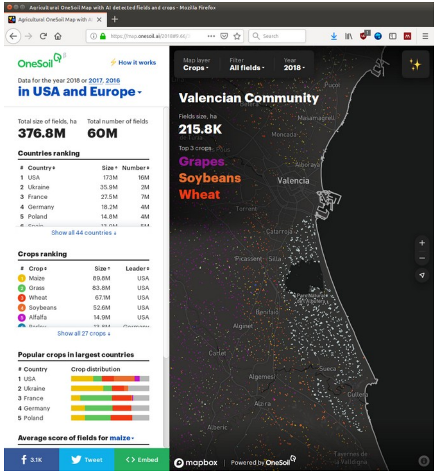


What about our atmosphere and air quality?

Some machine learning applications

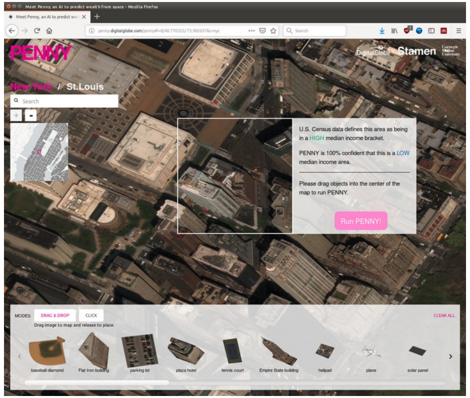
One soil map

https://map.onesoil.ai



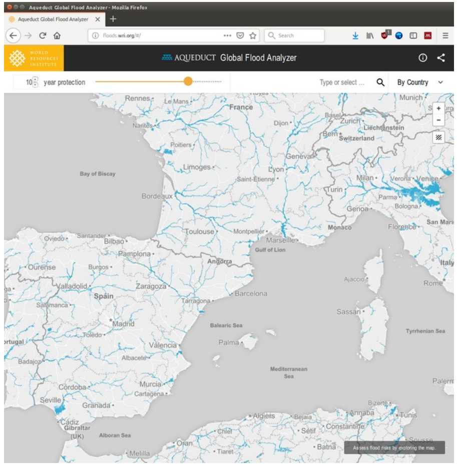
Global wealth map

http://penny.digitalglobe.com



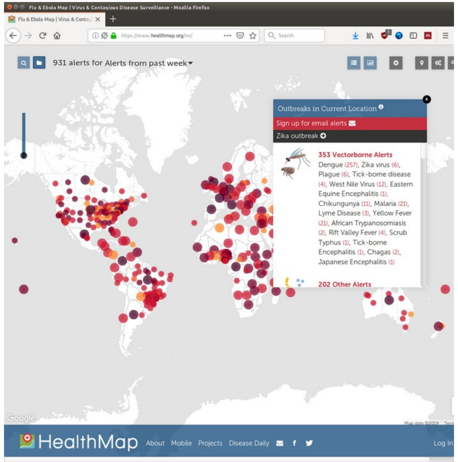
Flood analyzer

http://floods.wri.org



Disease mapping

https://www.healthmap.org



The challenges

ML in Earth science rocks... *only* when some things happen!

- Strong spatial and temporal correlations
- Big data accessible
- Cheap computing resources available
- Fast scalable ML models available
- No expert knowledge needed
- High prediction accuracy is enough
- Understanding the system is not that relevant

Machine/deep learning challenges

- Do Models respect Physics Laws?
- What did the ML model learn?
- Do they get cause-effect relations?


The New York Times

Opinion

OP-ED CONTRIBUTORS

Eight (No, Nine!) Problems With Big Data

By Gary Marcus and Ernest Davis


nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio &

Archive > Volume 538 > Issue 7623 > News Feature > Article

NATURE | NEWS FEATURE

Can we open the black box of AI?

Artificial intelligence is everywhere. But before scientists trust it, they first need t
understand how machines learn.

