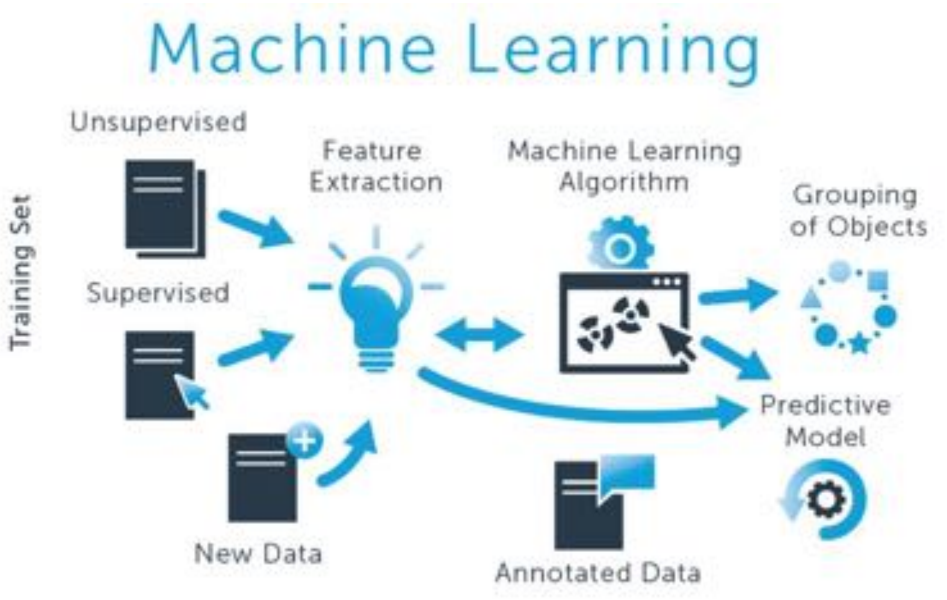
Physics-aware and Explainable Machine Learning

Gustau Camps-Valls, Luis Gómez-Chova, Daniel Svendsen
Diego Bueso, Luca Martino, Adrian Pérez-Suay
María Piles, Valero Laparra, Ana B. Ruescas

Image Processing Lab (IPL)
Universitat de València – **http://isp.uv.es**

Standard machine learning

- **Supervised learning**
  - Classification
  - Regression & model inversion
  - Anomaly/target detection
- **Unsupervised learning**
  - Density estimation
  - Dimensionality reduction
  - Clustering



Standard supervised machine learning

$$F(X) = y$$

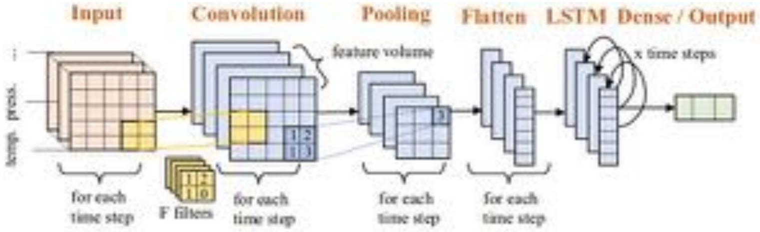
- X: observations, independent covariates
- Y: target, dependent variable
- F: machine learning model (nonlinear, nonparametric, flexible, learned from data)

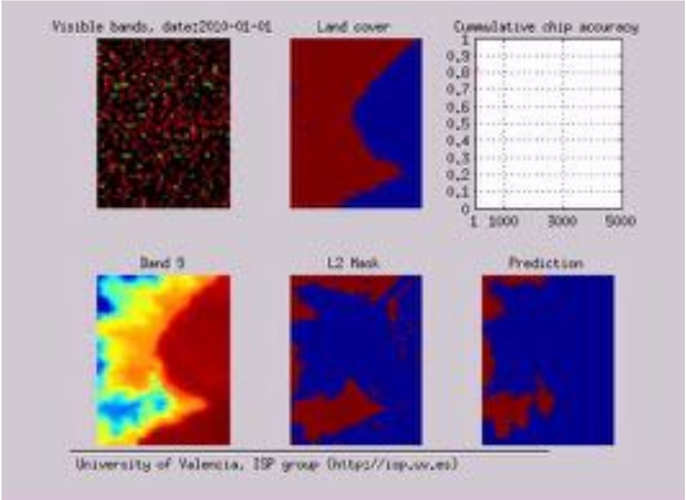
#1 - Spatio-temporal image classification

