Harnessing sensing technologies for a smarter agriculture

Andrea Maino

Tutor:
Prof. Fabio Mantovani

INFN

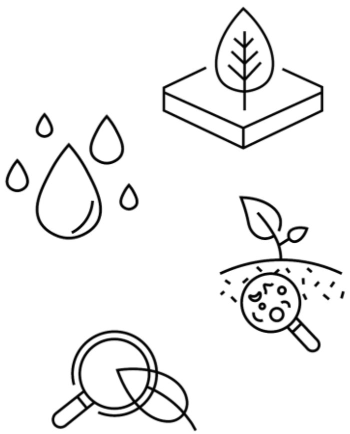
Università degli Studi di Ferrara

Laboratory for Nuclear Technologies Applied to the Environment

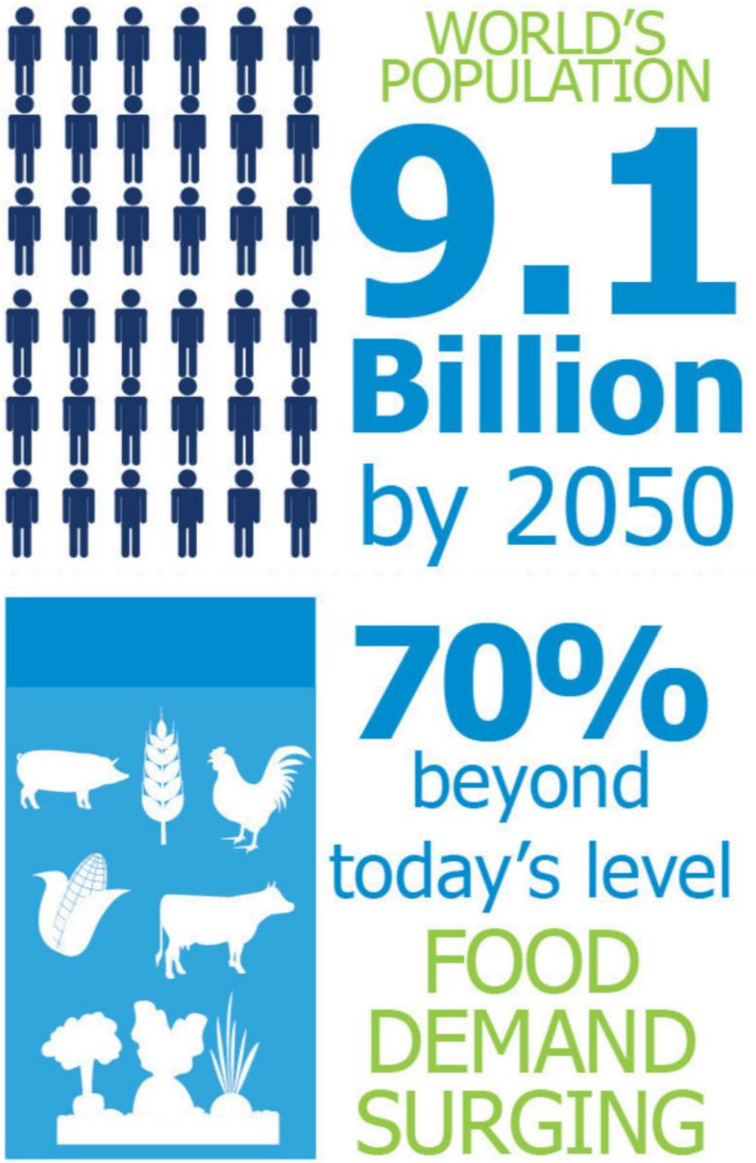
Scientific motivation

- World’s population is expected to grow to **9.1 Billion people by 2050**, with a **70% increase in food demands** across the world (FAO, 2014)
- Unlike demands, **land and water resources are limited**
- **70% of global freshwater** is already used by agriculture

- **Resource efficiency** must increase:
    - **Better land use** for high yield productions
    - Reduced **water wastes**
    - Reduced use of **fertilizers and pesticides**
    - **Disease monitoring and prevention** strategies

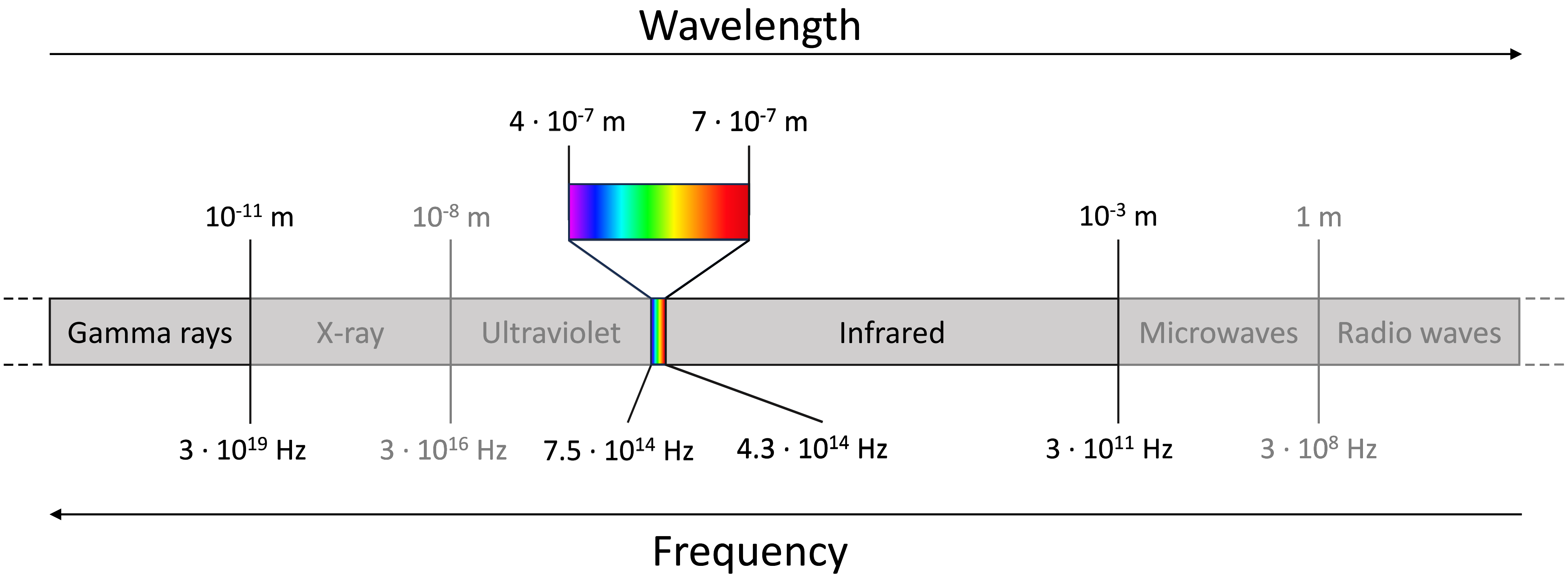


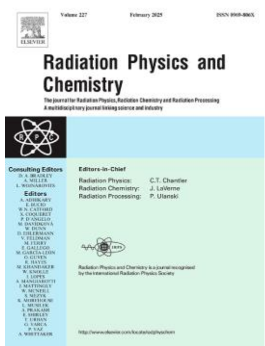
- **Smart agriculture** through **quantitative and measured data**
- Highly accurate **sensing technologies**



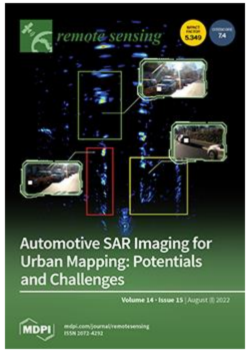
Icons by Vecteezy: www.vecteezy.com/free-vector/agriculture-icon

Some highlights from my PhD journey

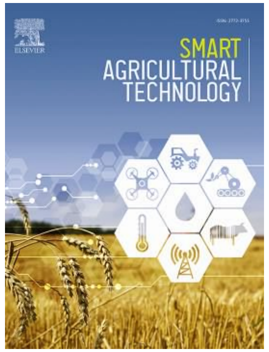




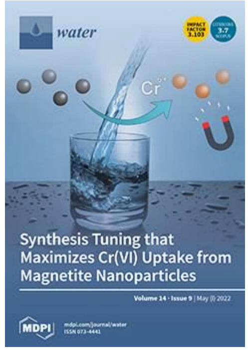
A deep neural network for predicting soil texture using airborne radiometric data



Airborne Radiometric Surveys and Machine Learning Algorithms for Revealing Soil Texture



Moving forward the automatic detection of Flavescence Dorée in vineyards with airborne imaging



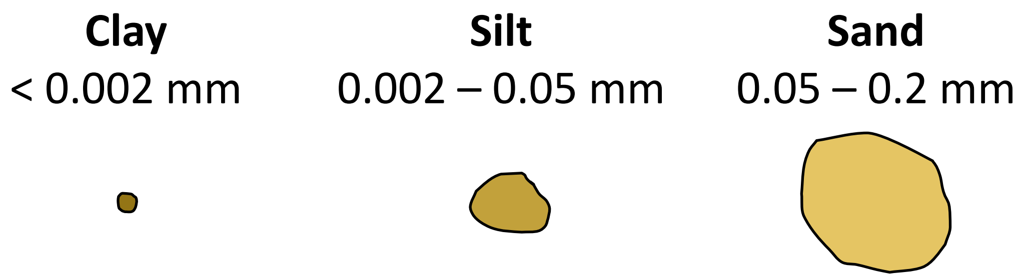
Combining Precision Viticulture Technologies and Economic Indices to Sustainable Water Use Management