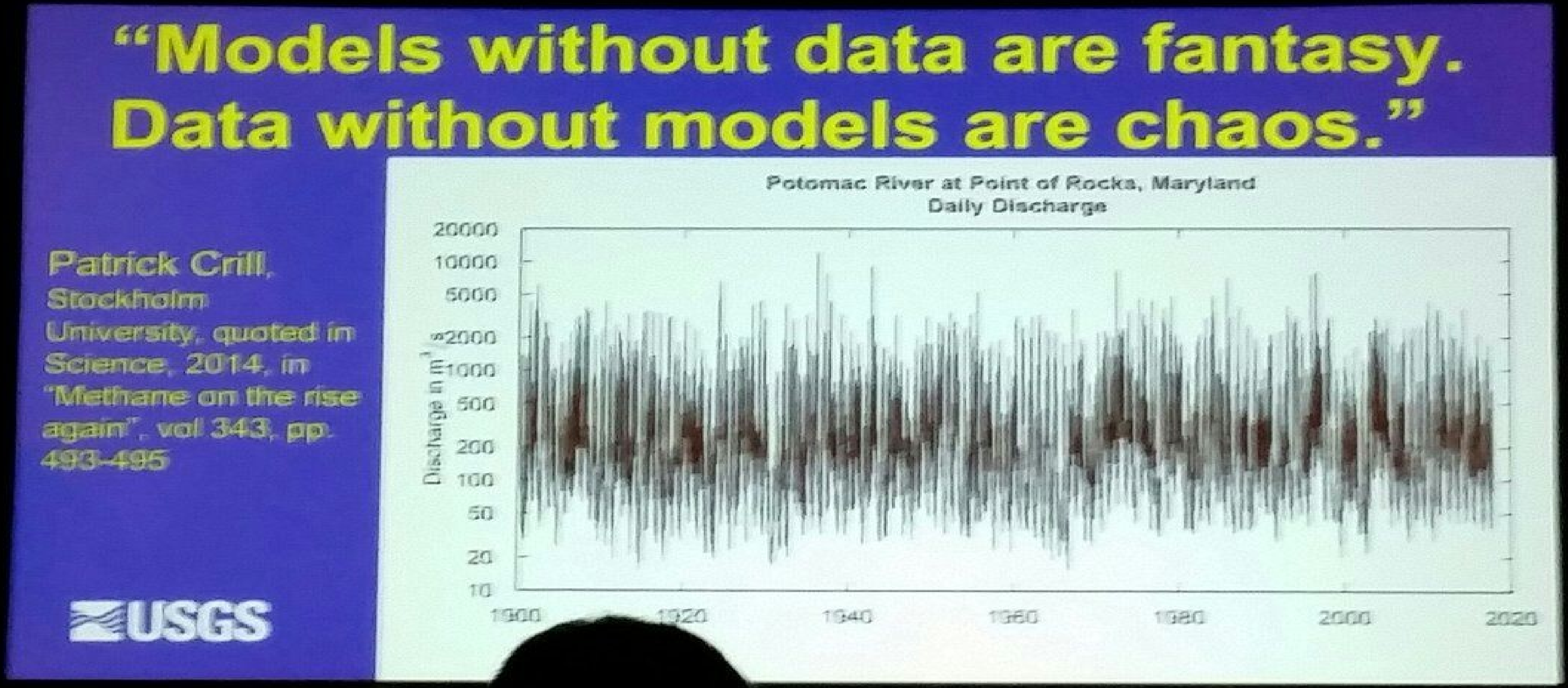
Davide Castelvecchi

Physics-*aware** machine learning

$$F(X, \frac{\partial c}{\partial t} + \mathbf{v}\nabla c = 0) = y$$

* aka physics-guided, physics-informed, ...

The truth is that...



At AGU 2017, New Orleans, USA

A simple taxonomy

A — Data-model blending
- Joint Gaussian processes
- Distribution regression

B — Surrogate modeling
- Gaussian processes
- Bayesian optimization

C — Learning to parameterize
- Variational inference
- Monte Carlo, Gibbs

D — Learning physics
- Sparse regression
- Latent force models

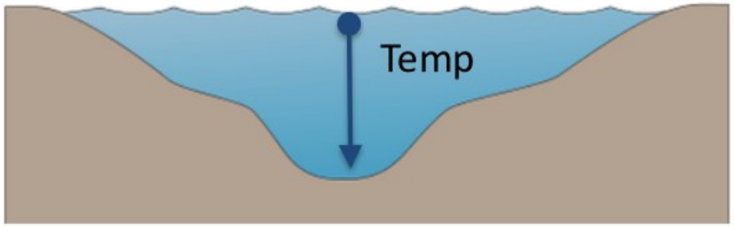
"Living in the Physics - Machine Learning Interplay for Earth Observation"
Camps-Valls et al. AAAI Fall Series 2020 Symposium on Physics-guided AI for Accelerating Scientific Discovery, 2020. arxiv.org/abs/2010.09031

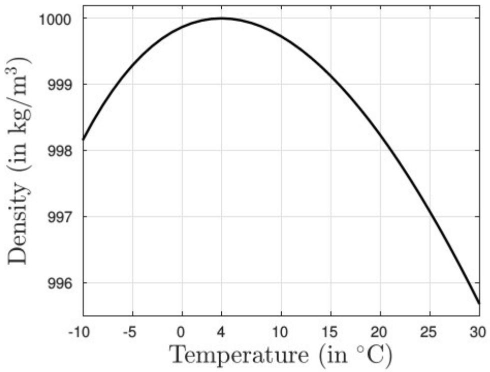
1- Constrained optimization

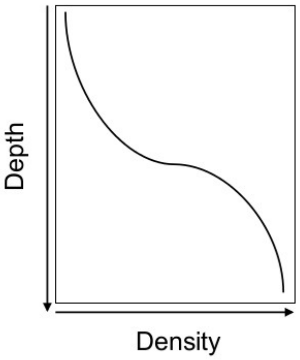
- ML minimizing model errors & violations of the physical laws

$$\text{PhysLoss} = \text{Cost}(y, \hat{y}) + \lambda_1 \|w\|_2^2 + \gamma \Omega(\hat{y}, \Phi)$$

$$\Omega(\hat{y}, \Phi) = \text{sum of physical violations of } \hat{y}$$







"Theory-guided Data Science", Karpatne, A. et al. IEEE Trans. Know. Data Eng., 2017.

1- Constrained optimization

- ML minimizing errors & predictions independent of sensitive factors

$$\mathrm{FairLoss} = \mathrm{Cost}(y, \hat{y}) + \lambda_1 \|w\|_2^2 + \gamma I(\hat{y}, s)$$

- Independence measured with HSIC

$$I := \mathrm{HSIC}(\mathcal{Y}, \mathcal{H}, \mathbb{P}_{\mathbf{ys}}) = \|\mathbf{C}_{ys}\|_{\mathrm{HS}}^2$$

- Closed form solution with kernels

$$\boldsymbol{\Lambda} = (\tilde{\mathbf{K}} + \lambda \mathbf{I} + \frac{\mu}{n^2} \tilde{\mathbf{K}} \tilde{\mathbf{K}}_S)^{-1} \mathbf{Y}$$

- Probabilistic interpretation with GPs:

$$f \sim \mathcal{GP}\left(0, k(\cdot, \cdot) - k_{\mathbf{X}\cdot}^{\top} (\mathbf{KHLH} + \delta^{-1} \mathbf{I})^{-1} \mathbf{HLH} k_{\mathbf{X}\cdot}\right)$$

"Fair Kernel Learning" Perez, Laparra, Gomez, Camps-Valls, G. ECML, 2017.

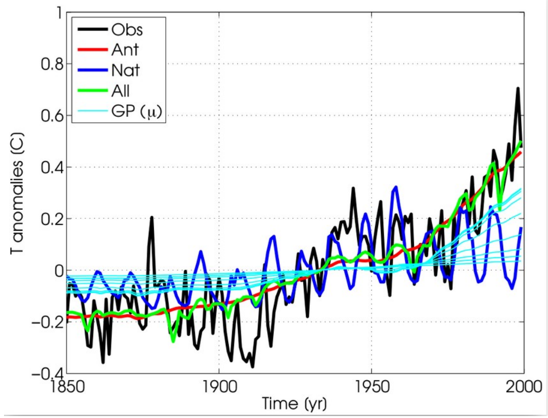
"Consistent Regression of Biophysical Parameters with Kernel Methods" Díaz, Peréz-Suay, Laparra, Camps-Valls, IGARSS 2018