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


Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting

Abstract

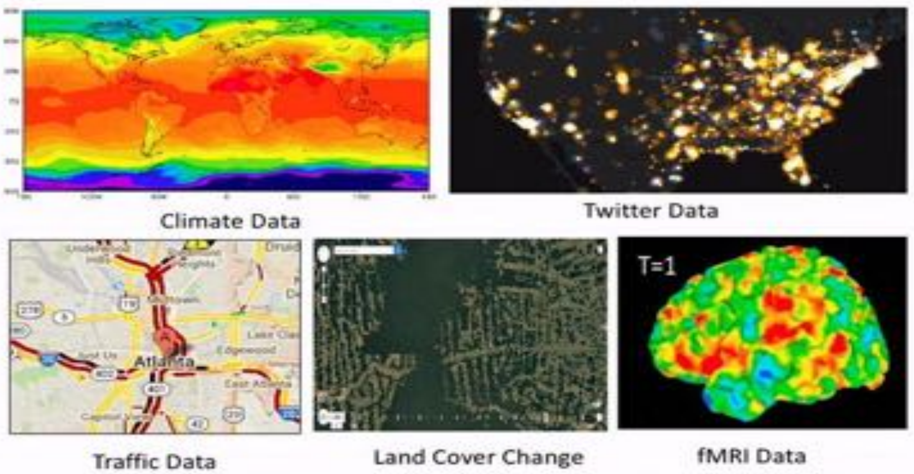




"A Deep Network Approach to Multitemporal Cloud Detection"
Tuia, Perez-Suay and Camps-Valls, IEEE IGARSS 2018, http://isp.uv.es/code/landmarks.html

#2- Spatio-temporal variable prediction

- **STA** is common place in climate informatics, neuroscience, video processing, NLP, ...
- **Current approaches**
  - CNN + LSTM
  - Space-time (deep) GPs
  - Combine CNN+GPs
- **Many applications in Remote Sensing**
  - Parameter retrieval / estimation
  - Time series gap filling
  - Sensor fusion / Data assimilation



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

"Statistical Retrieval of Atmospheric Profiles with Deep Convolutional Neural Networks", Malmgren-Hansen, D. and Laparra, V. and Camps-Valls, G., IEEE TGARS, 2018

Physics-aware machine learning

F(X, ) = y

#1 - Physics-driven ML: constrained optimization

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

<u>Minimize model violations</u>

$$\text{Loss} = \text{Cost}(y, \hat{y}) + \lambda_1 \|w\|_2^2$$

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

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

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

<u>Fair ML</u>

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

"Fair Kernel Learning" Perez-Suay, Laparra, Gomez-Chova, Camps-Valls, G. et al. ECML, 2017.

<u>Joint Model-Data ML</u>

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

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

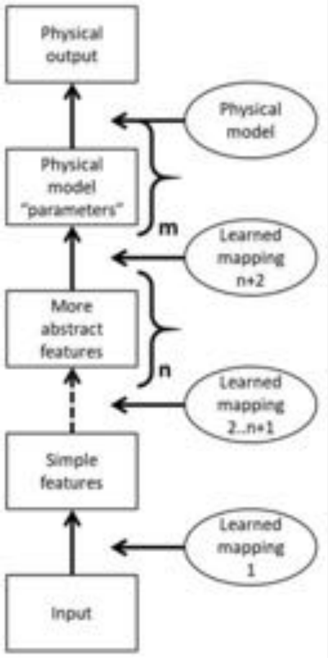
"Physics-aware Gaussian processes in remote sensing" Camps-Valls, G. et al. Applied Soft Computing, 2018.

