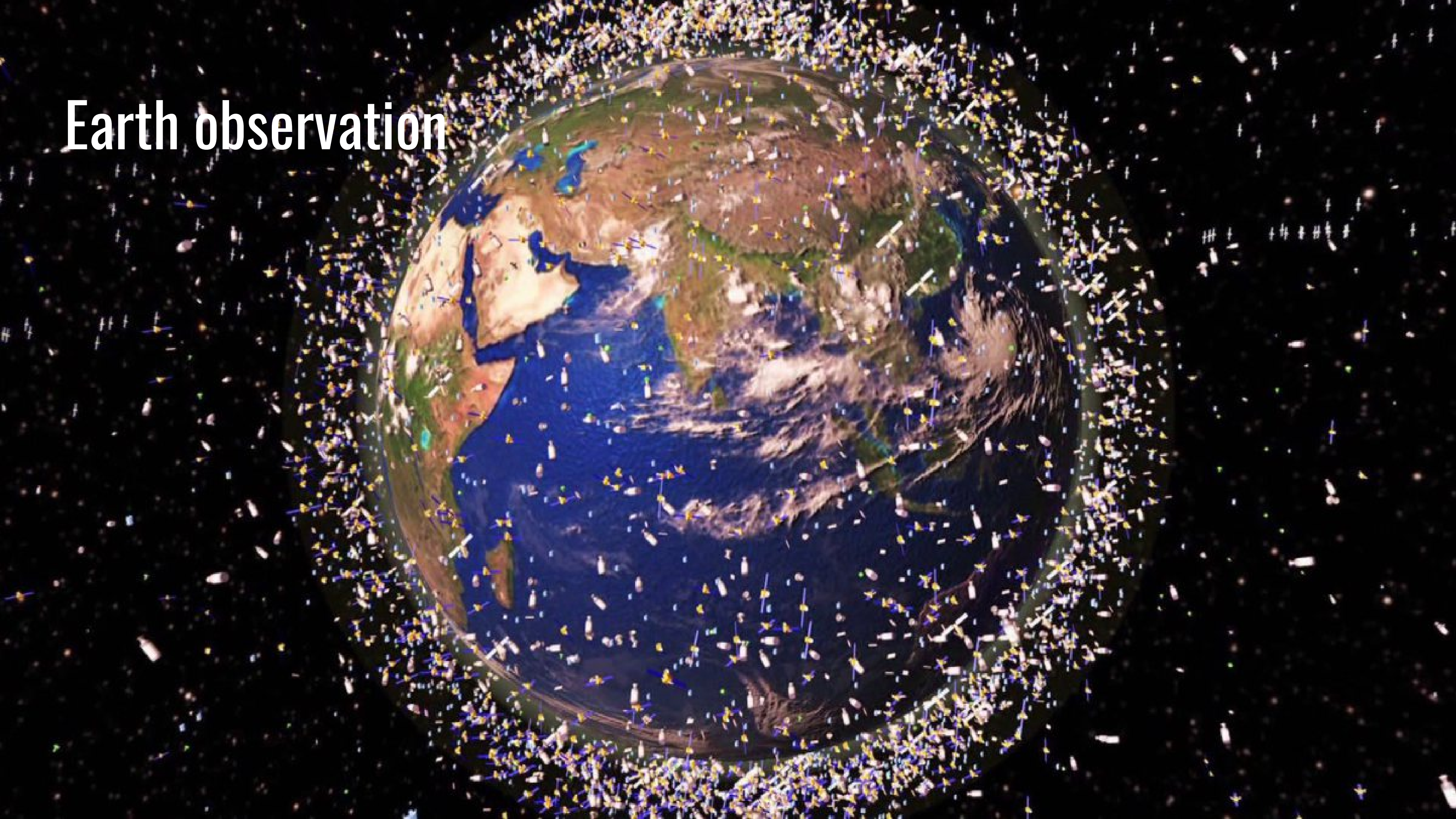
Machine Learning for Earth and Climate Sciences

Gustau Camps-Valls
Image Processing Laboratory
Universitat de València

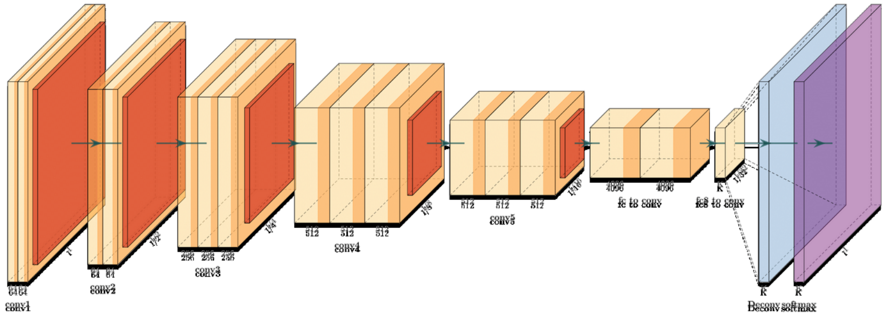
http://isp.uv.es @isp_uv_es gustau.camps@uv.es