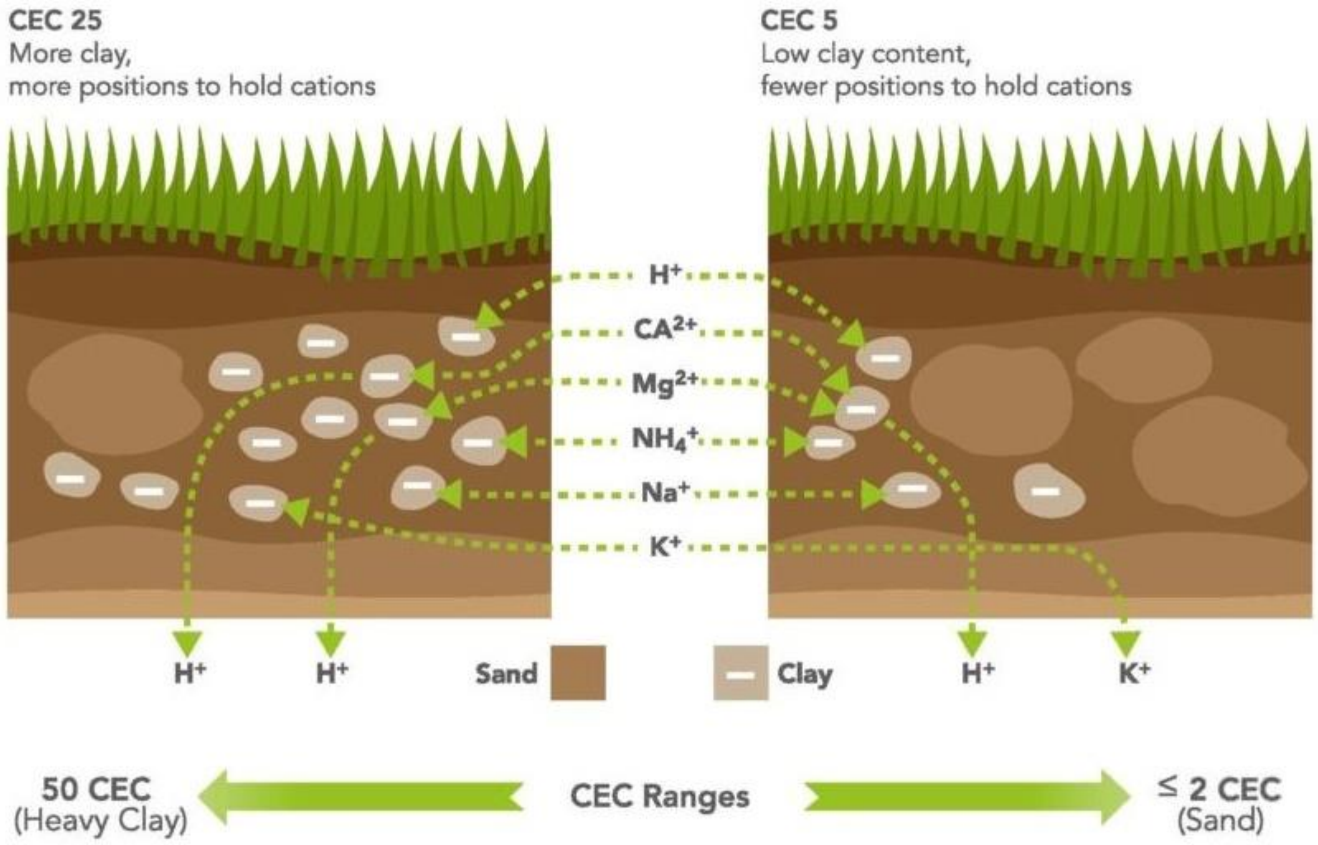
Radioelements and soil texture

- Small soil particles (mostly those smaller than **1 µm**) can act as **colloids** and **adsorb cations**.
- The sum of the exchangeable cations in the soil is called the **Cation Exchange Capacity (CEC)**.
- **Clay** and **silt** soil fractions have high **specific surface** and high **CEC** so they can adsorb cations and specifically positively ionized natural radioelements (**K^+, U^{4+}, U^{6+}, Th^{4+}**).

↓

Natural radioelements concentration in the soil is **correlated to soil texture**.

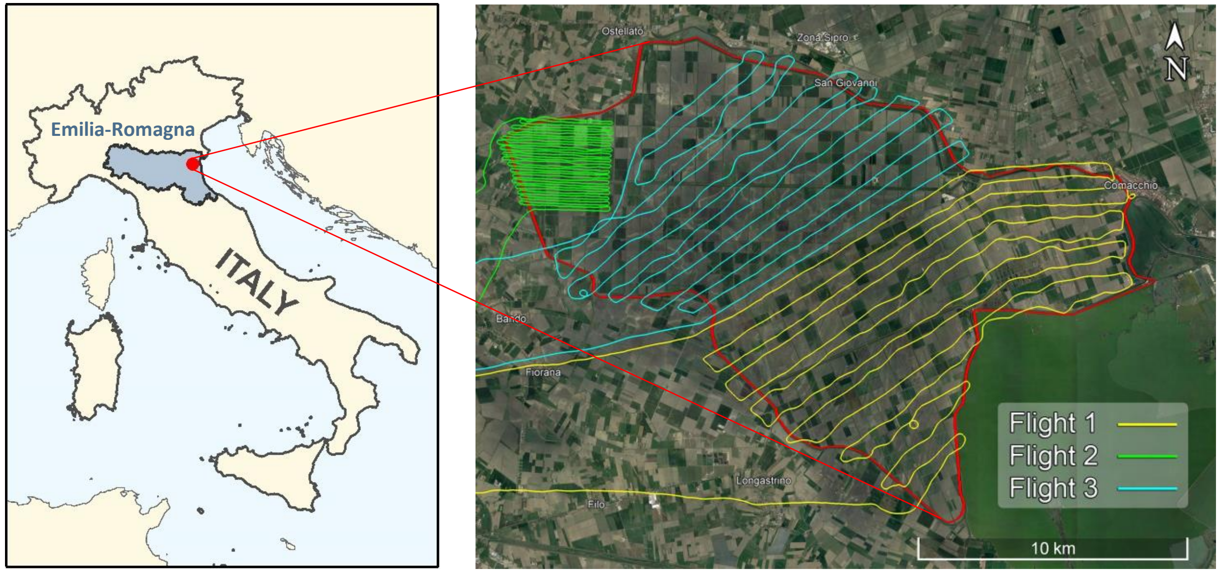




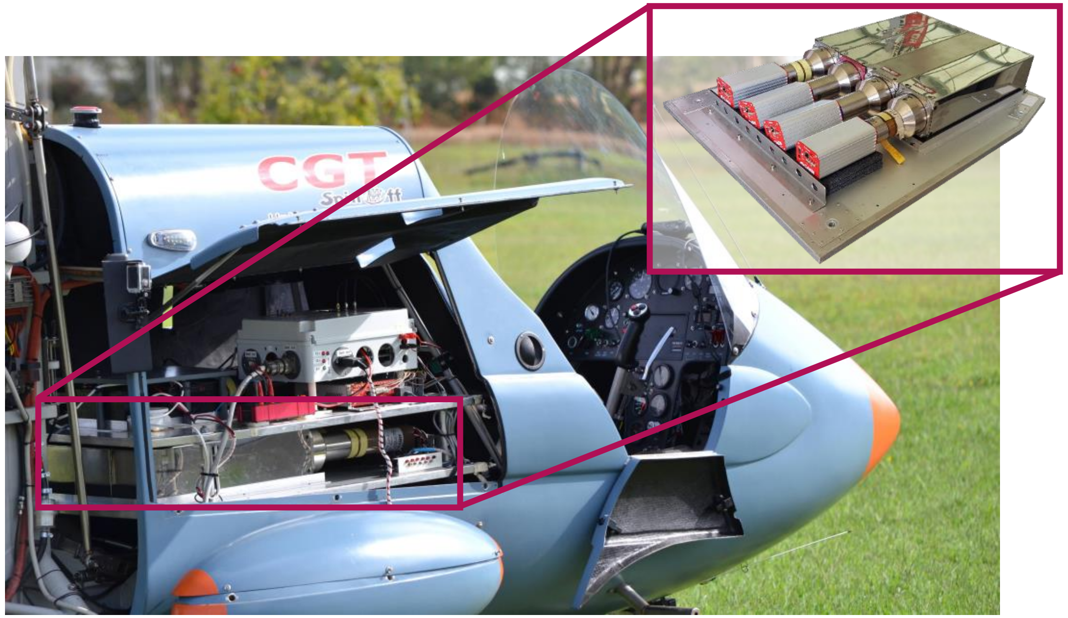
Data taking surveys

- Accurate **soil texture mapping** is key in precision agriculture for planning cultivations and targeting interventions.
- For large areas, **direct measurement** methods are **wasteful** in terms of time and money.
- **Airborne Gamma Ray Spectroscopy (AGRS)** overcomes these issues allowing for a fast and efficient mapping of large areas.

- 3 surveys over the Mezzano Lowland, Ferrara (~ **189 km^2**) for a total of **4 hours and 45 minutes**.
  - Mean **flight height** of 104 ± 21 m
  - Mean **velocity** of 102 ± 13 km/h
  - Field of view (FOV) of **radius 300 m**
  - Measurements every **300 m**.
- Gamma spectra acquired with a time resolution of 10 s for a total of **1469 spectra**.