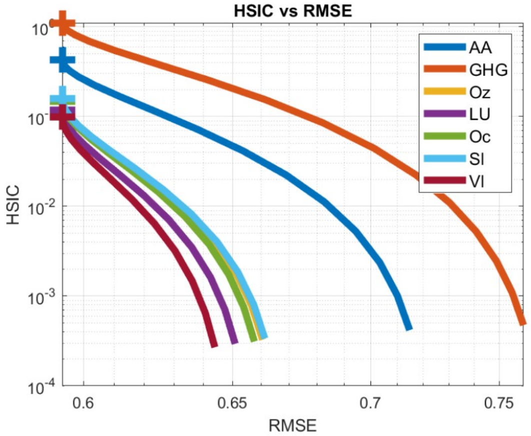
"Kernel Dependence Regularizers and Gaussian Processes with application to Algorithmic Fairness" Zhu Li, Perez-Suay, Camps-Valls and Sejdinovic, Submitted 2018

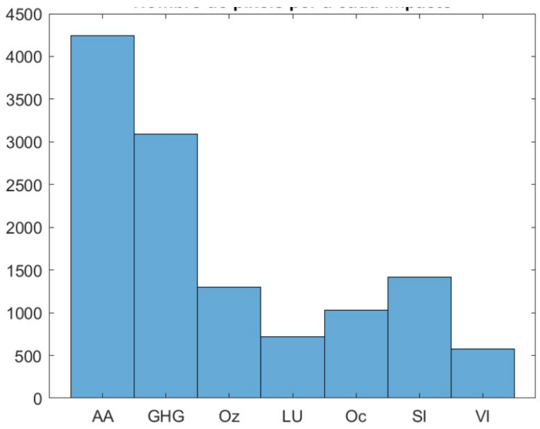
1- Constrained optimization

- ML minimizing errors & predictions independent of human factors

$$\mathrm{FairLoss} = \mathrm{Cost}(y, \hat{y}) + \lambda_1 \|w\|_2^2 + \gamma I(\hat{y}, s)$$







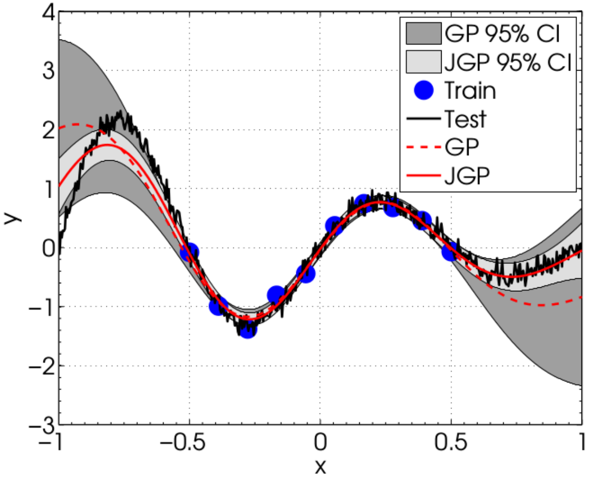
"Fair Kernel Learning" Perez, Laparra, Gomez, Camps-Valls, G. ECML, 2017.
"Consistent Regression of Biophysical Parameters with Kernel Methods" Díaz, Peréz-Suay, Laparra, Camps-Valls, IGARSS 2018
"Kernel Dependence Regularizers and Gaussian Processes with application to Algorithmic Fairness" Zhu Li, Perez-Suay, Camps-Valls and Sejdinovic, Submitted 2018

2- Blending observations and simulations for extrapolation

- Let ML talk to physical models for extrapolation

$$\text{JointLoss} = \text{Cost}(y, \hat{y}) + \lambda_1 \|w\|_2^2 + \gamma \Omega(\hat{y}, \Phi)$$

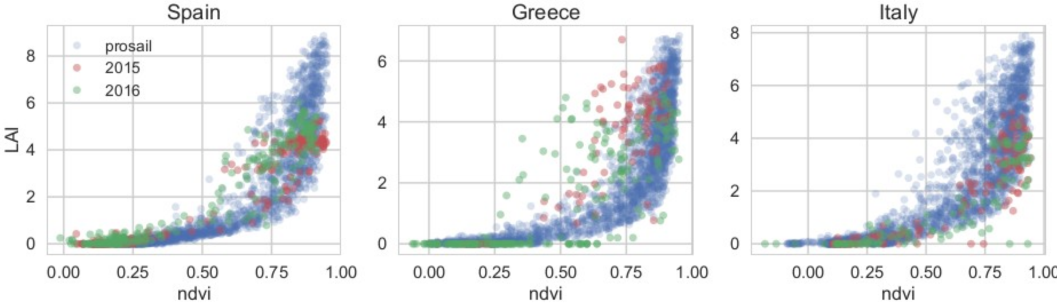
$$\Omega(\hat{y}, \Phi) = \text{Cost}_s(y_s, \hat{y}_s)$$



Setup

- ERMES project: 3 rice sites, 85% European production
- Landsat 8 + in situ measurements + PROSAIL simulations
- In situ LAI measurements: r = 70-300 (3 countries, 2 years)
- Simulations: s = 2000 (Landsat 8 spectra and LAI)

Filling the space ...

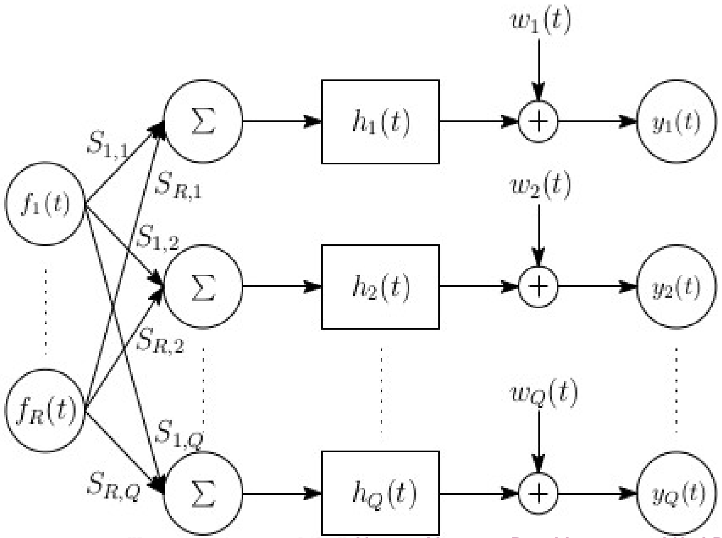


"**Joint Gaussian Processes for Biophysical Parameter Retrieval**" Svendsen, Martino, Camps-Valls, IEEE TGARS 2018

"**Physics-aware Gaussian processes in remote sensing**" Camps-Valls, G. et al. Applied Soft Computing, 2018.