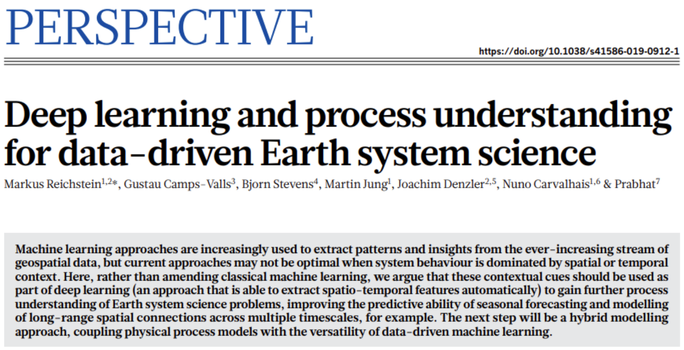
Earth science

Earth observation



Deep learning





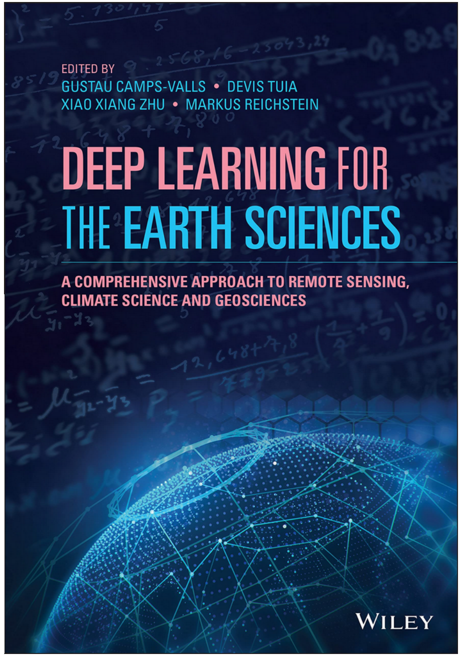
PERSPECTIVE

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

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

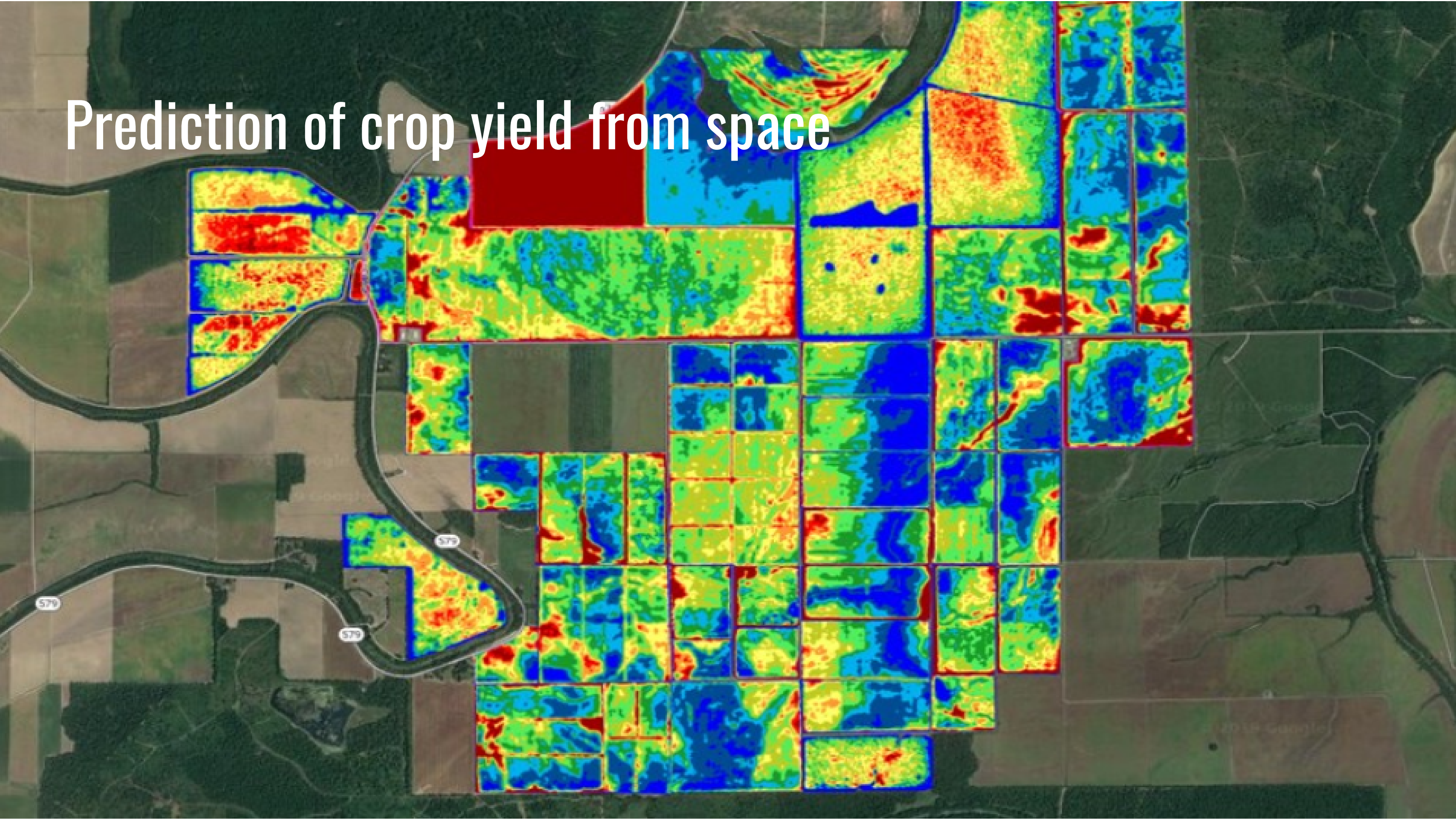
Markus Reichstein[1,2*], Gustau Camps-Valls[3], Bjorn Stevens[4], Martin Jung[1], Joachim Denzler[2,5], Nuno Carvalhais[1,6] & Prabhat[7]