Instrumentation and analysis

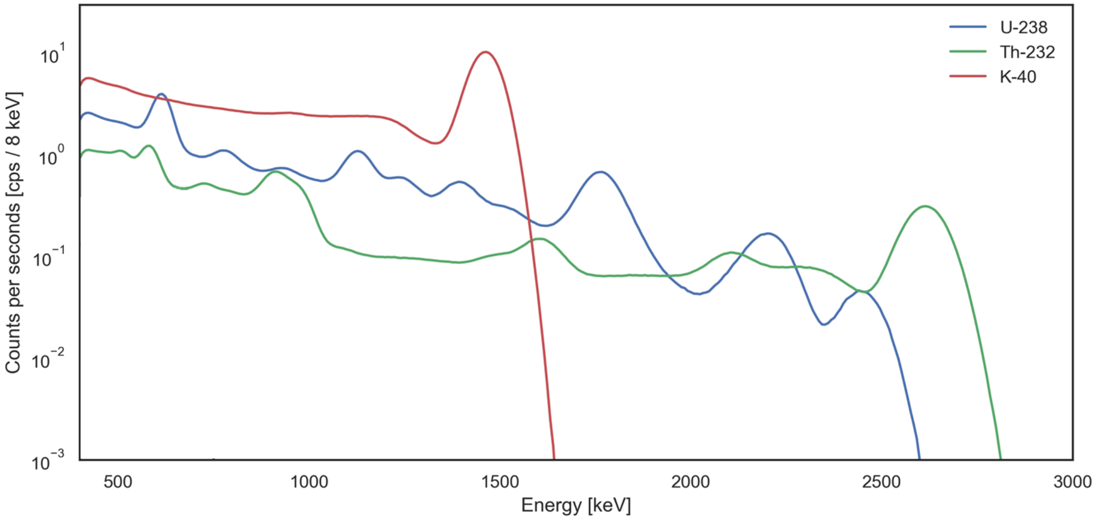


- **16 L NaI(Tl) crystals** surrounded by 1 mm thick stainless steel housings (4 modules).
- **6.8 % energy resolution** at 662 keV (^{137}Cs).
- Minimum Detectable **activity Concentration (MDC)** and **Abundance (MDA)**:

	MDC
^{40}K	16 Bq/kg
^{238}U	4.94 Bq/kg
^{232}Th	3.25 Bq/kg

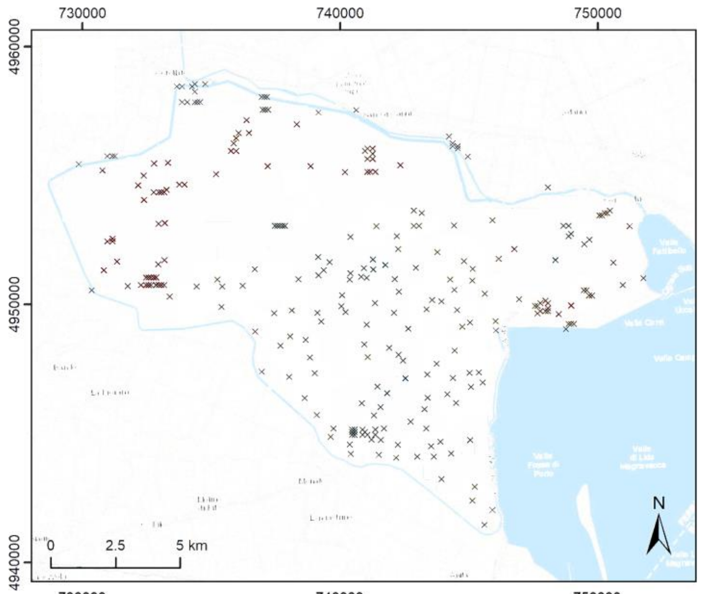
→

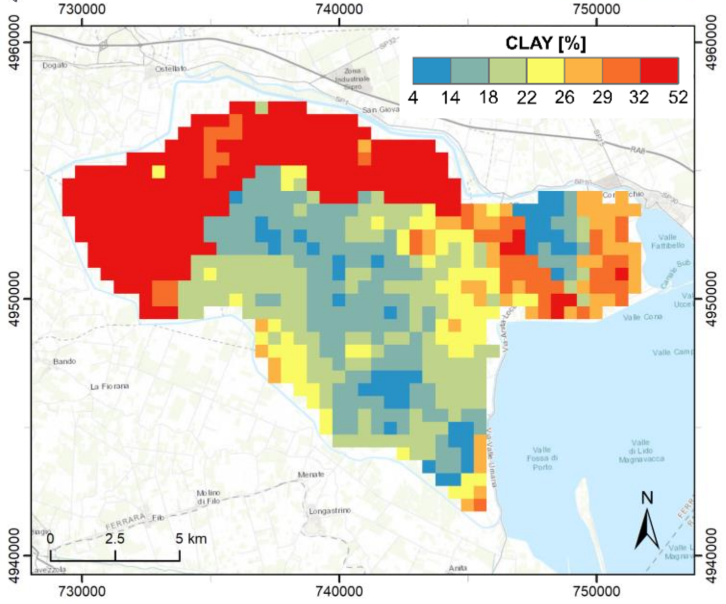
MDA	
0.05 · 10^{-2} g/g	K
0.4 µg/g	U
0.8 µg/g	Th



- **Full spectrum analysis** with ^{40}**K**, ^{238}**U** and ^{232}**Th** simulated fundamental spectra (assuming secular equilibrium conditions).

Model inputs and data manipulation



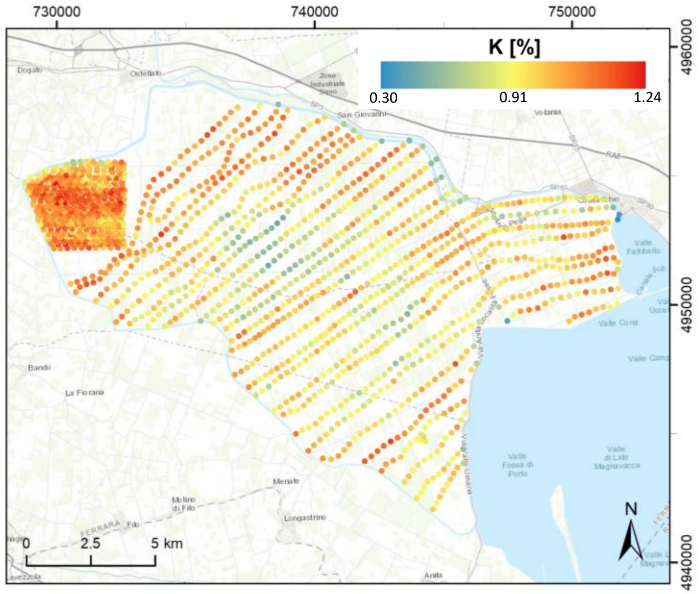


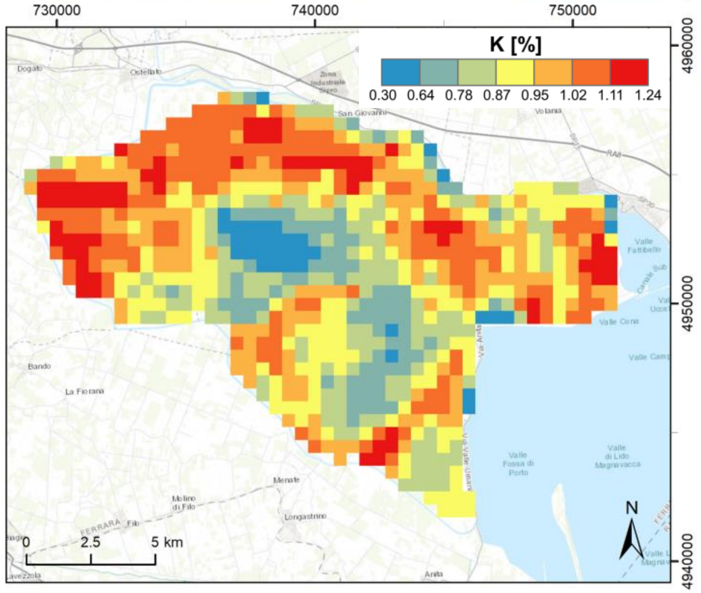
273 punctual soil texture measurements from **Regione Emilia-Romagna (RER)**

Kriging spatial interpolation

500 m x 500 m resolution **clay and sand soil content** maps

723 square meshes





1469 geolocalized K and Th abundance data points obtained via **airborne surveys**

Kriging spatial interpolation

500 m x 500 m resolution **K and Th abundance** maps with matching grid

723 square meshes

Correlation between texture and gamma

Linear Regression (LR)

Previous studies:

Van Der Klooster, E. et al. [0.1111/j.1365-2389.2011.01381.x]