#2 - Physics-driven ML: hybrid modeling framework

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

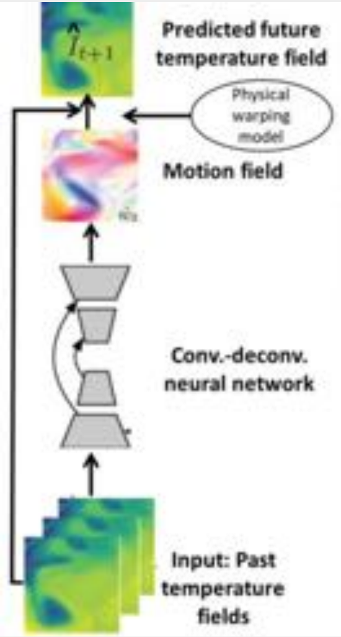
A: "Physisizing" a deep learning architecture by adding one or several physical layers after the multilayer neural network

"Deep learning and process understanding for data-driven Earth System Science"
Reichstein, Camps-Valls et al. Nature, 2018.



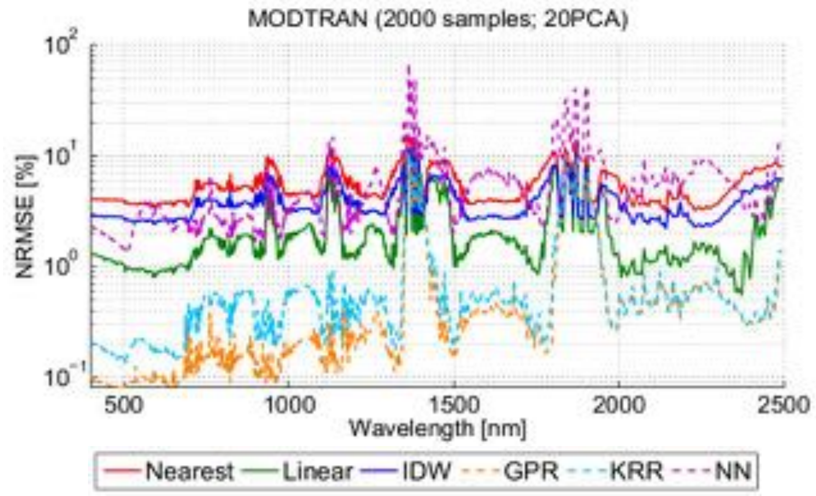
B: A motion field is learned with a convolutional-deconvolutional net, and the motion field is further processed with a physical model

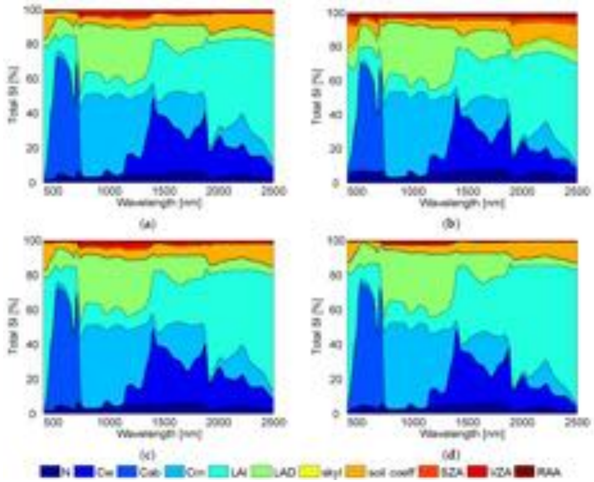
"Deep Learning for Physical Processes: Incorporating Prior Scientific Knowledge".
de Bezenac, Pajot, & Gallinari, arXiv:1711.07970 (2017).