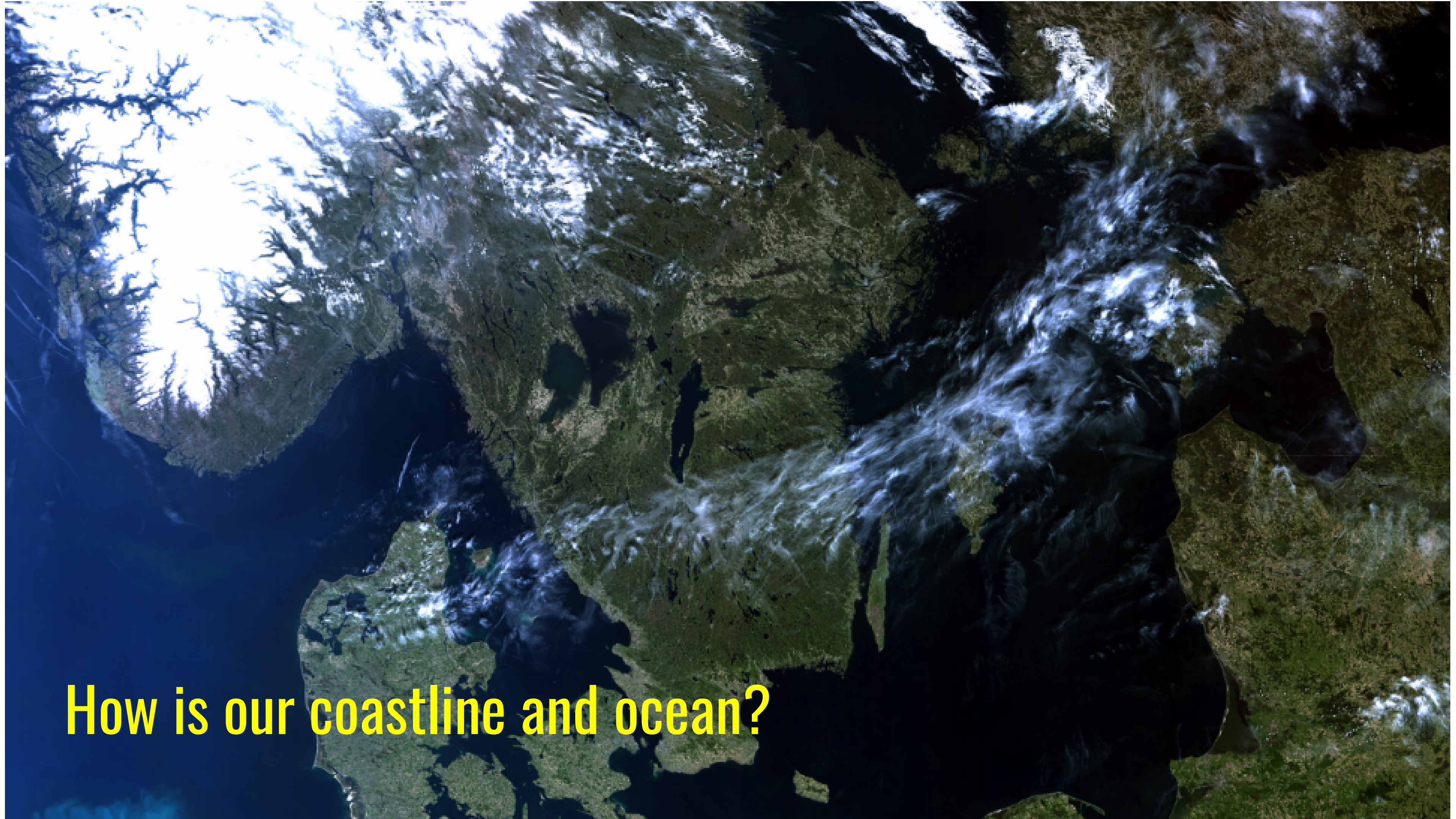
Machine learning approaches are increasingly used to extract patterns and insights from the ever-increasing stream of geospatial data, but current approaches may not be optimal when system behaviour is dominated by spatial or temporal context. Here, rather than amending classical machine learning, we 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.



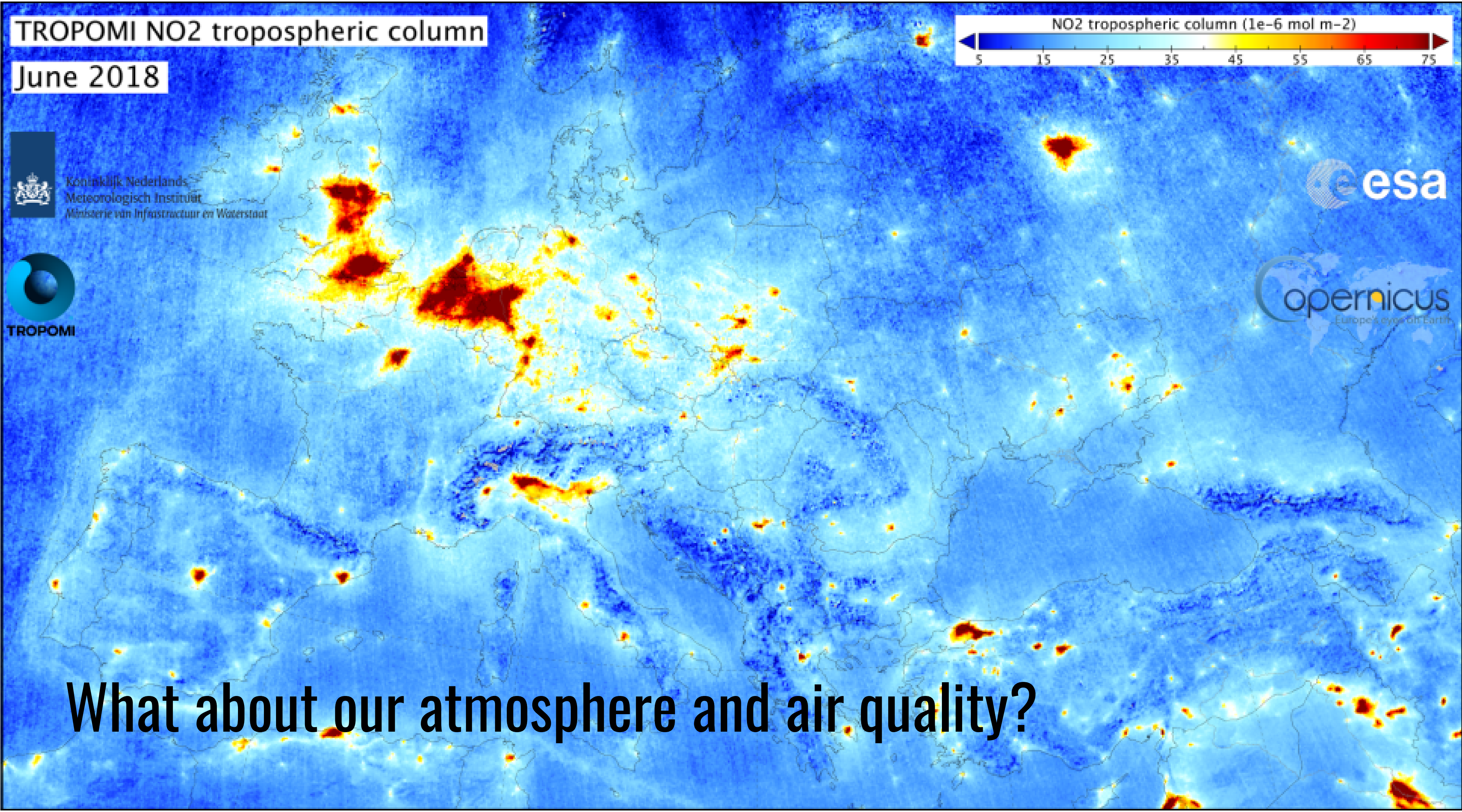
Reichstein, Camps-Valls et al, Nature, 2019
Camps-Valls, Tuia, Xiang, Reichstein. Wiley & Sons book, 2021