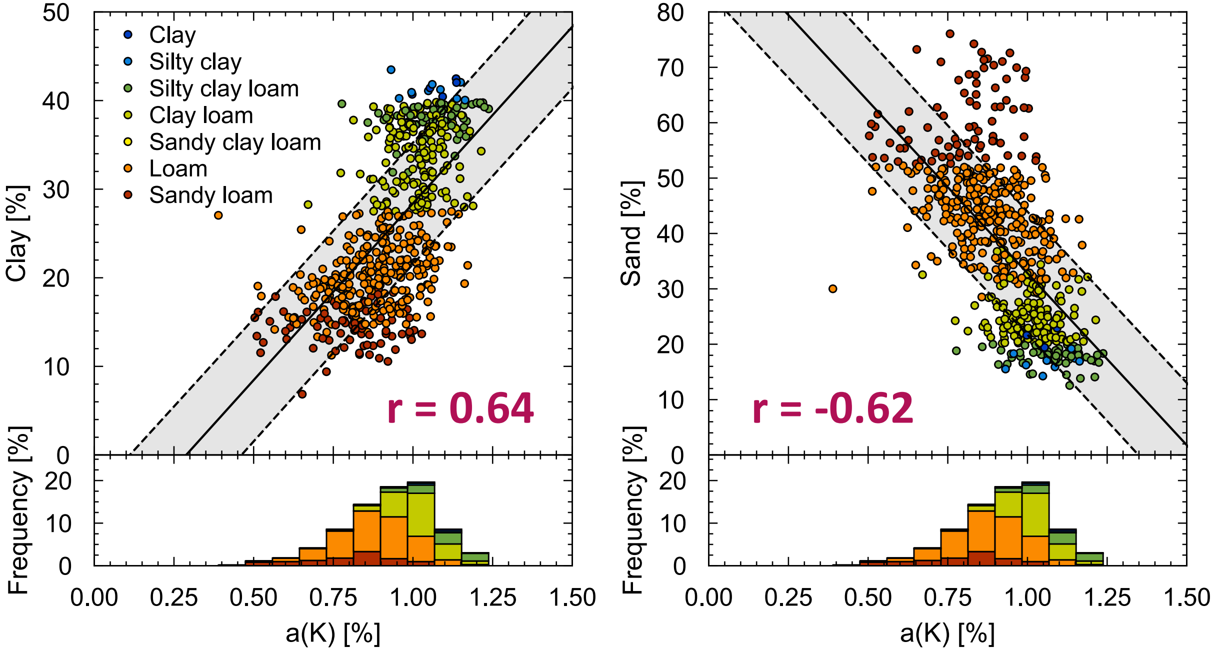
Mahmood, H. S. et al.[10.3390/s131216263]

Spadoni, M. and Voltaggio, M. [0.1016/j.gexplo.2012.10.016]

Elbaalawy, A. M. et al. [10.5958/2395-146X.2016.00038.7]

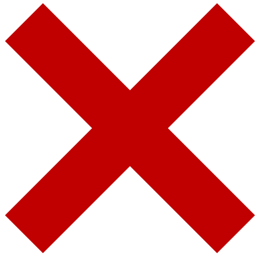
Petersen, H. et al. [10.1002/jpln.201100408]

My study:

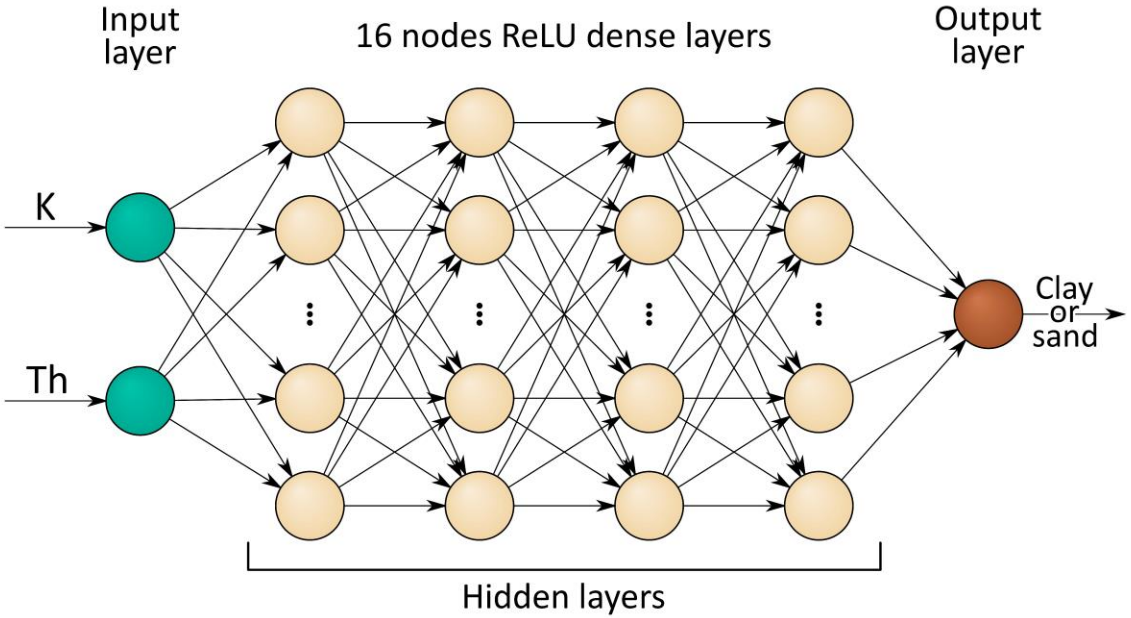


Non-Linear Machine Learning (NLML)

Previous studies:

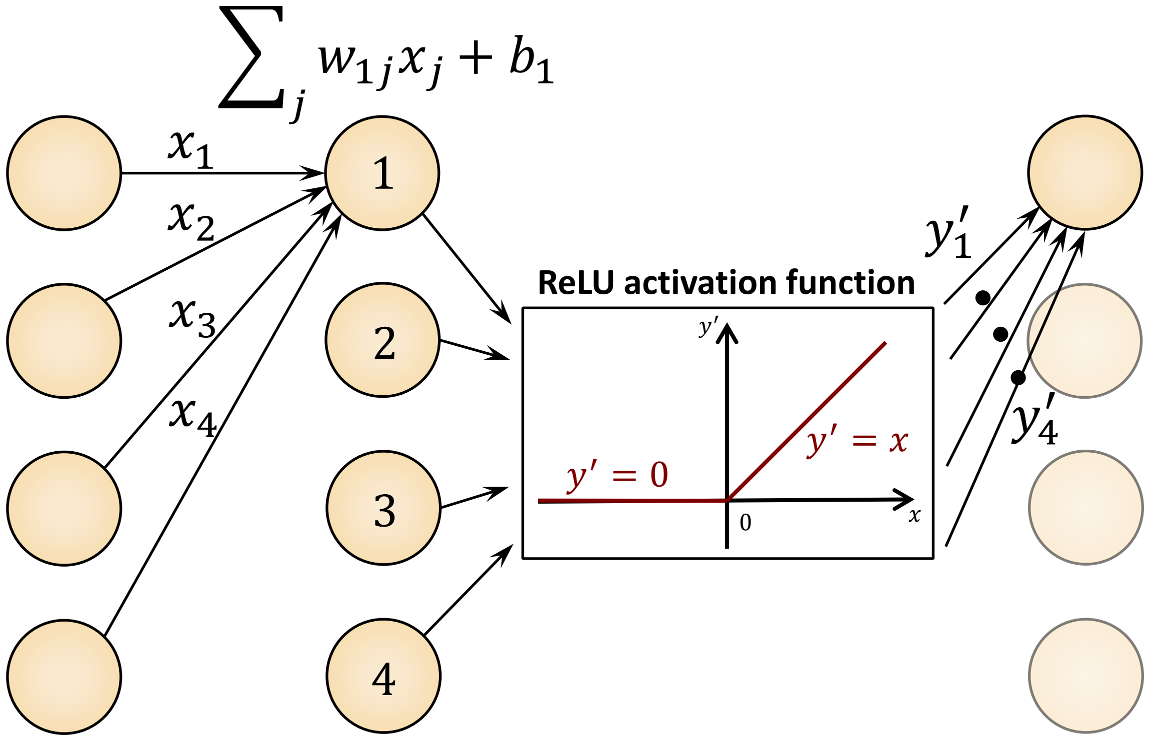


My study:



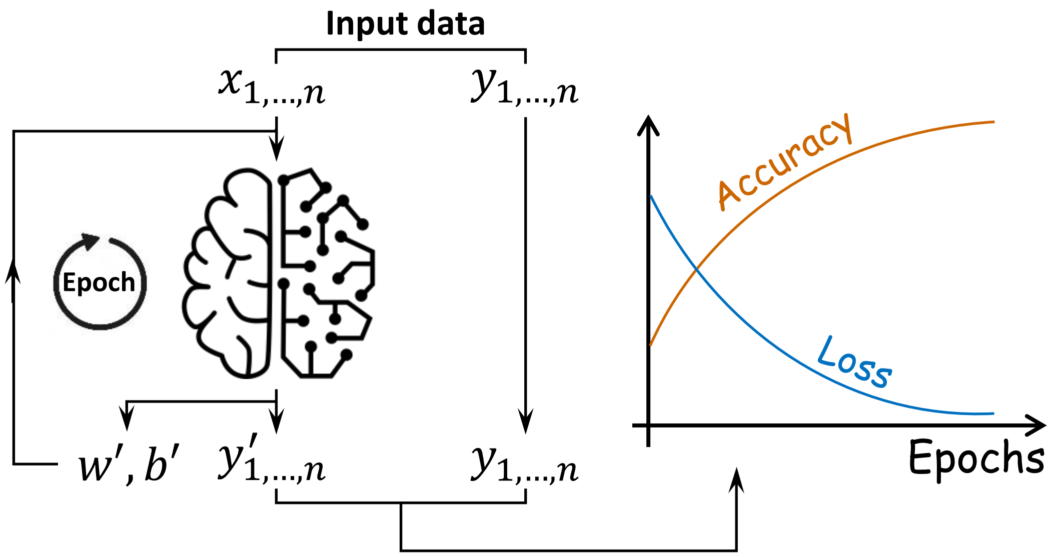
Unpacking the Machine Learning algorithm

- Data flows from **left to right** in the algorithm's architecture.
- In a Deep Neural Network, **each node** in a layer *i* receives data **from each node** of the previous layer *j*.
- Each node performs the following operations:
  - Each received input is multiplied by a **weight**.
  - The weighted values are **summed**.
  - A **bias** is added.
  - The result is passed through an **activation function** to the next layer.



