Learning ordinary differential equations from remote sensing data

EGU General Assembly 2020

Jose E. Adsuara, Adrián Pérez-Suay, Alvaro Moreno-Martínez, Anna Mateo-Sanchis, Maria Piles, Guido Kraemer, Markus Reichstein, Miguel D. Mahecha, Gustau Camps-Valls.

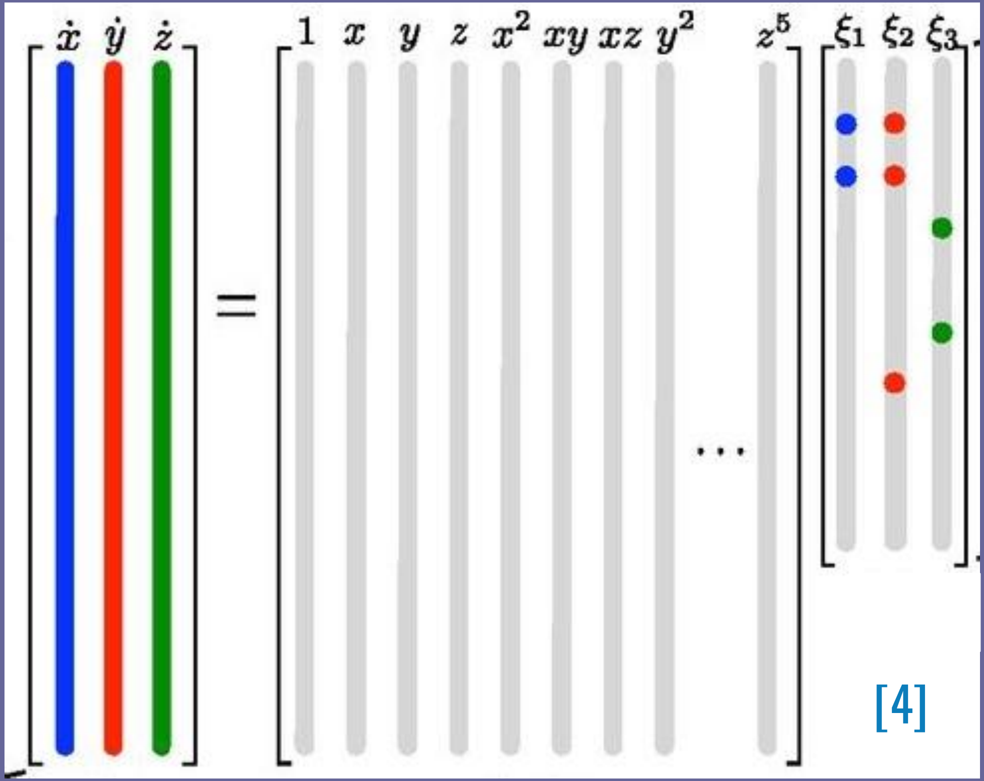
IPL Image Processing Laboratory
erc
iDiv
Image Signal Processing - ISP
Max Planck Institute for Biogeochemistry

http://isp.uv.es   @isp_uv_es   jose.adsuara@uv.es

Introduction

- Modeling and understanding the Earth system is a constant and challenging scientific endeavour.
- Learn from observational data using machine learning can be an alternative, but understanding is more difficult that fitting. [1,2,3]
- We introduce sparse regression to uncover a set of governing equations in the form of a system of ordinary differential equations (ODEs)… [4]
- … and used to explicitly describe a simplest ODEs explaining data to model relevant components of the biosphere. [5]

Sparse identification of dynamical systems



[4]

Sparse identification of dynamical systems

We consider dynamical systems of the form:

$$\frac{d}{dt}\boldsymbol{x}(t) = \boldsymbol{f}(\boldsymbol{x}(t))$$

$$\boldsymbol{x}(t) \in \mathbb{R}^n$$

that can be expanded as:

$$\frac{d}{dt}x_1(t) = f_1(x_1, ..., x_n)$$
$$\frac{d}{dt}x_2(t) = f_2(x_1, ..., x_n)$$
$$\vdots$$
$$\frac{d}{dt}x_n(t) = f_n(x_1, ..., x_n)$$