Prediction of crop yield from space



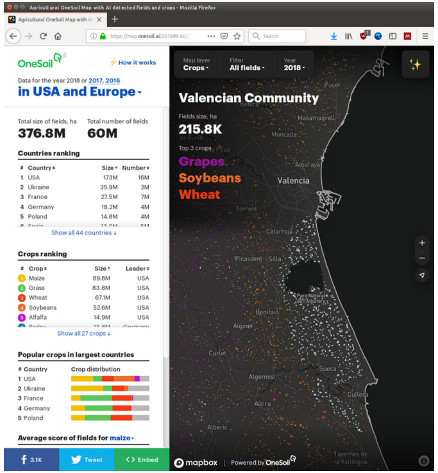

How is our coastline and ocean?

What about our atmosphere and air quality?

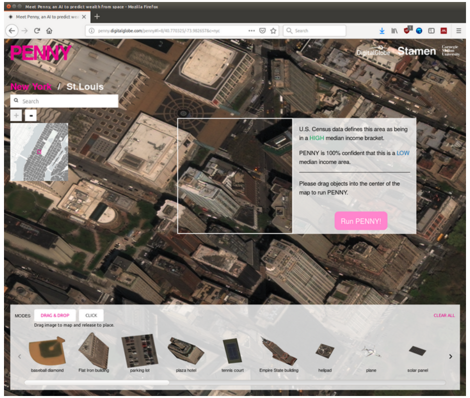


Some machine learning applications

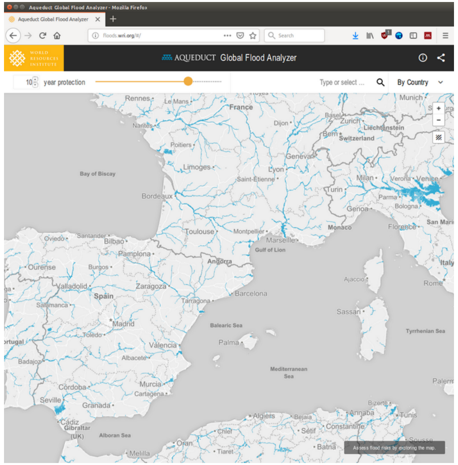
One soil map
https://map.onesoil.ai



Global wealth map
http://penny.digitalglobe.com



Flood analyzer
http://floods.wri.org

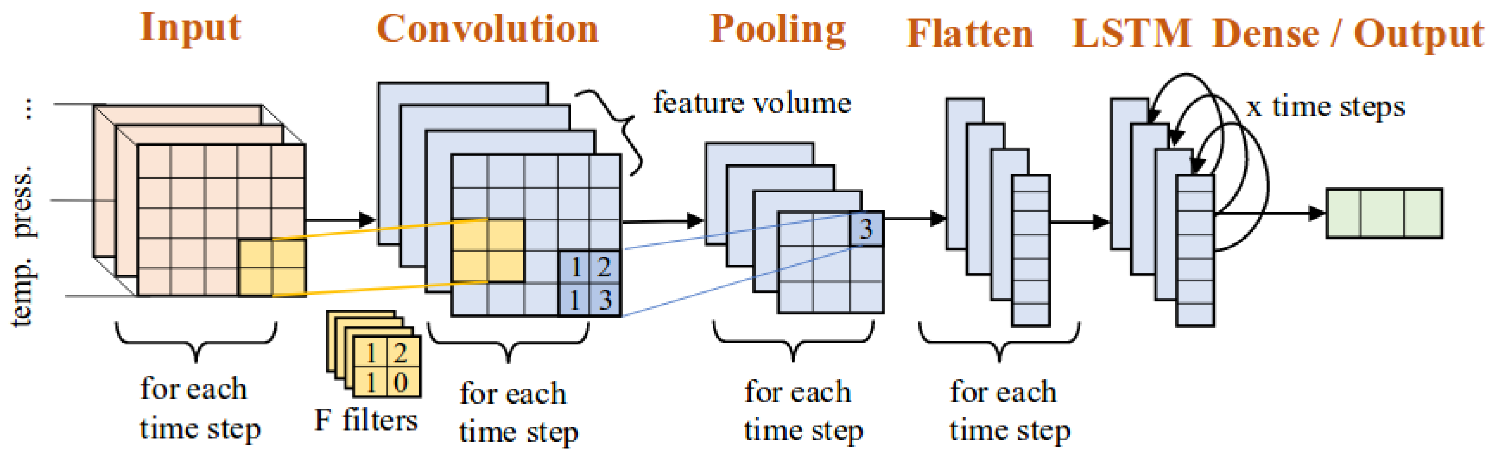


Disease mapping
https://www.healthmap.org

AI opportunities

Neural networks for spatio-temporal classification

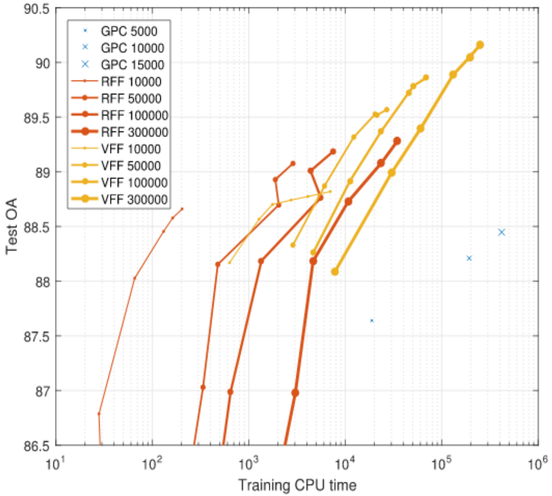
- Convolutional neural nets (CNN): hierarchical structure exploits spatial relations
- Long short-term memory (LSTM): recurrent network that accounts for memory/dynamics

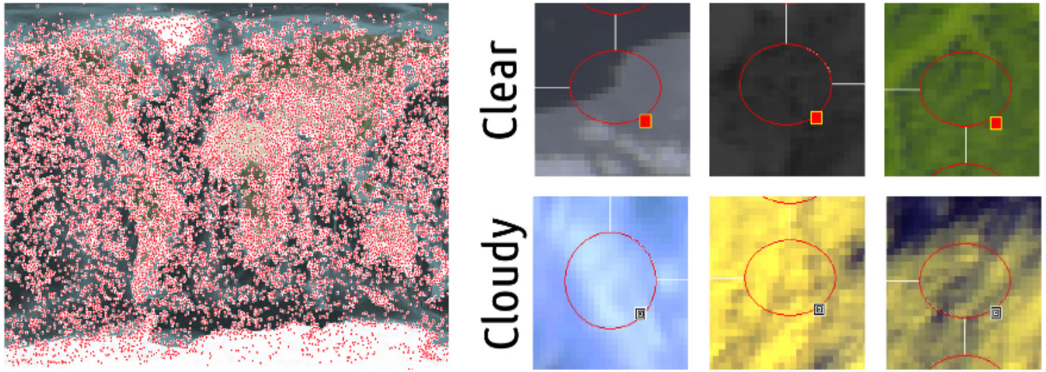


"A Deep Network Approach to Multitemporal Cloud Detection"