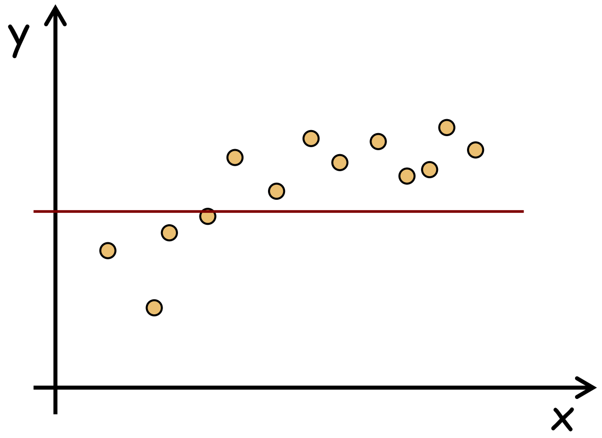
$$y'_i = f\left(\sum_j w_{ij}x_j + b_i\right)$$

Learning epoch by epoch

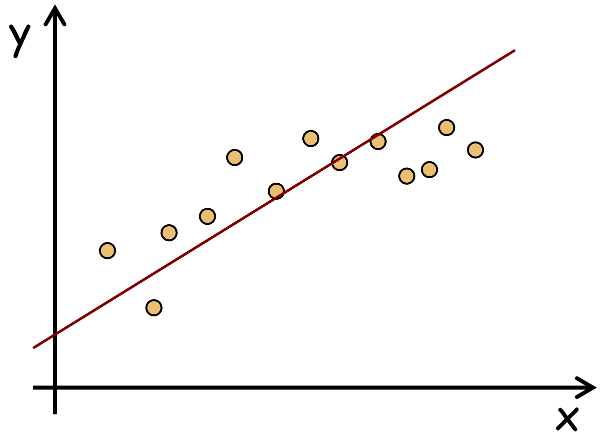


- The data is fed to the algorithm multiple times, called **epochs**.
- Between each epoch, the algorithm changes its learned parameters (**weights** and **biases**) to better model the input data.
- **After each epoch**, performances are evaluated in terms of **accuracy** (quality of the predictions $y'_{1,\dots,n}$) and/or **loss** (difference from the ground-truth $y_{1,\dots,n}$).

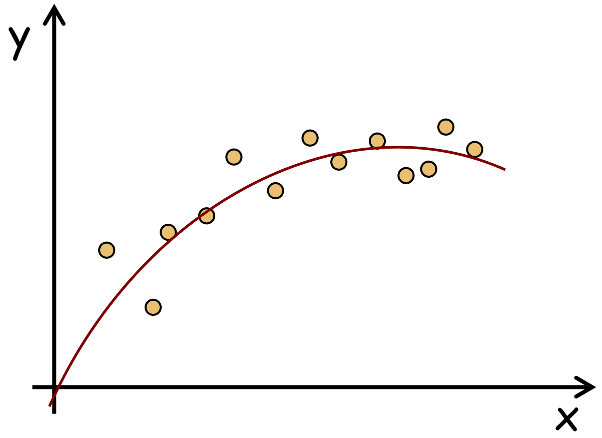
Epoch 1



Epoch 2

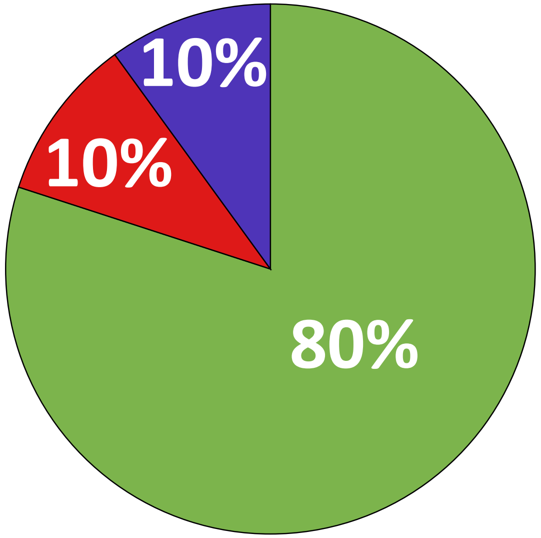


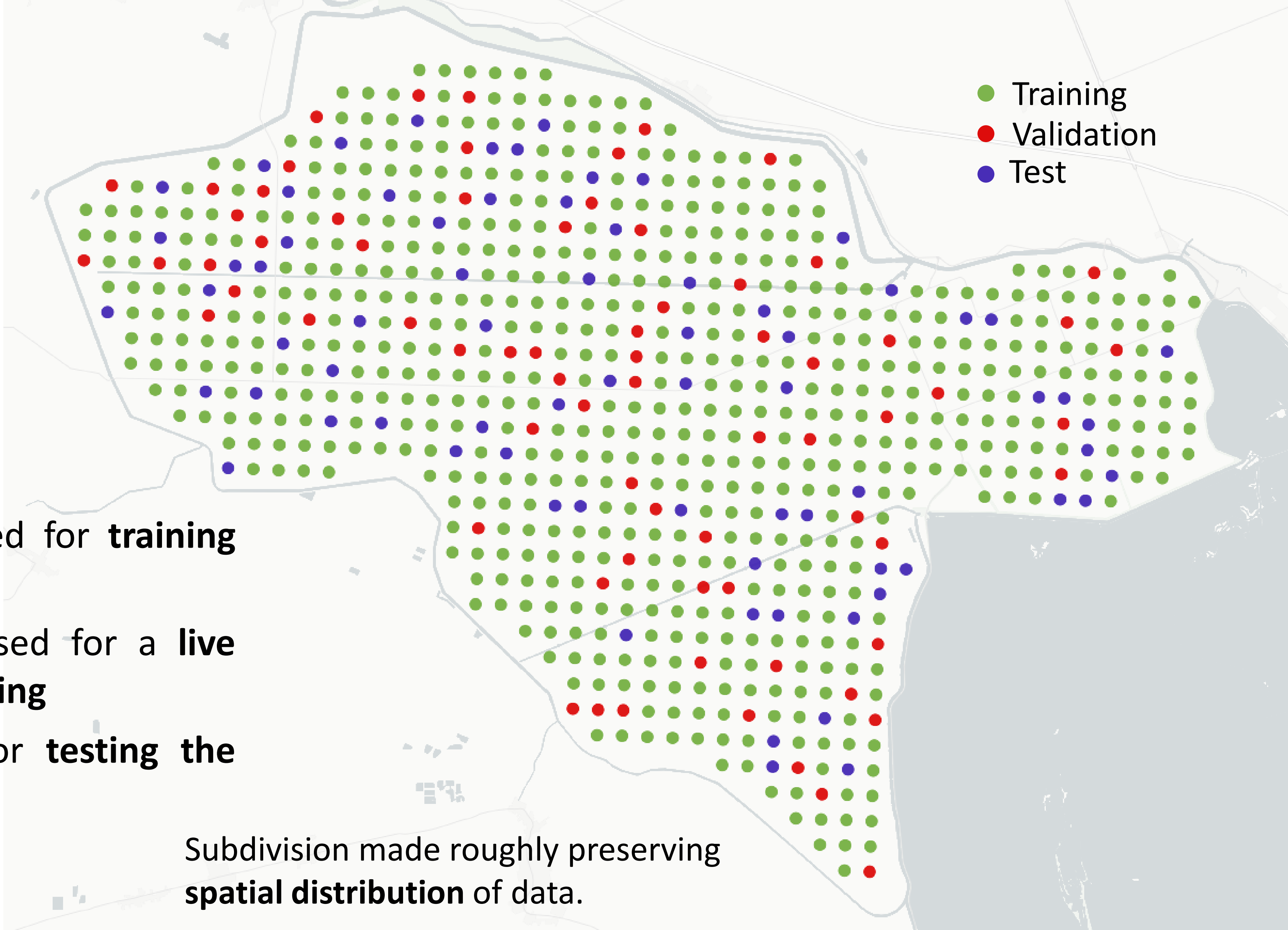
Epoch 3



…

Preparing the dataset for analysis





Training dataset: 578 data used for **training the algorithm**

Validation dataset: 72 data used for a **live performance check during training**