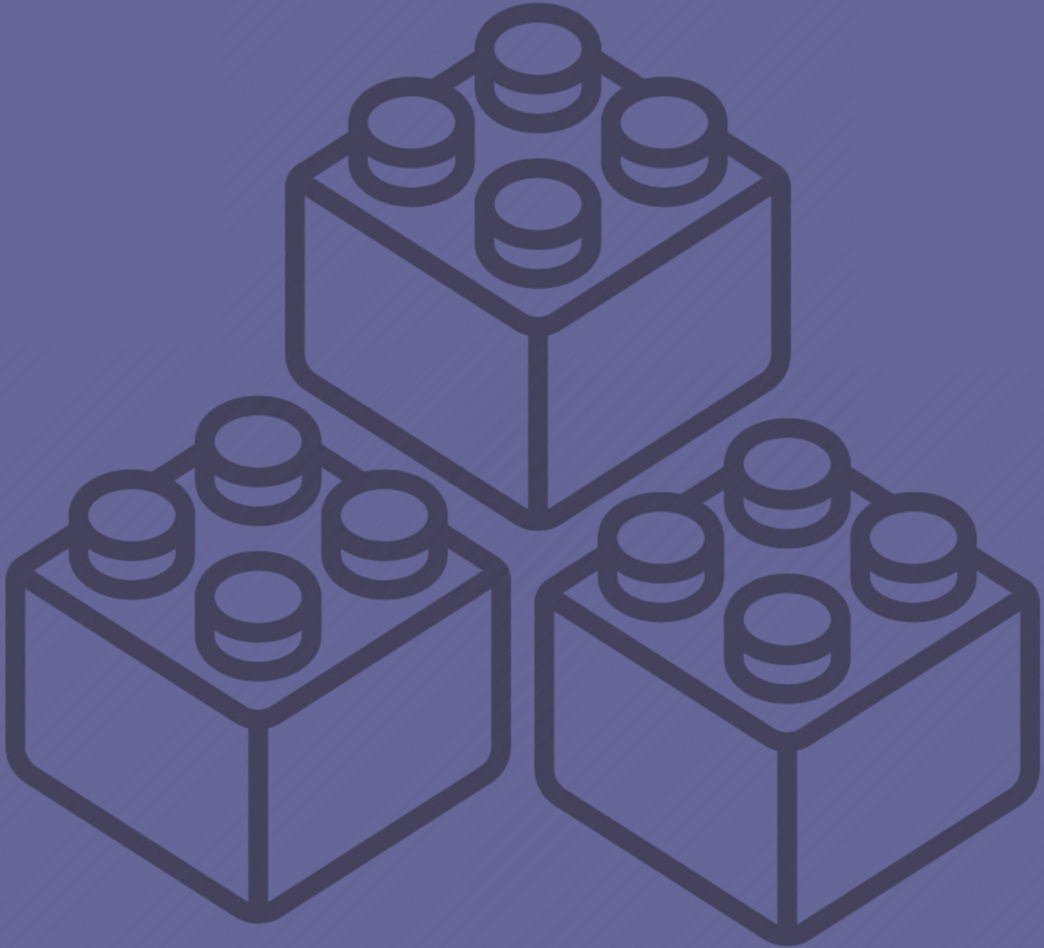
[4] For many systems, the right part of the equations are sparse in the space of possible functions, so a library of candidate functions l_i is needed:

$$\frac{d}{dt}x_1(t) = \varepsilon_{11}l_1(x_1,\dots,x_n) + \varepsilon_{11}l_2(x_1,\dots,x_n) + \dots + \varepsilon_{1m}l_m(x_1,\dots,x_n)$$
$$\frac{d}{dt}x_2(t) = \varepsilon_{21}l_1(x_1,\dots,x_n) + \varepsilon_{21}l_2(x_1,\dots,x_n) + \dots + \varepsilon_{2m}l_m(x_1,\dots,x_n)$$
$$\vdots$$
$$\frac{d}{dt}x_n(t) = \varepsilon_{n1}l_1(x_1,\dots,x_n) + \varepsilon_{n1}l_2(x_1,\dots,x_n) + \dots + \varepsilon_{nm}l_m(x_1,\dots,x_n)$$

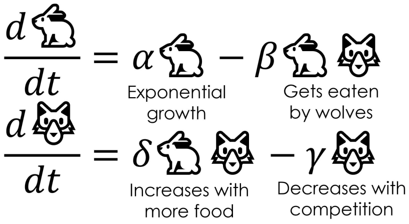
- l_1, l_2, ..., l_m is a predefined finite library of candidate functions.
- express x_i as a linear combinations of them with coefficients.
- learn scalar coefficients ε_{ij} using ridge regression (RLR), LASSO (maximizes the number of zeros) or Elastic Net (ENet) (convex combination of the above)
- we need computing derivatives of the data: finite differences or by kernel regression aka Gaussian processes. [3, 6]