3- Convolution processes & encoding ODEs

- Transfer learning across time and space



"Latent force models." Alvarez, Mauricio, David Luengo, and Neil D. Lawrence. Artificial Intelligence and Statistics. PMLR, 2009.
"Integrating Domain Knowledge in Data-driven Earth Observation with Process Convolutions" Svendsen, Muñoz, Piles, Camps-Valls, IEEE TGARS. 2021

1. **Latent forces** $f_r(t)$**:** zero-mean GPs with covariance function

$$k_{f_r f_r}(t' - t) \propto \exp\left(-\frac{(t'-t)^2}{2\ell_r^2}\right),$$

as vegetation should be smooth and exhibit local relations

2. **Coupling mechanism** $f_r(t) \leftrightarrow y_q(t)$**:** linear convolution operator with $h_q(t)$

$$h_q(t) \propto \exp\left(-\frac{t^2}{2\nu_q^2}\right)$$

Green's func. of heat diffusion eq.

as rate of change of $y \propto$ curvature of y

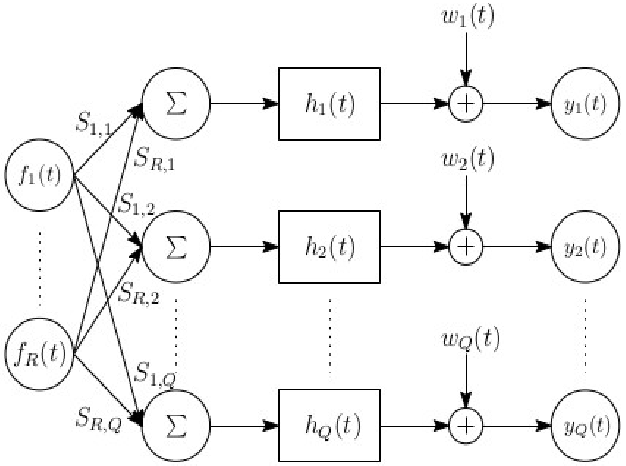
3. **Outputs** as lin. combination of pseudo-outputs plus AWGN:

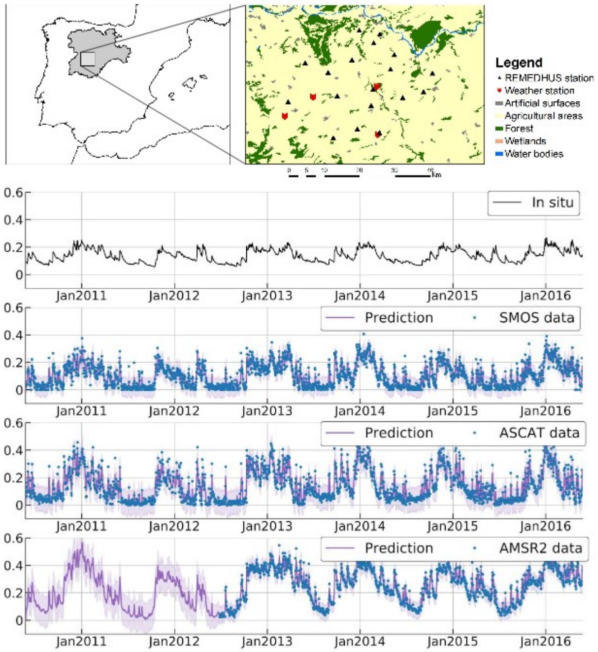
$$y_q(t) = \sum_{r=1}^{R} S_{rq} y_{rq}(t) + w_q(t), \quad w_q(t) \sim \mathcal{N}(0, \eta_q^2)$$

where S_{rq} accounts for the coupling strength

3- Convolution processes & encoding ODEs

- Encode ODEs governing the system + Learn latent forces driving it

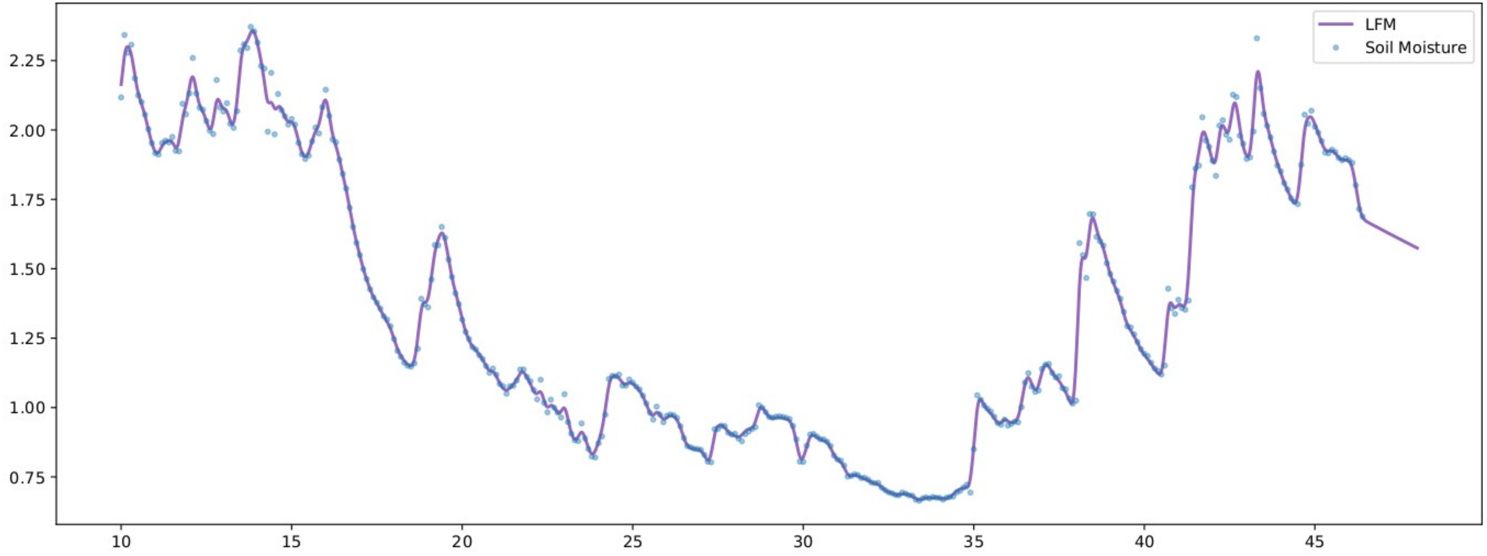




"Latent force models." Alvarez, Mauricio, David Luengo, and Neil D. Lawrence. Artificial Intelligence and Statistics. PMLR, 2009.
"Integrating Domain Knowledge in Data-driven Earth Observation with Process Convolutions" Svendsen, Muñoz, Piles, Camps-Valls, IEEE TGARS. 2021