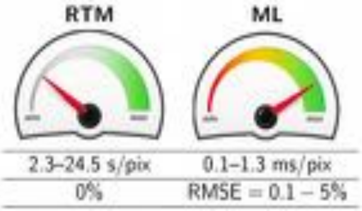


#3- Physics-driven ML: emulation of complex codes

- GP Emulation = Mathematical tractability + Global sensitivity analysis + Speed





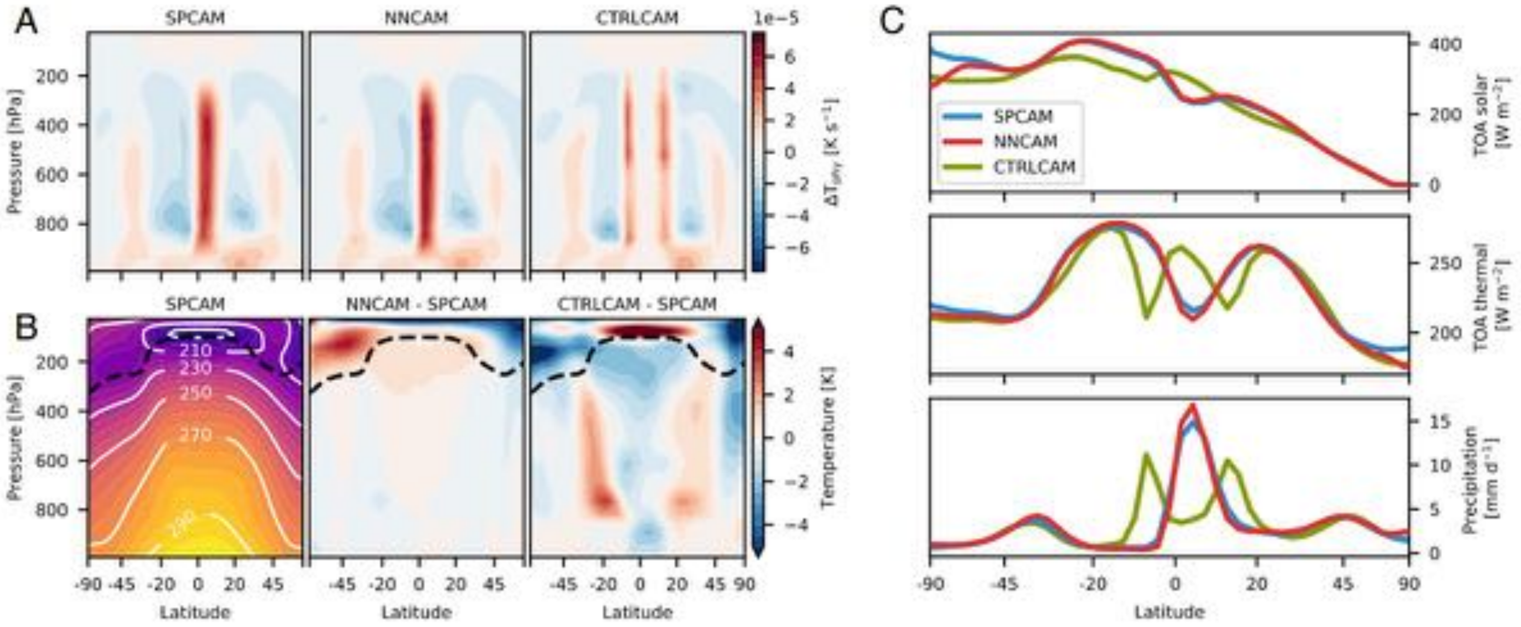


"Emulation of Leaf, Canopy and Atmosphere Radiative Transfer Models for Fast Global Sensitivity Analysis", Verrelst, Camps-Valls et al Remote Sensing of Environment, 2016

"Emulation as an accurate alternative to interpolation in sampling radiative transfer codes", Vicent and Camps-Valls, IEEE Journal Sel. Topics Rem. Sens, Apps. 2018