Results

- Lotka-Volterra
- Biosphere indicators
- Exploring other data



Toy problem: Lotka-Volterra system



- Prey-predator model in ecology:

$$\frac{d}{dt}x = \alpha x - \beta x y$$
$$\frac{d}{dt}y = -\gamma y + \delta x y$$

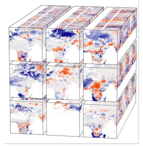
- We set α = 3/2, β = 1, γ = 3, δ = ½.
- Synthetic data + two levels of AWGN: 40dB, 5dB.
- ridge regression, finite differences.
- train/test in 75%-25% samples.

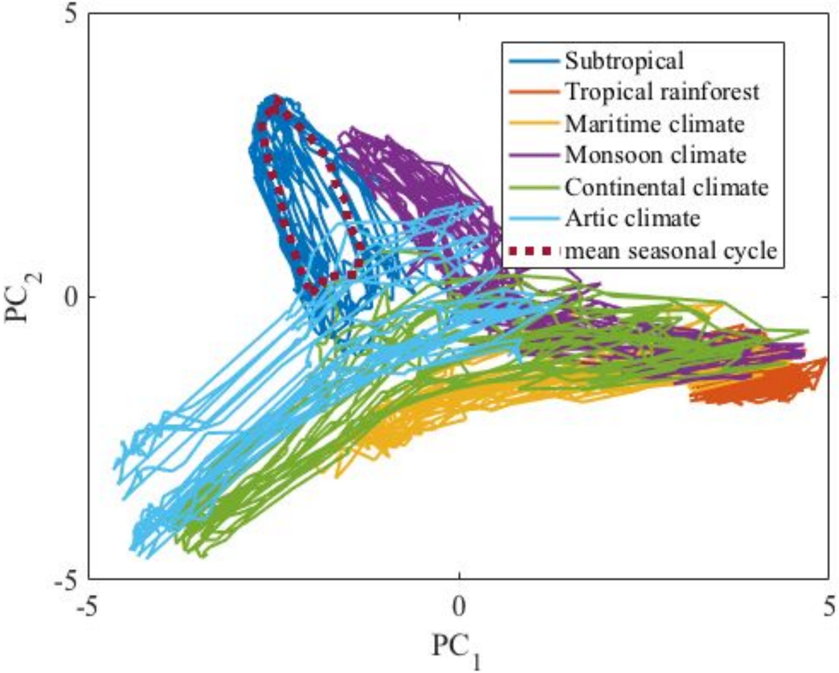
Library functions	Coefficients 40 dB $\frac{d}{dt}x$	Coefficients 40 dB $\frac{d}{dt}y$	Coefficients 5 dB $\frac{d}{dt}x$	Coefficients 5 dB $\frac{d}{dt}y$
x	1.3822	0	1.1404	0
y	0	-2.9123	0	-2.7946
x^2	0	0	0	0
xy	-0.9797	0.4849	-0.9520	0.4710
y^2	0	0	0	0
x^3	0	0	0	0
y^3	0	0	0	-0.0001
1	0	0	0	0
R	0.9999		0.8674	

- accurated correlation coefficient (R)...
- ... so, ODE coefficients recovered accurately.
- critial points recovered without a qualitative change in their type (a saddle point & a center of cycles).

Biospheric indicators: data

- We use the biosphere indices proposed in [5] for summarizing the state of an ecosystem.
- Earth System Data Lab (ESDL)
  - 12 variables (common spatiotemporal grid: 0.25º in space, 8 days in time)
  - PCA incorporating information about latitude
- First two principal components explained 73% of variance→two biosphere indicators:
  - PC_1 summary of vegetation productivity
  - PC_2 summary of water availability





Trajectories in the phase space of this first two PCs for the most paradigmatic ecosystems along 11 years.