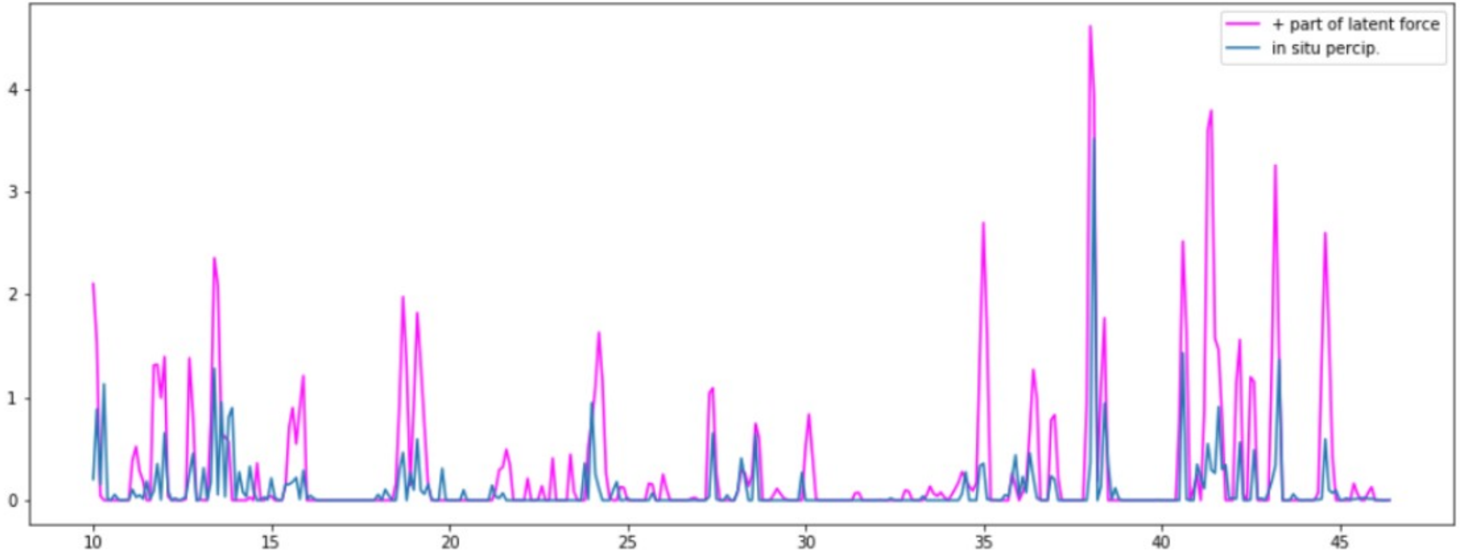
3- Convolution processes & encoding ODEs

- LFM-GP learns to estimate SM from uneven sampled time series ...



3- Convolution processes & encoding ODEs

- … and also learns driving forces, and one resembles precipitation
- … plus the time-decay constant of the ODE!



4- Hybrid modeling framework

ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Computers and Chemical Engineering 8 (2004) 7 5–766

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

Combining first principles modelling and artificial neural networks: a general framework

R. Oliveira*

REQUIMTE/CQFB, Centro de Química Fina e Biotecnologia, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, P-2829-516 Caparica, Portugal

$$\frac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} = \mathbf{KH}(\mathbf{c})\boldsymbol{\rho} - D\mathbf{c} + \boldsymbol{u} \tag{3a}$$

$$\boldsymbol{\rho} = \mathbf{N}(\mathbf{c}, \mathbf{W}) \tag{3b}$$

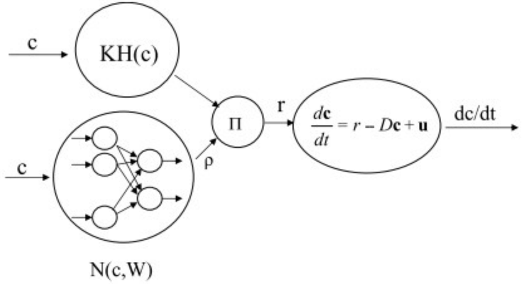
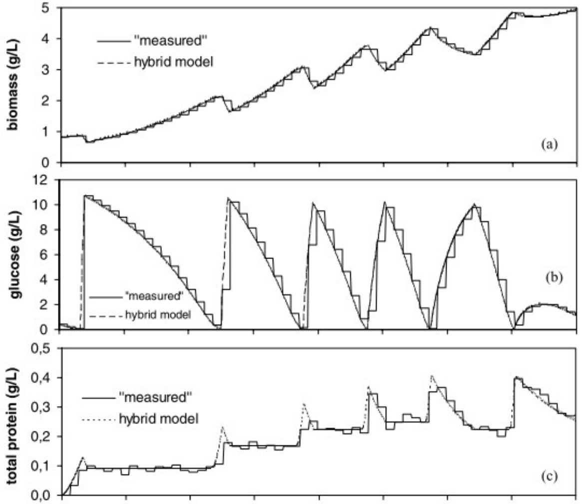


Fig. 1. General hybrid model structure for stirred tank bioreactors.