#3- Physics-driven ML: emulation of superparam. models

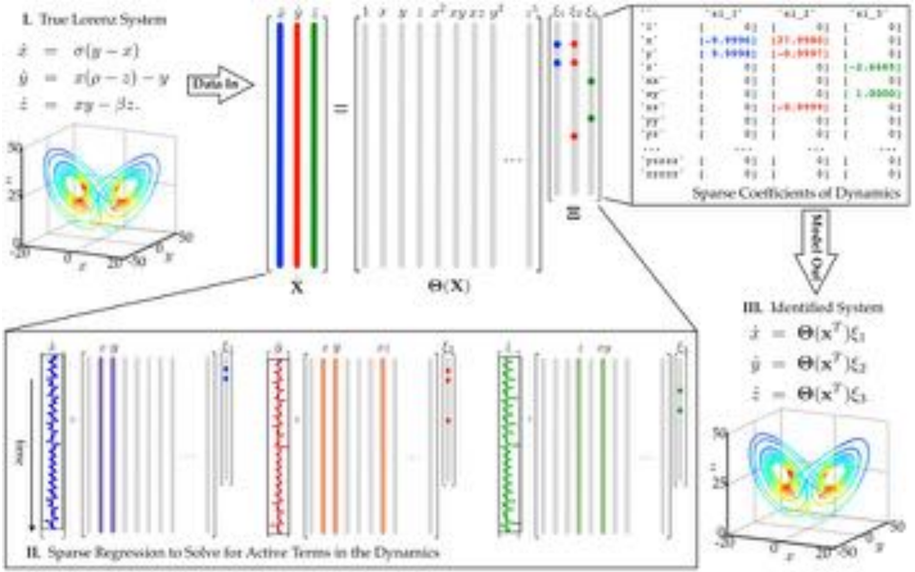
- **NN Emulation of a superparameterized model**: no virtual error, 10x faster



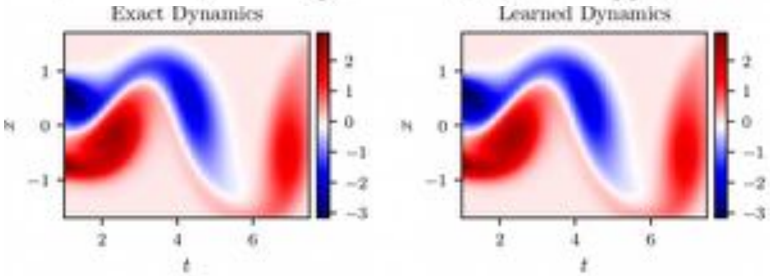
“Deep learning to represent subgrid processes in climate models” Rasp, Pritchard, Pierre Gentine, PNAS 2018

#4- Physics-driven ML: encoding and learning ODE/PDEs

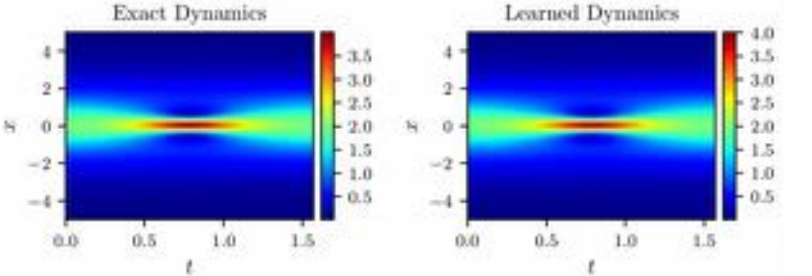
- Who needs Lorenz?



- Who needs Navier Stokes? $w_t = -uw_x - vw_y + 0.01(w_{xx} + w_{yy})$



- Who needs Schrödinger? $\psi_t = 0.5i\psi_{xx} + i|\psi|^2\psi$



"Discovering governing equations from data by sparse identification of nonlinear dynamical systems" Brunton, Proctor, Kutz, PNAS 2016