Biospheric indicators: learned dynamical model

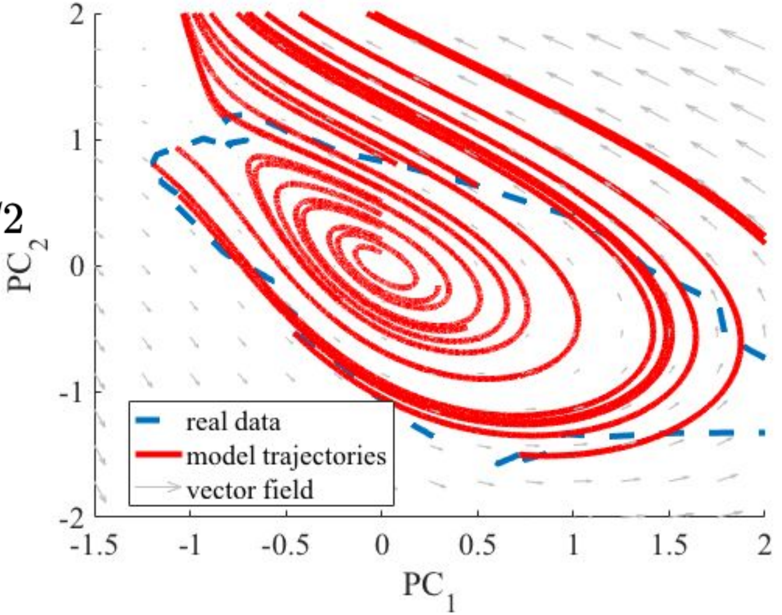
- We start by focusing on the particular subtropical ecosystem.
- Work with the mean seasonal cycle trajectory, which summarizes the state of the ecosystem throughout the year.

$$\begin{aligned}\frac{d}{dt}\mathrm{PC}_1 &= \varepsilon_{11} l_1(\mathrm{PC}_1, \mathrm{PC}_2) + \ldots + \varepsilon_{1m} l_m(\mathrm{PC}_1, \mathrm{PC}_2)\\ \frac{d}{dt}\mathrm{PC}_2 &= \varepsilon_{21} l_1(\mathrm{PC}_1, \mathrm{PC}_2) + \ldots + \varepsilon_{2m} l_m(\mathrm{PC}_1, \mathrm{PC}_2).\end{aligned}$$

$$\frac{d}{dt}PC1 = -37.5PC1 - 55.6PC2 - 31.9PC1PC2$$

$$\frac{d}{dt}PC2 = 67.2PC1 + 44.8PC2 - 74.0PC1PC2$$

(x 10^{-4})



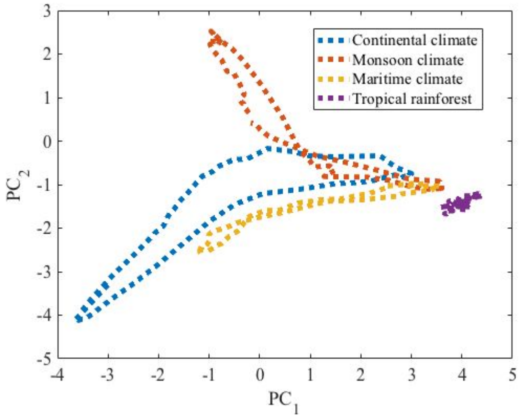
- System analysis: attractor at 0.000365± 0.00451995j
  - (removing real part and recover the new system?)

R^2 = 0.73 (Coefficient of determination)

Biospheric indicators: learned dynamical model (II)

- We complete the study for the rest of the paradigmatic ecosystems.
- We learn the models not only using the first two PCs, but also for the first three PCs.
- Monomials with max. deg. from 2 up to 12. (bid.: $\{x^2,xy,y^2\}$, $\{x^3,x^2y,xy^2,y^3\}$,...,$\{x^{12},x^{11}y,...,y^{12}\}$; trid.: $\{x^iy^jz^k:i+j+k=12\}$)
- We repeat each experiment 10 times with 10 different train/test partitions (in 75%-25%)

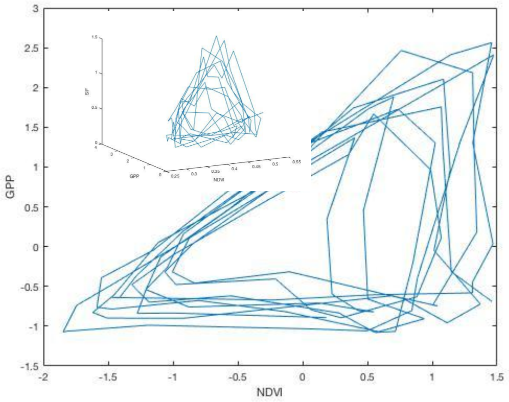
bidimensional: PC_1, PC_2			tridimensional: PC_1, PC_2, PC_3		
Ecosystem	max. deg. (method)	R2	Ecosystem	max. deg. (method)	R2
Continental	3 (LASSO)	*0.87 ± 0.06*	Continental	3 (LASSO)	0.79 ± 0.23
Monsoon	4 (RLR)	0.25 ± 0.23	Moonsoon	3 (LASSO)	*0.60 ± 0.29*
Maritime	3 (RLR)	0.38 ± 0.25	Maritime	2 (RLR)	0.35 ± 0.16
Tropical	2 (RLR)	0.20 ± 0.12	Tropical	2 (RLR)	0.40 ± 0.17