Tuia and Camps-Valls, IEEE IGARSS 2018, **http://isp.uv.es/code/landmarks.html**

Probabilistic and scalable classifiers

- Gaussian processes as an alternative to neural nets
- GPs allow a probabilistic treatment, confidence intervals, feature ranking, deep too!
- Gaussian processes start to be scalable ...

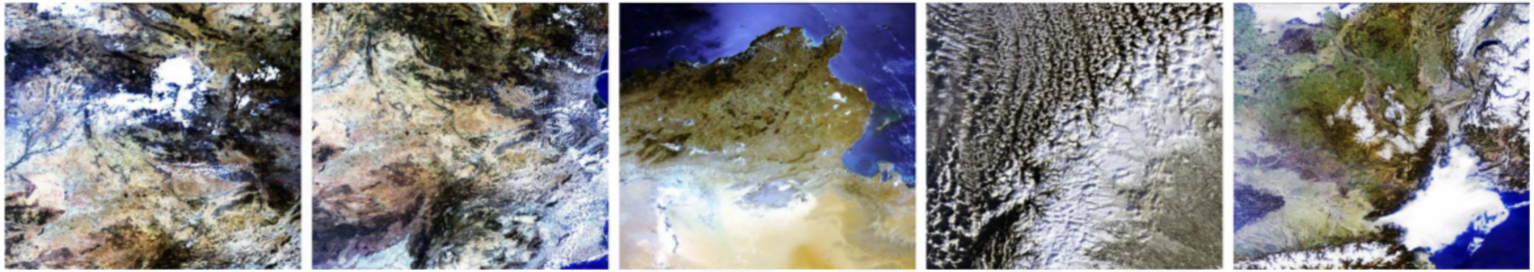


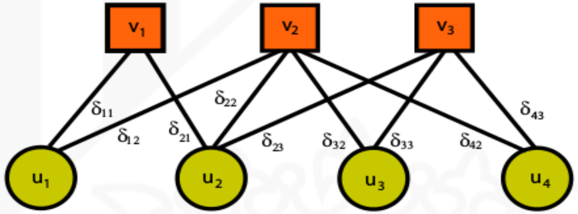


"Remote Sensing Image Classification With Large-Scale Variational Gaussian Processes,"
Morales, Molina and Camps-Valls, IEEE Trans. Geosc. Rem. Sens, 2018

Multitask learning

- Multiple inter-related outputs? Data from multiple sources?
- Learn to fuse heterogeneous information



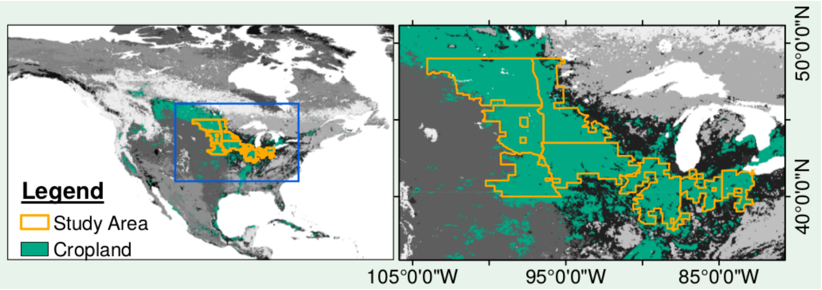


"Multitask Remote Sensing Data Classification"
Leiva and Camps-Valls, IEEE Trans. Geosc, Rem. Sens 2015