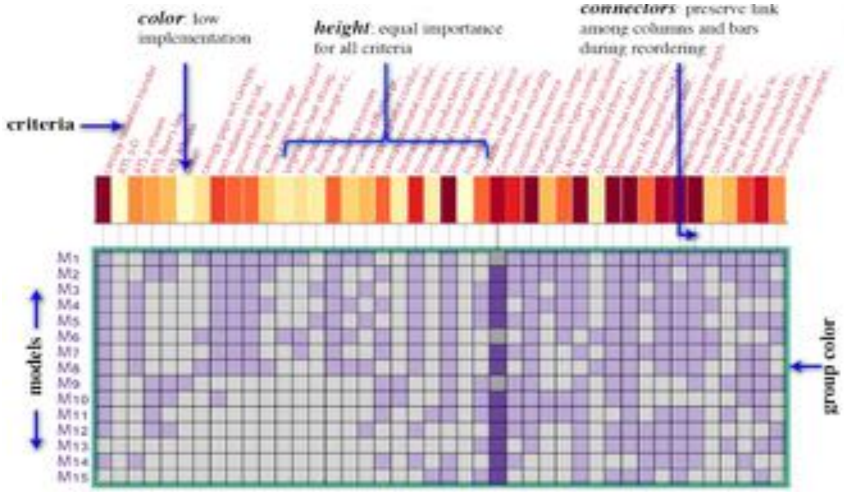
"Deep Hidden Physics Models: Deep Learning of Nonlinear Partial Differential Equations" Raissi, JMLR 2018

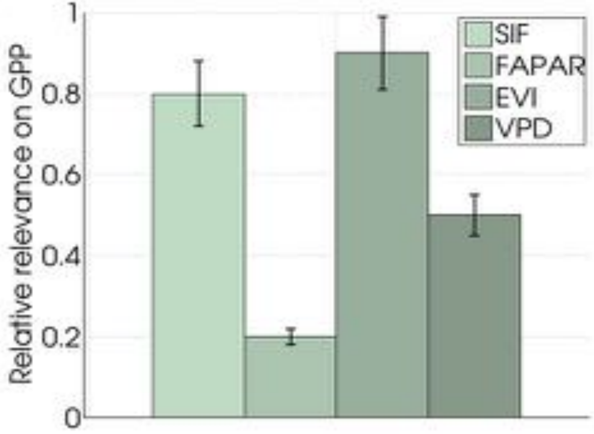
Understanding is more important than fitting

$F(X,y) =$ 

#1- Feature selection & ranking

- **Filters & wrappers**

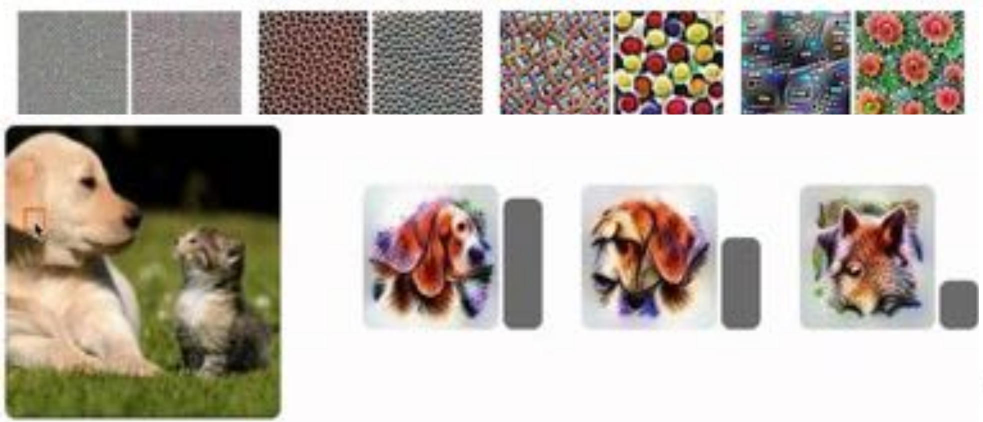


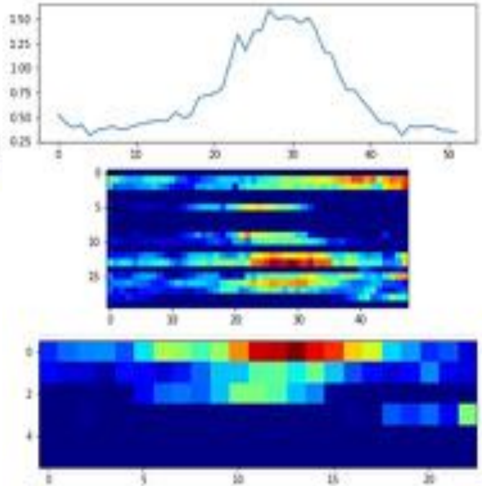


"Remote Sensing Feature Selection by Kernel Dependence Estimation", Camps-Valls, G. Mooij, JM. Schölkopf, IEEE-GRSL, 2010.
"A guided hybrid genetic algorithm for feature selection with expensive cost functions", M. Jung, J. Zscheischler, Procedia, 2013.