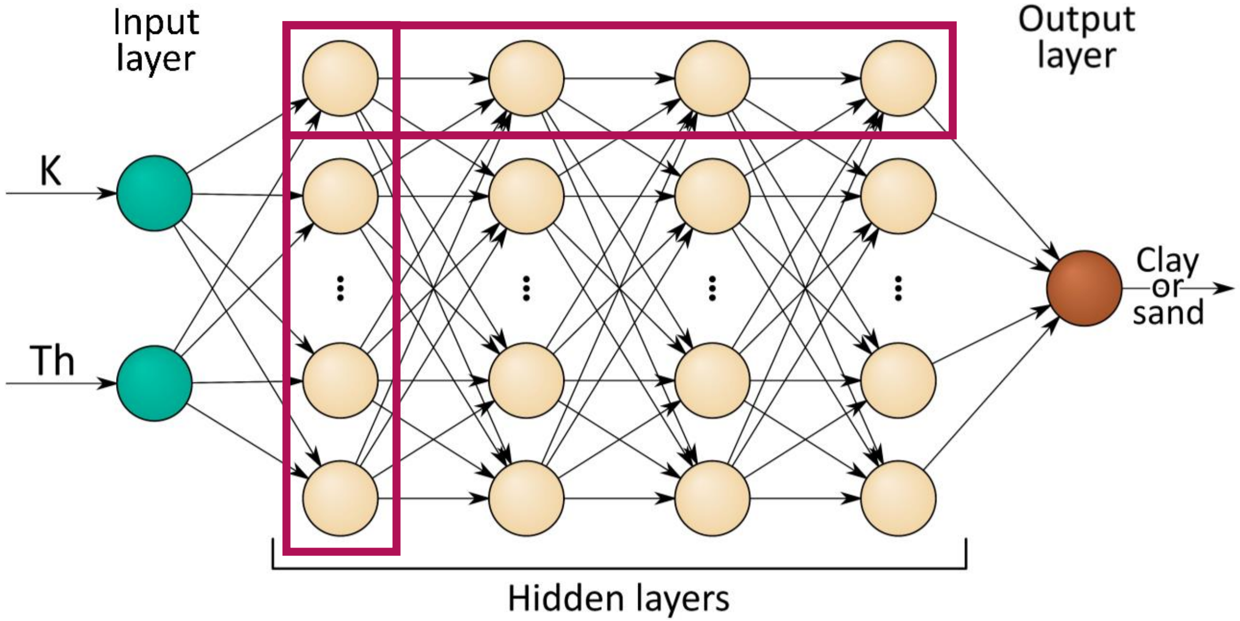
Test dataset: 73 data used for **testing the algorithm's final performances**

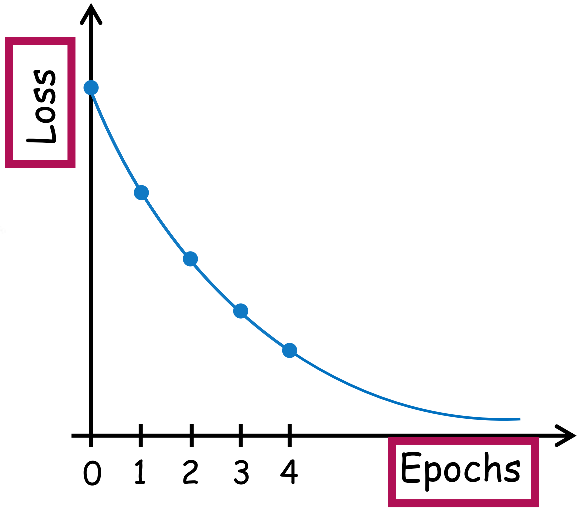
Subdivision made roughly preserving **spatial distribution** of data.

Hyperparameters explained

Hyperparameter	Definition
Width	Number of nodes in a given layer
Depth	Number of hidden layers
Batch size	Size of input data “packages”
Activation function	Modifies the outputs of a layer
Loss function	Quantifies prediction error
Optimizer	Minimizes the loss function during training
Learning rate	Fraction of the parameters’ updates applied after each batch
Number of Epochs	Number of learning cycles

ID	K	Th	Clay	Sand
1	0.9	6.7	24	47
2	1.2	5.8	35	25
3	1.3	7.1	23	36
4	0.7	6.5	12	34
5	0.5	6.6	27	73
6	1.1	8.2	31	64
7	0.8	5.9	16	53





Learning rate

$$\theta_1 \rightarrow \theta_1 - \alpha \frac{\partial}{\partial \theta_1} Loss(\theta_1)$$

If **α** is **too small**:

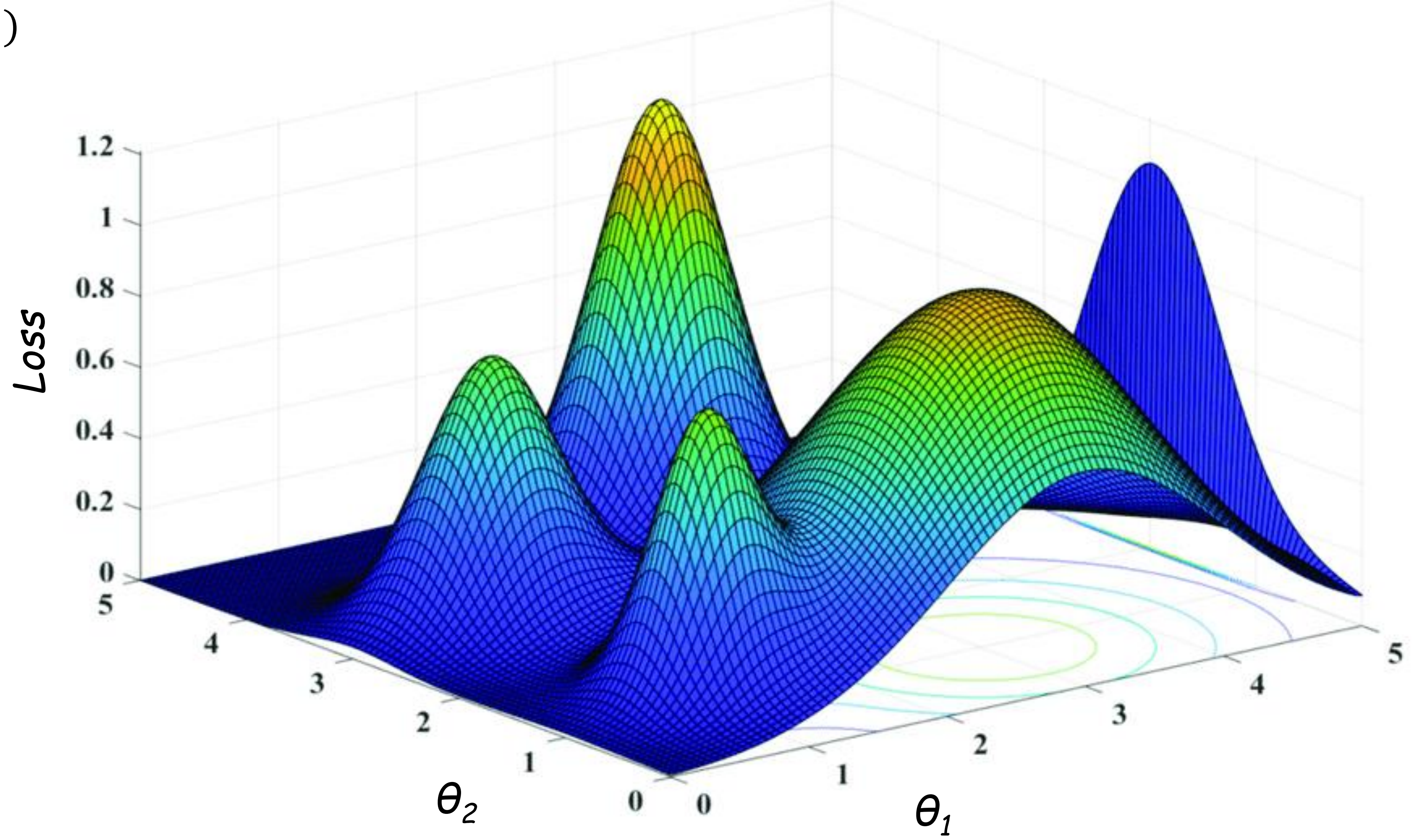
- might take a **very long time**
- might end up **stuck** in a **local minimum**

If **α** is **too big**:

- might have **trouble converging**
- might **skip global minimum**

If **α** is **just right**:

- willl converge **quickly**
- will **evade local minima**

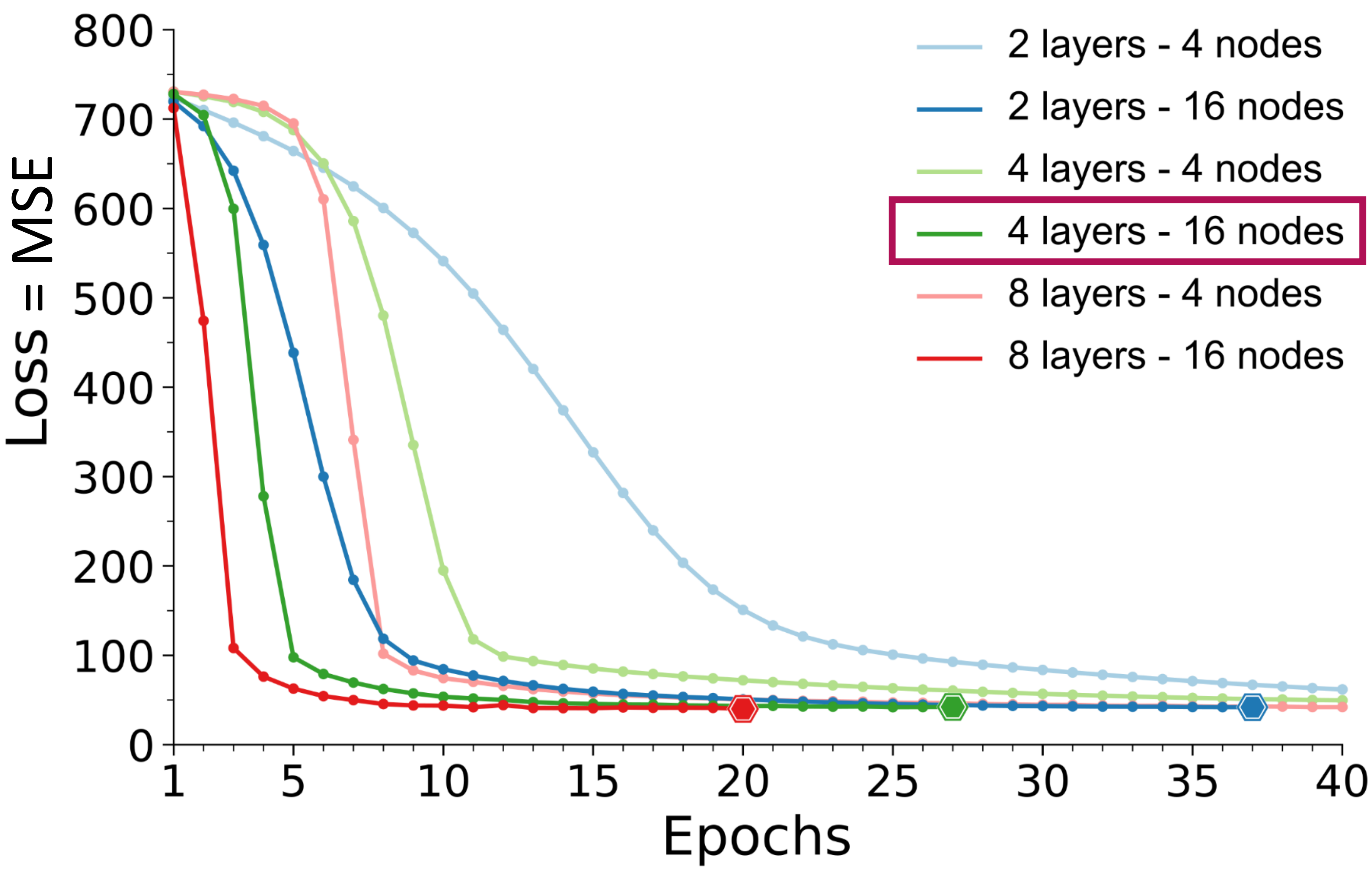


Hyperparameters tuning: network structure

- The lightest configuration **(2 layers – 4 nodes, 37 learnable parameters) doesn't reach convergence** within 40 epochs

- The heaviest configuration **(8 layers – 16 nodes) converges quickly** but at the expense of the number of learnable parameters **(1969)**

- The middle-ground configuration **(4 layers – 16 nodes)** is a **good balance** between **converging speed (27 epochs)** and **complexity (881 learnable parameters)**



Hyperparameters tuning: batch size

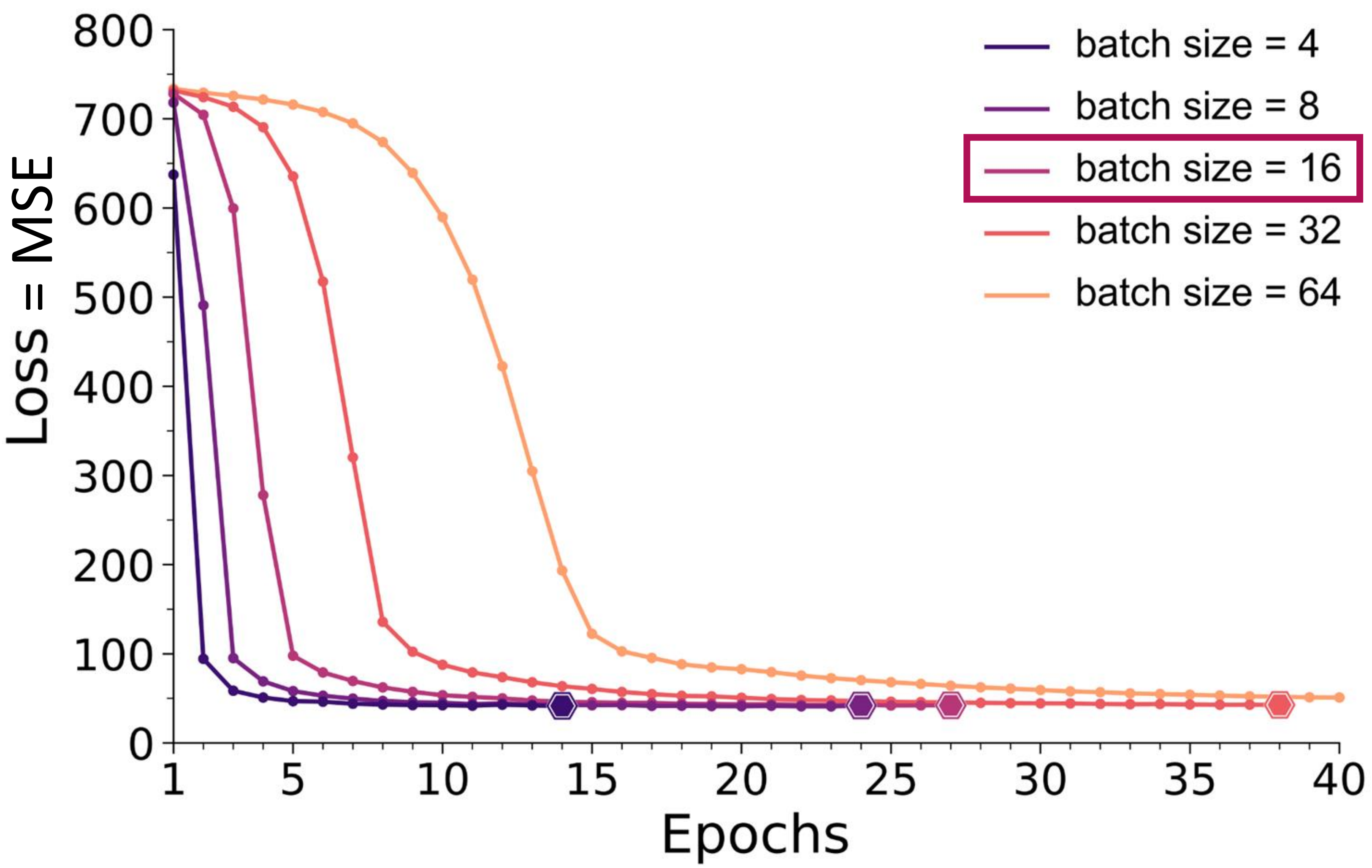
- Batch size follows a non-linear **direct relation** with convergence speed (**in epochs**)

  Behavior **not reflected** in total computational time:

  t(bs = 4) [s] > 2x t(bs = 64) [s]

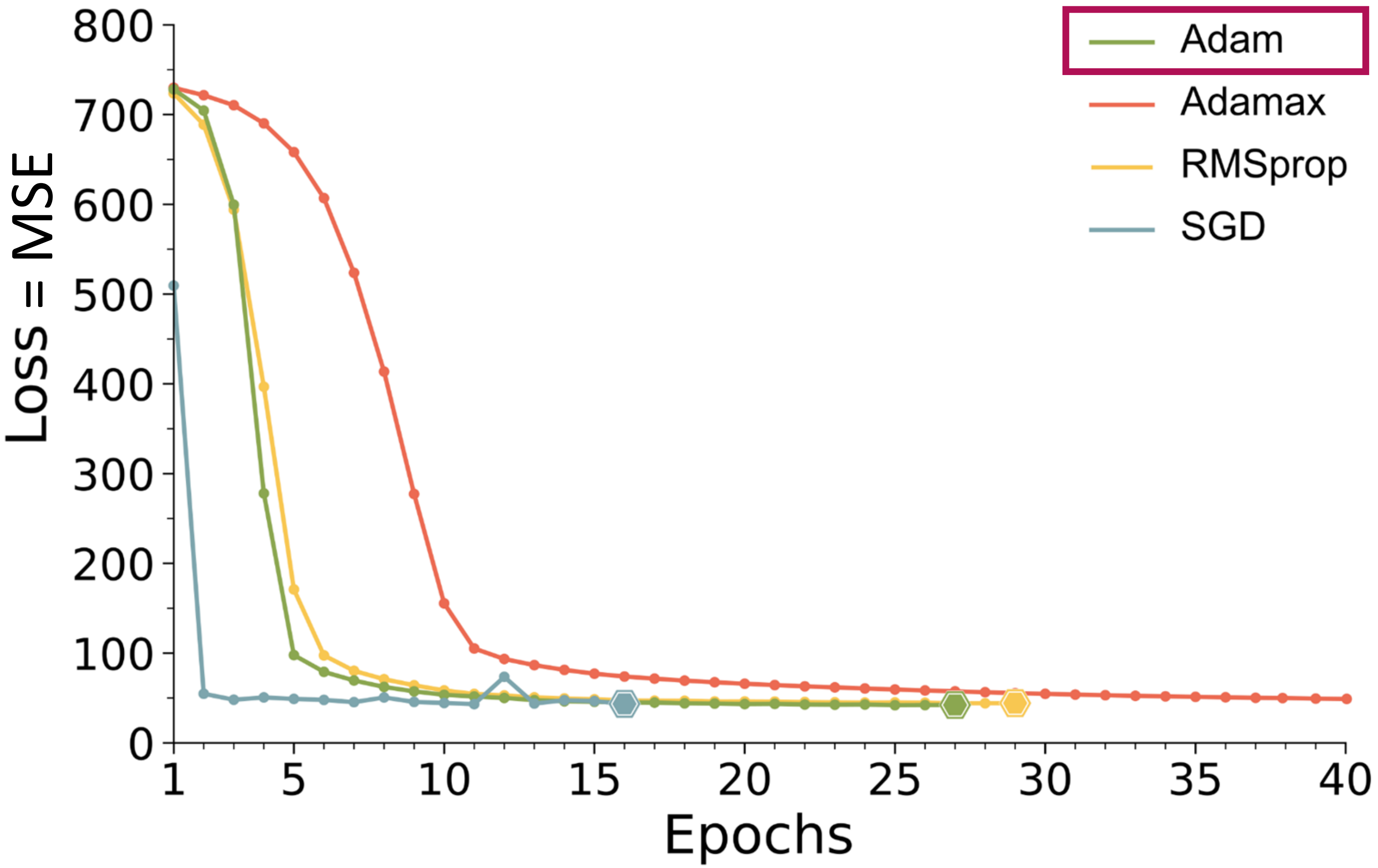
- Large batch size values **compromise generalizing ability** of the network