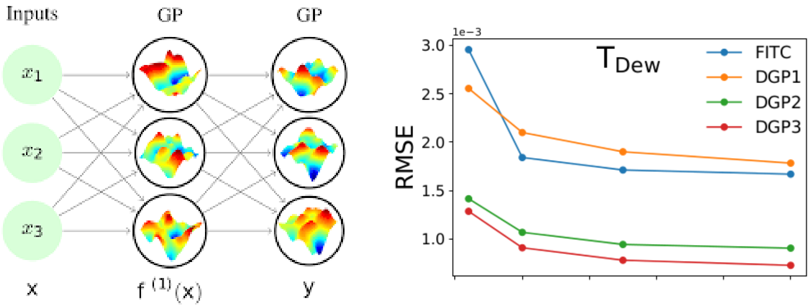
Spatio-temporal variable prediction

- **STA** is common place in climate informatics, neuroscience, video processing, NLP, ...
- **Current approaches:** CNN + LSTM, space-time Gaussian processes
- **Novel approaches:** distribution regression and variational deep GPs

$$\mathrm{P} \mapsto \mu_k(\mathcal{P}) \rightarrow \mathcal{P} \mapsto [\mathbb{E}\phi_1(X), \ldots, \mathbb{E}\phi_s(X)] \in \mathbb{R}^s$$
$$\langle \mu_k(\mathcal{P}), \mu_k(\mathcal{Q}) \rangle_{\mathcal{H}_k} = \mathbb{E}_{X\sim\mathcal{P}, Y\sim\mathcal{Q}} k(X, Y)$$



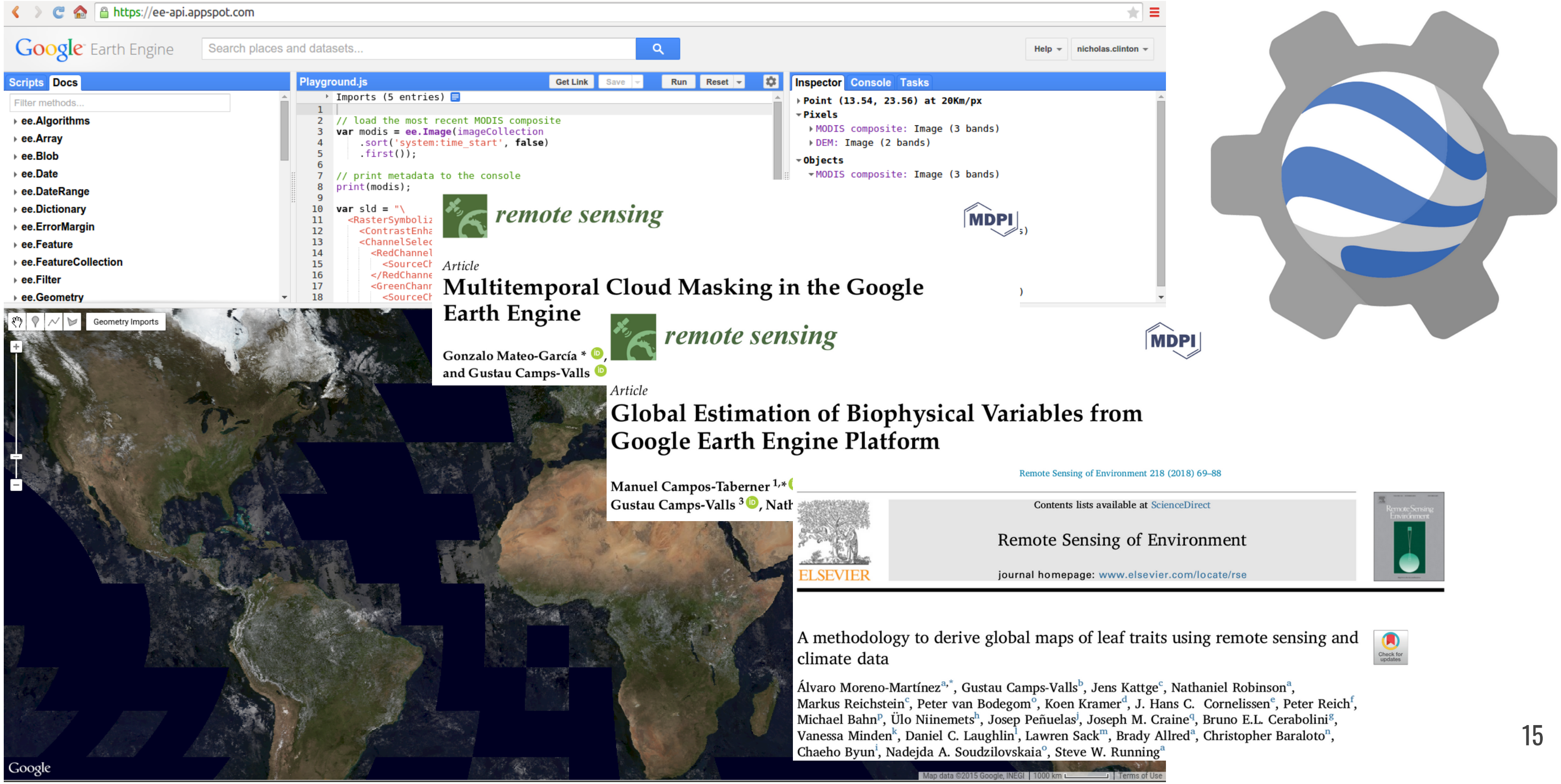
"Nonlinear Distribution Regression for Remote Sensing Applications"
Adsuara, Perez,Muñoz, Mateo, Piles, Camps-Valls, IEEE TGARS 2019



"A Survey on Gaussian Processes for Earth Observation Data Analysis"
Camps-Valls et al. IEEE Geoscience and Remote Sensing Magazine 2016

"Deep Gaussian Processes for Retrieval of bio-geo-physical parameters",
Svendsen, Ruescas and Camps-Valls, IEEE Trans. Geosc. Rem. Sens, 2019

Google Earth Engine



Potential risks and 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? Physics-aware ML
- What did the ML model learn? Explainable AI (XAI)
- Do they get cause-effect relations? Causal discovery and inference