#2- Neuron and bases visualization

- What did the network learn?
- How do bases change in time, with real/simulations/together, under extremes?

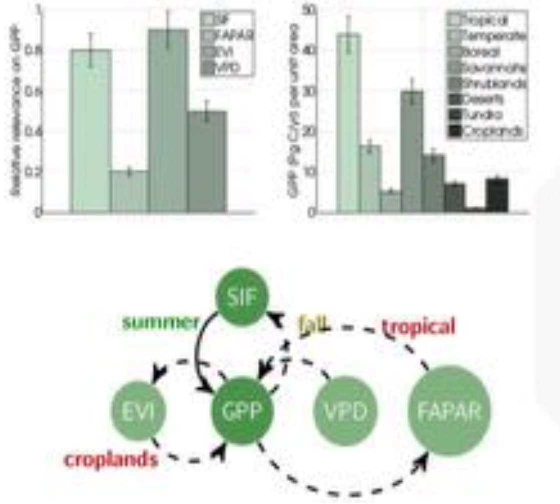


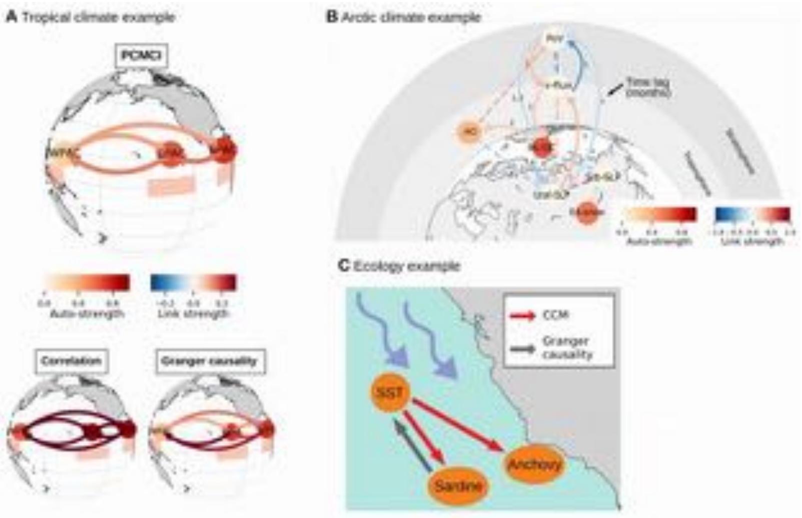


"Visualizing and Understanding Convolutional Networks", Zeriler, et al 2013
"Processing of Extremely high resolution LiDAR and optical data", Campos-Taberner, Camps-Valls et al, 2016
"DeeplyOut: What did your network learn under anomalies and adaptation? ," Camps-Valls et al, JMLR (2019)

#3- Graphical models and causality

- **Causality discovery** learns cause and effects relations from data
- **What for?** Hypothesis testing, model-data comparison, causes of extreme impacts





"Inferring causation from time series with perspectives in Earth system sciences", Runge, Bathiany, Bollt, Camps-Valls, et al. Nat Comm (submitted), 2018.
"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**

Conclusions

Conclusions

- **Machine learning in EO and climate**
  - Many techniques ready to use
  - Huge community, exciting tools
- **Solid mathematical framework**
  - Multivariate data
  - Multisource data
  - Structured spatio-temporal relations
  - Nonlinear feature relations
  - Fitting & understanding
- **Risks**: addictive, overfitting, overlooking
- **Remedies**: Physics-driven ML, Explainable AI, Causality