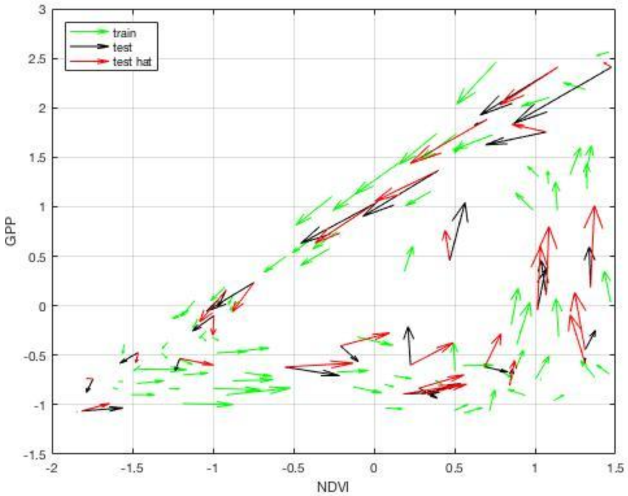
deg: best maximum degree
R2: coefficient of determination

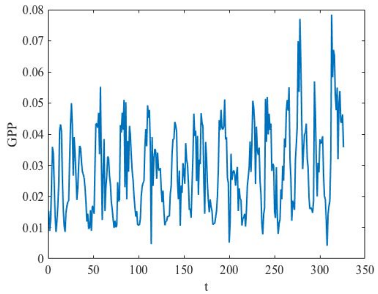
Exploring other data: NDVI/GPP models, GPP/VOD models, …

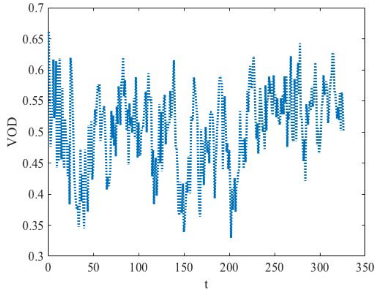
- NDVI/GPP
  - looking for “good” pixels.
  - 9 years of data.
  - One measure per month.
- GPP/VOD
  - GPP tower: lat = -35.6557, lon = 148.1521



R^2=0.75







R^2=0.81

Conclusions and future work

Conclusions and future work

- Presented a methodology for obtaining an analytic model with ODEs from data using sparse identification.
- Applied to both a toy model and a set of biosphere indices obtained from EO.
- Learned model captures the dynamics of the system: water availability and vegetation productivity strongly coupled with exponential grow/decays.
- Study other (less aggressive) sparse-promoting strategies.
- Study other types of (more physically-inspired or large) orthonormal basis of functions.
- Interventional studies like distortions in the eigenvalues and its effects on the phase space.
- More careful study of the models with new data.

References

[1] M. Reichstein, et al., Deep learning and process understanding for data-driven Earth system Science, Nature (2019)

[2] G. Camps, et al., Advances in hyperspectral image classification: Earth monitoring with statistical learning methods, SP (2014)

[3] G. Camps, et al. A survey on Gaussian processes for earth observation data analysis, GRS (2016)

[4] S. L. Brunton, et al., Discovering governing equations from data by sparse identification of nonlinear dynamical systems, PNAS (2016)

[5] G. Kraemer, et al., Summarizing the state of the terrestrial biosphere in few dimensions, Biogeosciences (2020)

[6] J. E. Johnson, et al., Disentangling derivatives, uncertainty and error in gaussian process models, IGARSS (2018)

Thanks!

EGU General Assembly 2020

Jose E. Adsuara, Adrián Pérez-Suay, Alvaro Moreno-Martínez, Anna Mateo-Sanchis, Maria Piles, Guido Kraemer, Markus Reichstein, Miguel D. Mahecha, Gustau Camps-Valls.

IPL IMAGE PROCESSING LABORATORY

erc

iDiv

http://isp.uv.es

@isp_uv_es

jose.adsuara@uv.es

Image Signal Processing - ISP

Max Planck Institute for Biogeochemistry