4- Hybrid machine learning

PERSPECTIVE

https://doi.org/10.1038/s41586-019-0912-1

Deep learning and process understanding for data-driven Earth system science

Markus Reichstein[1,2]*, Gustau Camps-Valls[3], Bjorn Stevens[4], Martin Jung[1], Joachim Denzler[2,5], Nuno Carvalhais[1,6] & Prabhat[7]

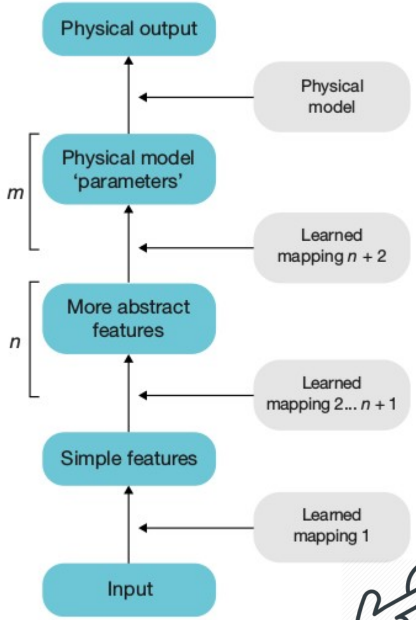
Machine learning approaches are increasingly used to extract patterns and insights from the ever-increasing stream of geospatial data, but current approaches may not be optimal when system behaviour is dominated by spatial or temporal context. Here, rather than amending classical machine learning, we argue that these contextual cues should be used as part of deep learning (an approach that is able to extract spatio-temporal features automatically) to gain further process understanding of Earth system science problems, improving the predictive ability of seasonal forecasting and modelling of long-range spatial connections across multiple timescales, for example. The next step will be a hybrid modelling approach, coupling physical process models with the versatility of data-driven machine learning.

“Deep learning and process understanding for data-driven Earth System Science”, Reichstein, Camps-Valls et al. Nature, 2019.

4- Hybrid machine learning

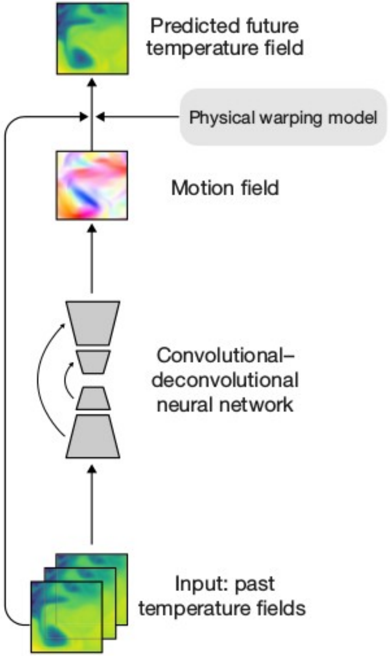
- ML that learns laws of physics (e.g. model-data consistency, mass and energy conservation)

A: “Physisizing” a deep learning architecture by adding one or several physical layers after the multilayer neural network



“Deep learning and process understanding for data-driven Earth System Science”
Reichstein, Camps-Valls et al. Nature, 2019.

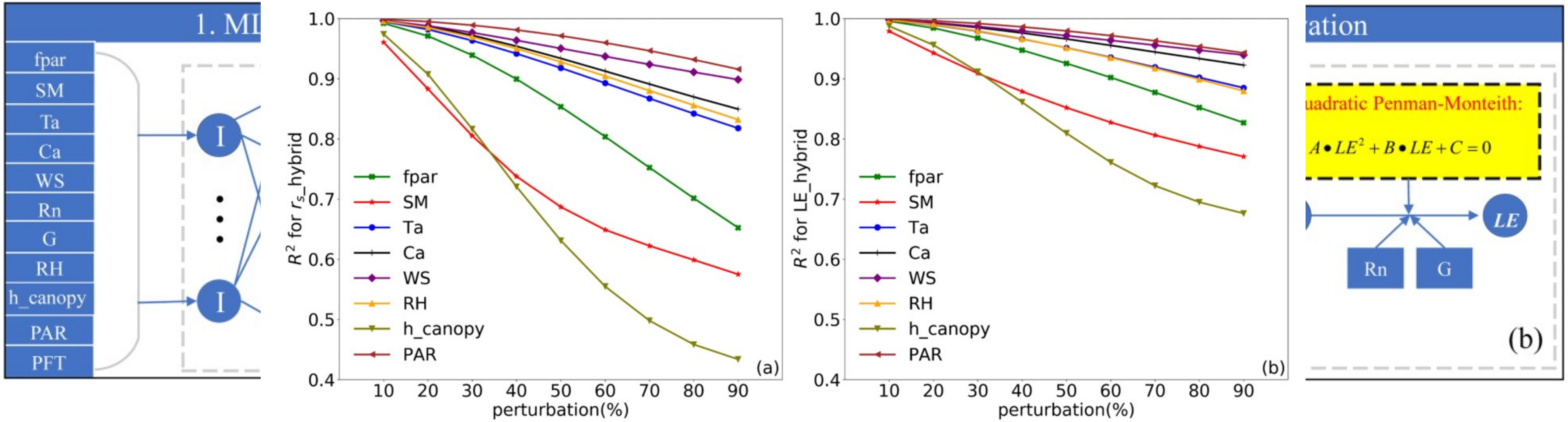
B: A motion field is learned with a convolutional-deconvolutional net, and the motion field is further processed with a physical model



“Deep Learning for Physical Processes: Incorporating Prior Scientific Knowledge”.
de Bezenac, Pajot, & Gallinari, arXiv:1711.07970 (2017).

4- Hybrid machine learning

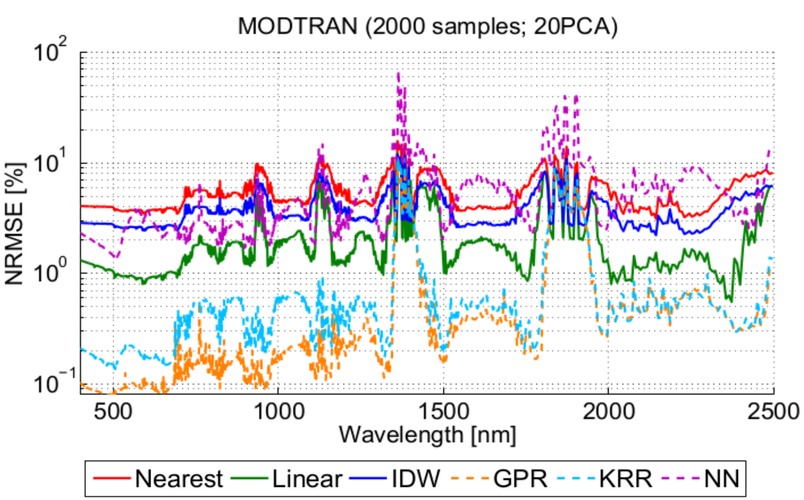
- ML that learns laws of physics (e.g. model-data consistency, mass and energy conservation)