THERE IS
light
AT THE END
of the
TUNNEL

XAI

Methods	HSI	ANN	Mechansim		
CAM with global average pooling [41], [90]	✗	✓	Decomposition	Saliency	Perceptive Interpretability
+ Grad-CAM [42] generalizes CAM, utilizing gradient	✓	✓			
+ Guided Grad-CAM and Feature Occlusion [67]	✗	✓			
+ Respond CAM [43]	✗	✓			
+ Multi-layer CAM [91]	✗	✓			
LRP (Layer-wise Relevance Propagation) [13], [52]	✗	N.A.			
+ Image classifications. PASCAL VOC 2009 etc [44]	✗	✓			
+ Audio classification. AudioMNIST [46]	✗	✓			
+ LRP on DeepLight. fMRI data from Human Connectome Project [47]	✗	✓			
+ LRP on CNN and on BoW(bag of words)/SVM [48]	✗	✓			
+ LRP on compressed domain action recognition algorithm [49]	✗	✗			
+ LRP on video deep learning, *selective relevance method* [51]	✗	✓			
+ BiLRP [50]	✗	✓			
DeepLIFT [56]	✗	✓			
Prediction Difference Analysis [57]	✗	✓			
Slot Activation Vectors [40]	✗	✓			
PRM (Peak Response Mapping) [58]	✗	✓			
LIME (Local Interpretable Model-agnostic Explanations) [14]	✓	✓	Sensitivity		
+ MUSE with LIME [84]	✓	✓			
+ Guidelinebased Additive eXplanation optimizes complexity, similar to LIME [92]	✓	✓			
# Also listed elsewhere: [55], [68], [70], [93]	N.A.	N.A.			
Others. Also listed elsewhere: [94]	N.A.	N.A.	Others		
+ Direct output labels. Training NN via multiple instance learning [64]	✗	✓			
+ Image corruption and testing Region of Interest statistically [65]	✗	✓			
+ Attention map with autofocus convolutional layer [66]	✗	✓			
DeconvNet [71]	✗	✓	Inversion	Signal	
Inverting representation with natural image prior [72]	✗	✓			
Inversion using CNN [73]	✗	✓			
Guided backpropagation [74], [90]	✗	✓			
Activation maximization/optimization [37]	✗	✓	Optimization		
+ Activation maximization on DBN (Deep Belief Network) [75]	✗	✓			
+ Activation maximization, multifaceted feature visualization [76]	✗	✓			
Visualization via regularized optimization [77]	✗	✓			
Semantic dictionary [38]	✗	✓			
Decision trees	N.A.	N.A.	Verbal		
Propositional logic, rule-based [81]	✗	✗			
Sparse decision list [82]	✗	✗			
Decision sets, rule sets [83], [84]	✓	✗			
Encoder-generator framework [85]	✗	✓			
Filter Attribute Probability Density Function [86]	✗	✗			
MUSE (Model Understanding through Subspace Explanations) [84]	✓	✓			

“A Survey on Explainable Artificial Intelligence(XAI): towards Medical XAI”, Tjoa 2019

“Advancing Deep Learning For Earth Sciences: From Hybrid Modeling To Interpretability”, Camps-Valls, G. and Reichstein, M. and Zhu, Z. and Tuia, D. IEEE IGARSS (2020)

Physics-aware ML

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

PERSPECTIVE

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

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

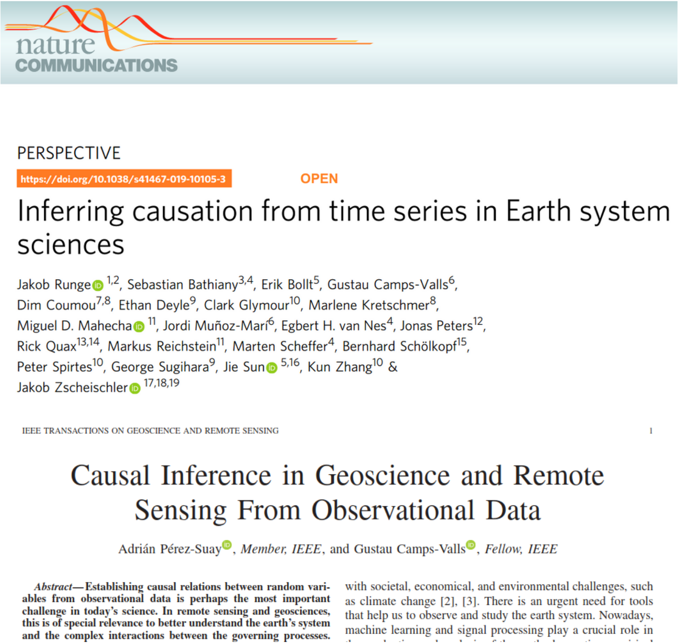
Markus Reichstein[1,2]*, Gustau Camps-Valls[3], Bjorn Stevens[4], Martin Jung[1], Joachim Denzler[2,5], Nuno Carvalhais[1,6] & Prabhat[7]

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.

"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

Causality



nature COMMUNICATIONS

PERSPECTIVE

https://doi.org/10.1038/s41467-019-10105-3 OPEN

Inferring causation from time series in Earth system sciences

Jakob Runge[1,2], Sebastian Bathiany[3,4], Erik Bollt[5], Gustau Camps-Valls[6], Dim Coumou[7,8], Ethan Deyle[9], Clark Glymour[10], Marlene Kretschmer[8], Miguel D. Mahecha[11], Jordi Muñoz-Marí[6], Egbert H. van Nes[4], Jonas Peters[12], Rick Quax[13,14], Markus Reichstein[11], Marten Scheffer[4], Bernhard Schölkopf[15], Peter Spirtes[10], George Sugihara[9], Jie Sun[5,16], Kun Zhang[10] & Jakob Zscheischler[17,18,19]