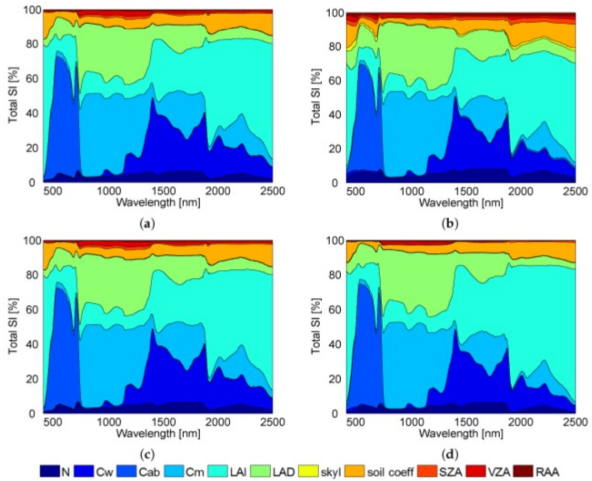


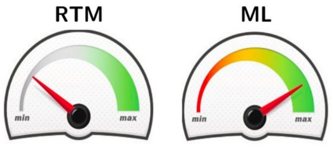
"Physics- constrained machine learning of evapotranspiration."
Zhao, Wen Li, et al. Geophysical Research Letters 46.24 (2019): 14496-14507.

5- Emulating complex codes with machine learning

- GP Emulation = Uncertainty quantification/propagation + Sensitivity analysis + Speed





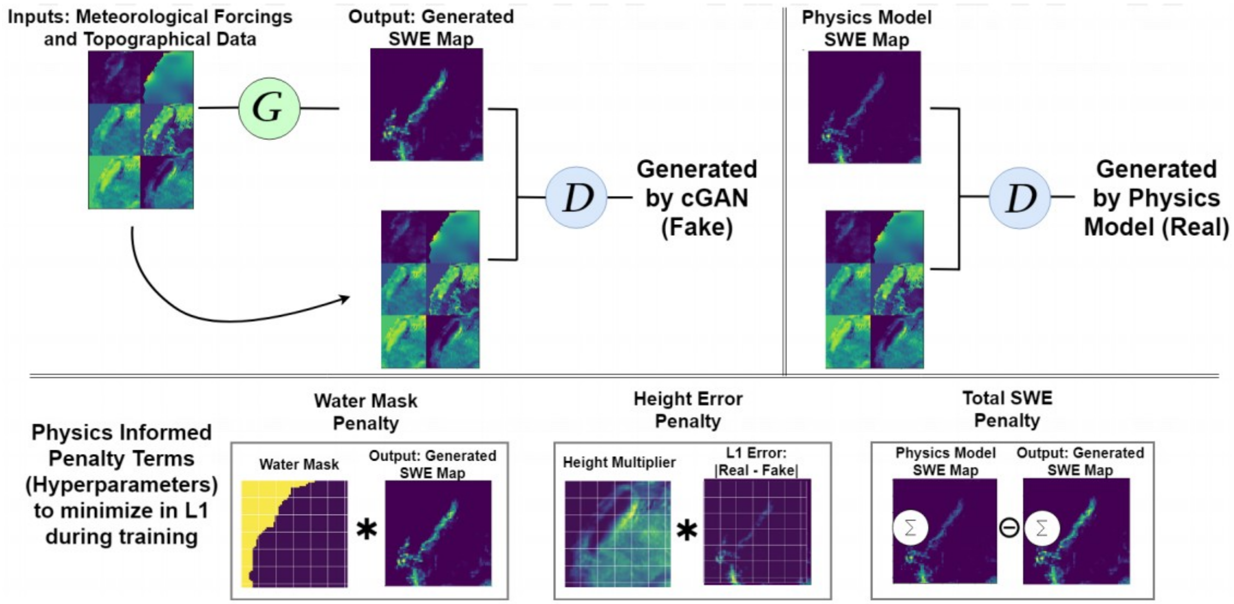


2.3–24.5 s/pix	0.1–1.3 ms/pix
0%	RMSE = 0.1 – 5%

“Emulation of Leaf, Canopy and Atmosphere Radiative Transfer Models for Fast Global Sensitivity Analysis”,
Verrelst, Camps-Valls et al Remote Sensing of Environment, 2016
“Emulation as an accurate alternative to interpolation in sampling radiative transfer codes”,
Vicent and Camps-Valls, IEEE Journal Sel. Topics Rem. Sens, Apps. 2018