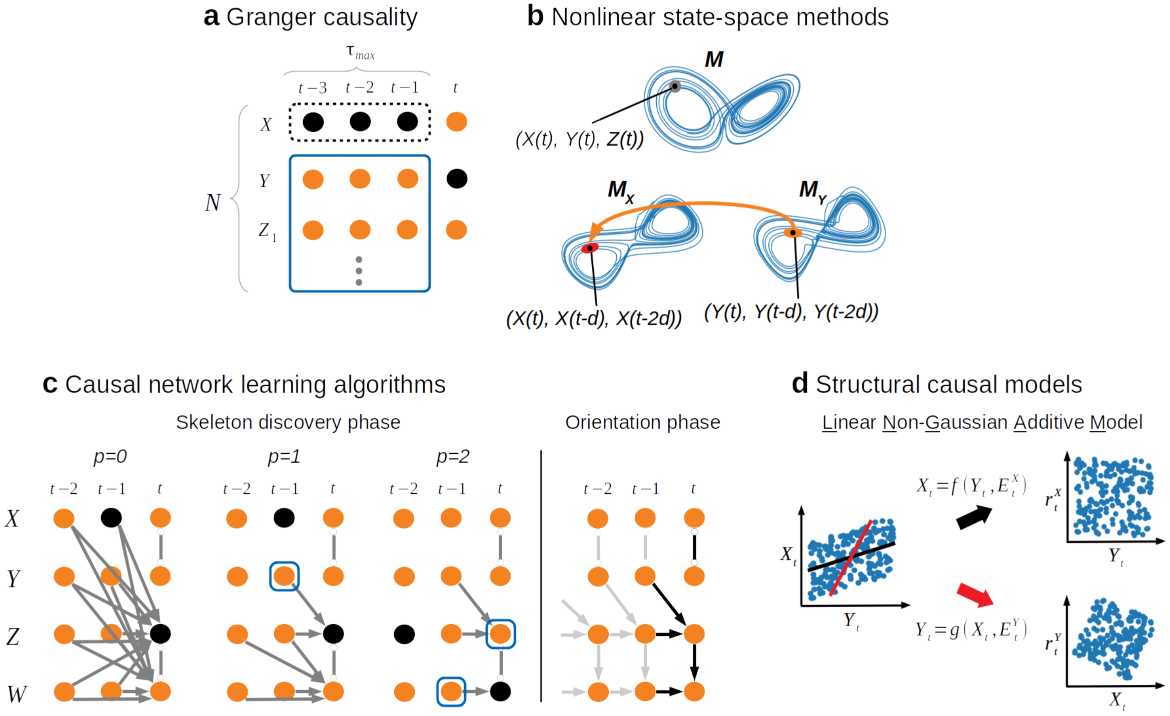
IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING 1

Causal Inference in Geoscience and Remote Sensing From Observational Data

Adrián Pérez-Suay, *Member, IEEE*, and Gustau Camps-Valls, *Fellow, IEEE*

Abstract—Establishing causal relations between random variables from observational data is perhaps the most important challenge in today's science. In remote sensing and geosciences, this is of special relevance to better understand the earth's system and the complex interactions between the governing processes.

with societal, economical, and environmental challenges, such as climate change [2], [3]. There is an urgent need for tools that help us to observe and study the earth system. Nowadays, machine learning and signal processing play a crucial role in



"Inferring causation from time series with perspectives in Earth system sciences", Runge, Bathiany, Bollt, Camps-Valls, et al. Nat Comm., 2019

"Causal Inference in Geoscience and Remote Sensing from Observational Data," Pérez-Suay and Camps-Valls, IEEE Trans. Geosc. Rem. Sens, 2018

"CauseMe: An online system for benchmarking causal inference methods," Muñoz-Marí, Mateo, Runge, Camps-Valls. In preparation (2019). **CauseMe: http://causeme.uv.es**

ELLIS, ELISE and AIDA

ELLIS.eu

- ELLIS: European Laboratory for Learning and Intelligent Systems
- Distributed center of excellence in AI
- ELLIS goals and mandates:
  - Coordinates research, fosters collaborations with industry and users, promotes technology adoption, support PhD fellowships & postdoc visits/stays, organize focused workshops
- ELLIS legacy:
  - place Europe in the global map of top AI research and transfer to industry
  - Make Europe an international talent magnet, incubator of innovation, and ecosystem
- Largely multidisciplinary & organized in Research Programs:
  - Robotics, robustness, health, language processing, earth, etc.

https://ellis.eu

ELLIS.eu & ELISE-AI

elise
European Network of AI Excellence Centres

ellis

- **ELLIS is Largely multidisciplinary & organized in Research Programs:**
  - Robotics, robustness, health, language processing, etc.
- **“Machine learning for Earth and climate sciences”** (Dir.: Gustau Camps & Markus Reichstein)
  - **Goal**: Model & understand the Earth with Machine Learning and Process Understanding
    - Spatio-temporal anomaly and extreme events detection, anticipation and attribution
    - Data-driven dynamic modelling and forecasting
    - Hybrid modeling: linking physics and machine learning models
    - Causal inference, Learning and explaining feature representations
    - Earth and Climate model emulation, generative modelling and data-model fusion
  - 20 ELLIS Fellow members, regular meetings/workshops, exchange students, teaching material
- Universitat de València is a core center in the ICT-48 project **ELISE**

https://www.elise-ai.eu/

AIDA - AI Doctoral Academy

- All ICT-48 networks (AI4Media, ELISE, HumanE-AI NET, TAILOR) + VISION consortium joined forces
- New joint instrument to support world-level AI education and research programme
- AIDA offers:
  - access to resources, knowledge & expertise for the latest developments and trends on AI research and innovation
  - operate as an umbrella organisation for AI PhD and Postdoc studies
  - support a new generation of alumni
  - enhance networking and collaboration among talented researchers
  - provide access to workshops/conferences, short courses, mobilities, training/job opportunities

https://www.i-aida.org/

Conclusions

Take-home message

- **AI is a paradigm shift**
  - Excel in classification, (change) detection, parameter retrieval
  - Automate & understand processes
- **Challenges:** interpretability + causal relations + physics consistency
- **Future:**
  - User-centric AI + trustworthiness + accountability
  - Holistic & interdisciplinary education

"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