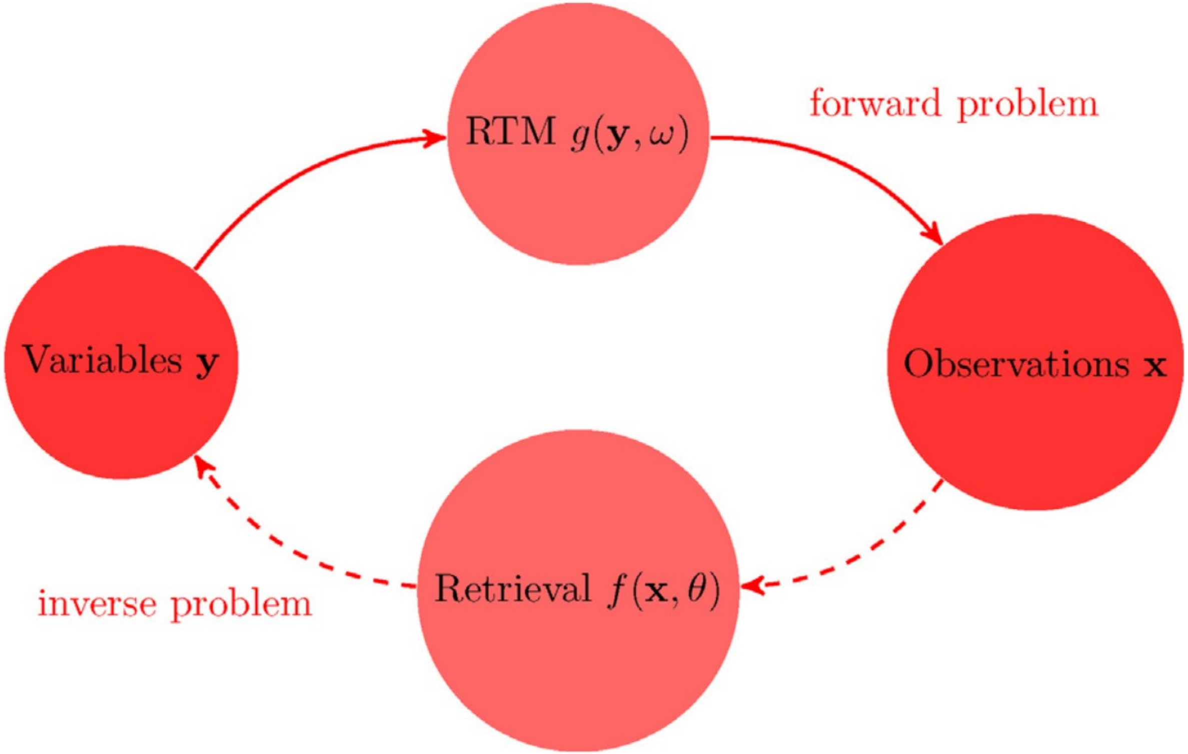
6- Going generative

- Estimate snow water equivalent (SWE)
- A conditional GAN with a physics-informed loss:
  - Higher elevation, more snow
  - Consistency with water mask
  - Penalize differences between cGAN and Phys Model



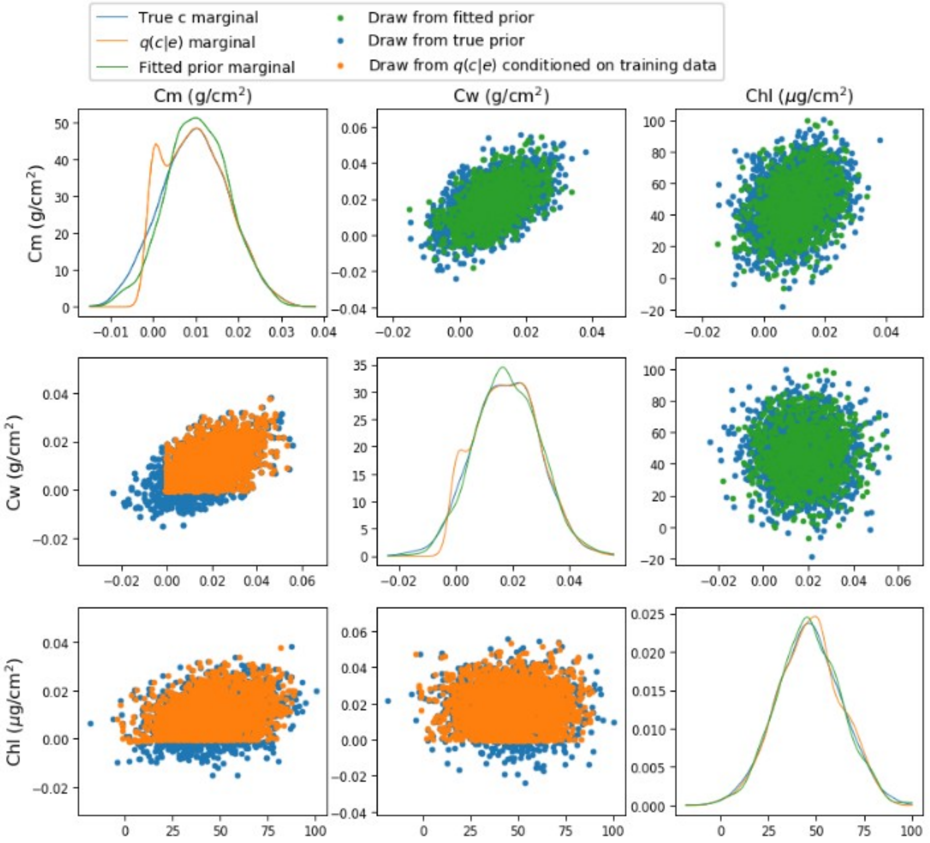
"Modelling mountain snowpack dynamics with cGANs", A. Albert et al, Arxiv, 2018

7- Parametrizations with variational inference



7- Parametrizations with variational inference

- An RTM is a deterministic model mapping parameters ('causes',*c*) to radiances ('effects', *E*) $P(E|c) = \mathcal{N}(E|RTM(c), \sigma \mathrm{I})$
- Assume a Gaussian prior $P(c) = \mathcal{N}(\mu_\phi, \Sigma_\phi)$
- The evidence/marginal likelihood is hard to integrate w/ RTM inside the Gaussian mean!
- VAE is orders of magnitude faster than MCMC, but problems with multimodal distributions

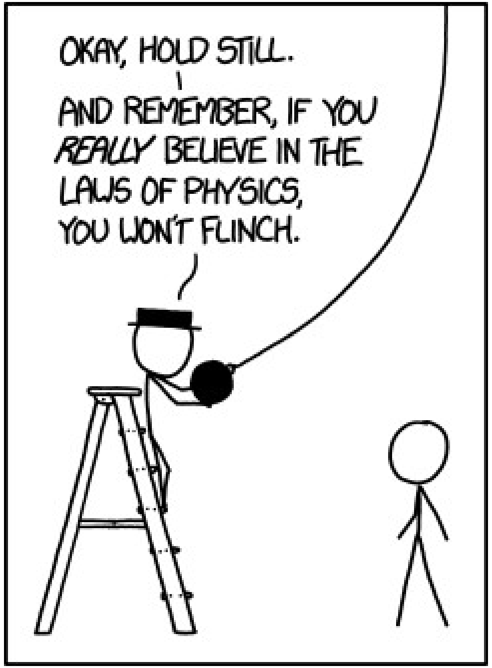


"Variational inference over radiative transfer model for biophysical parameter retrieval"
D. Svendsen, L. Martino, V. Laparra, G. Camps-Valls, *Machine Learning*, 2021

Conclusions

Take-home message

"AI is not deep learning, dude! Give Physics a Chance."



"Towards a Collective Agenda on AI for Earth Science Data Analysis"
Tuia, Roscher, Wegner, Jacobs, Zhu, and Camps-Valls, G. IEEE Geoscience and Remote Sensing Magazine 2021, arxiv.org/abs/2104.05107
"Living in the Physics - Machine Learning Interplay for Earth Observation"
Camps-Valls et al. AAAI Fall Series 2020 Symposium on Physics-guided AI for Accelerating Scientific Discovery, 2020. arxiv.org/abs/2010.09031