- The middle-ground value of 16 keeps computational times at reasonable values while preventing loss of generalizability



Hyperparameters tuning: optimizer

- The fastest convergence is shown by **SGD** (**16 epochs**), which presents **instabilities** in the loss values with a visible bump at epoch 12
- The second-best optimizer in terms of convergence speed (**27 epochs**) is **Adam**, which converges in a **stable** manner
- Other optimizers tested (Adadelta, Adagrad, Ftrl and Adafactor, not shown here) **did not converge**

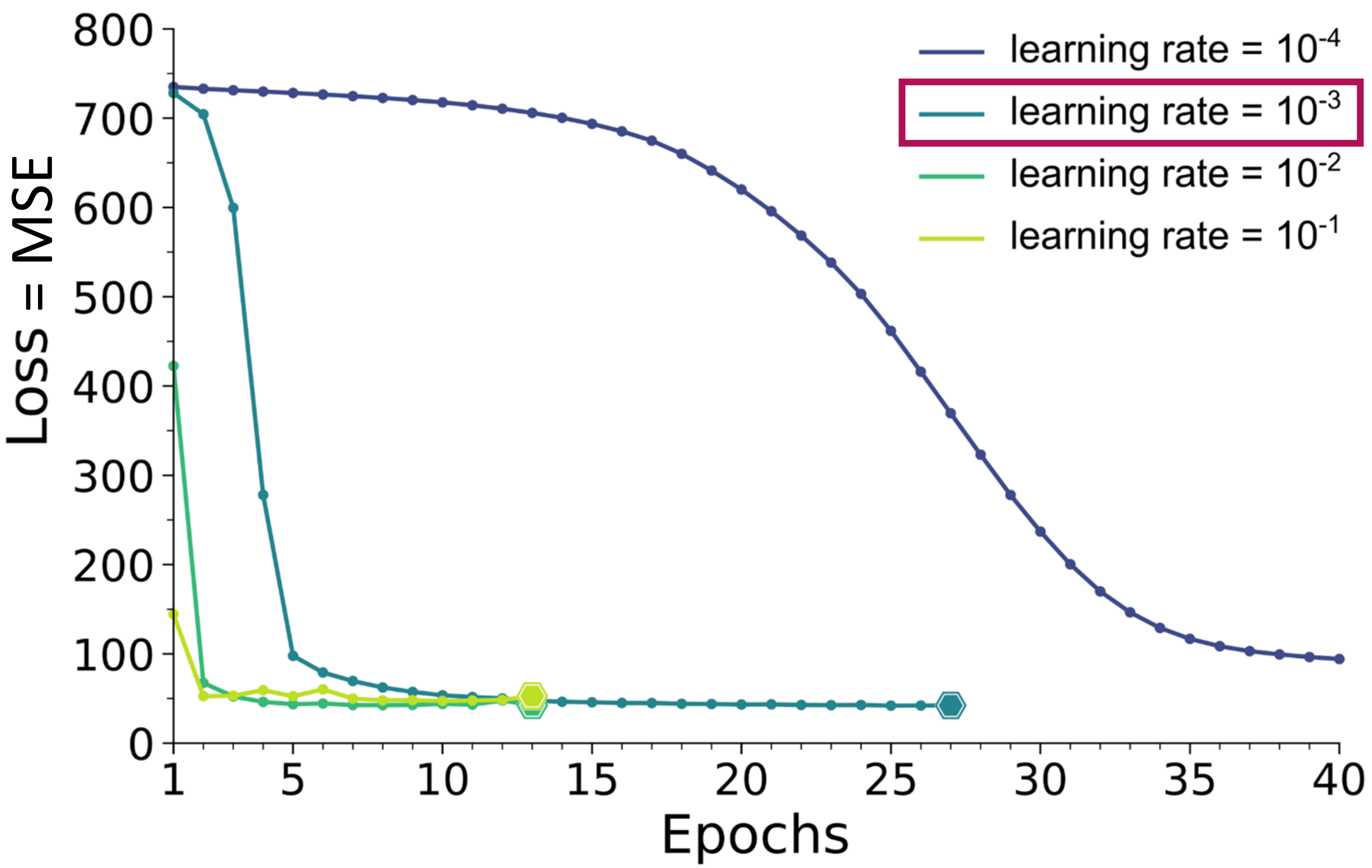


Hyperparameters tuning: learning rate

- Learning rate values of 10^{-1} and 10^{-2} show **fast convergence** (**13 epochs**) but unstable loss curves, with the 10^{-2} value showing a bump at epoch 12

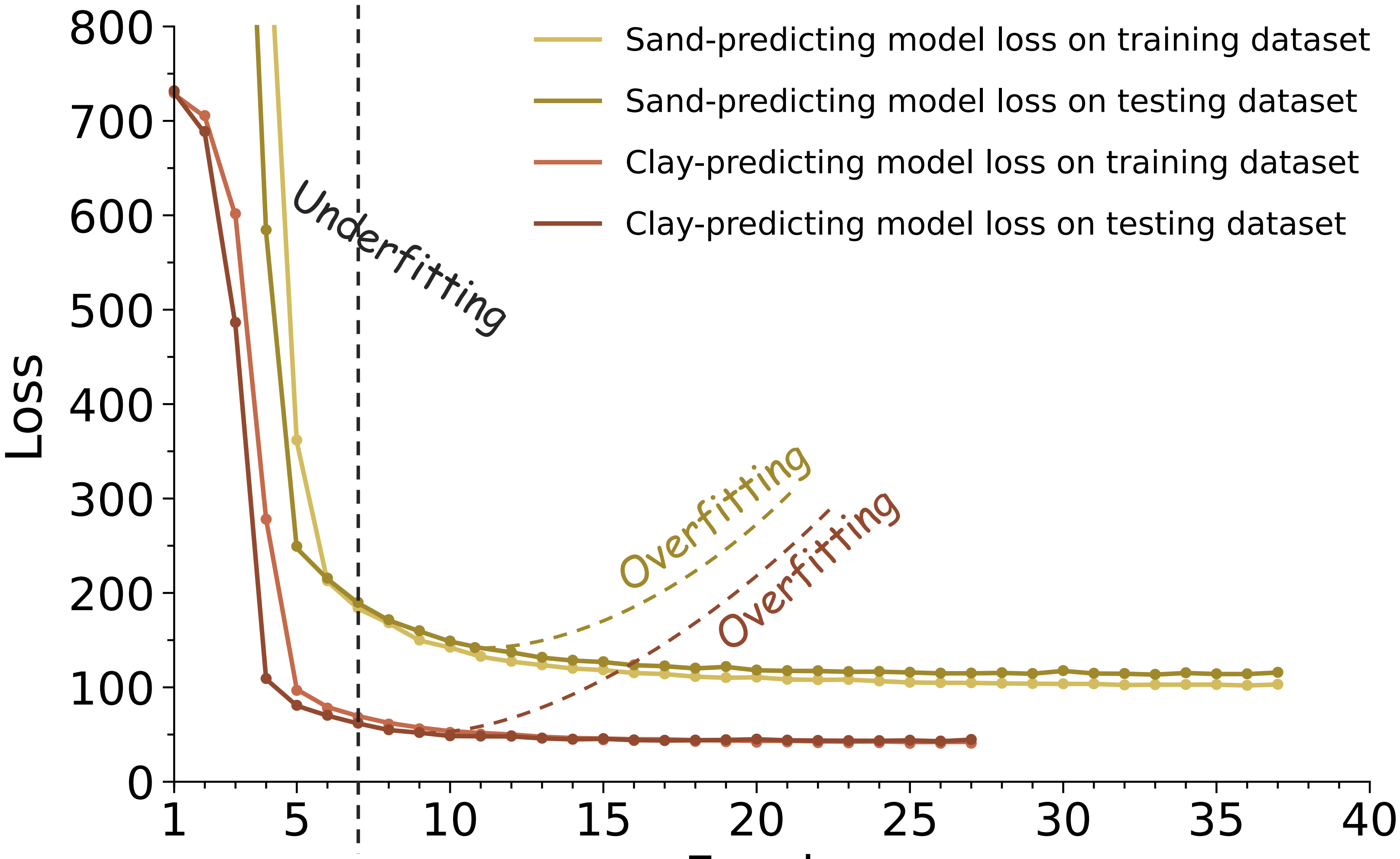
- The learning rate value of 10^{-4} **doesn't reach convergence** in the 40 epochs limit

- The best choice is therefore the value 10^{-3}, which reaches convergence at **epoch 27**



Final configuration of the network

- **Layer density** = 16
- **Layer number** = 4
- **Optimization parameter (Loss)** = Mean Squared Error (MSE)
- **Optimizer** = Adam
- **Batch size** = 16
- **Learning rate** = 10^{-3}
- **Activation function**
  - ‘**ReLU**’ for the 4 hidden layers
  - ‘**Linear**’ for the output layer



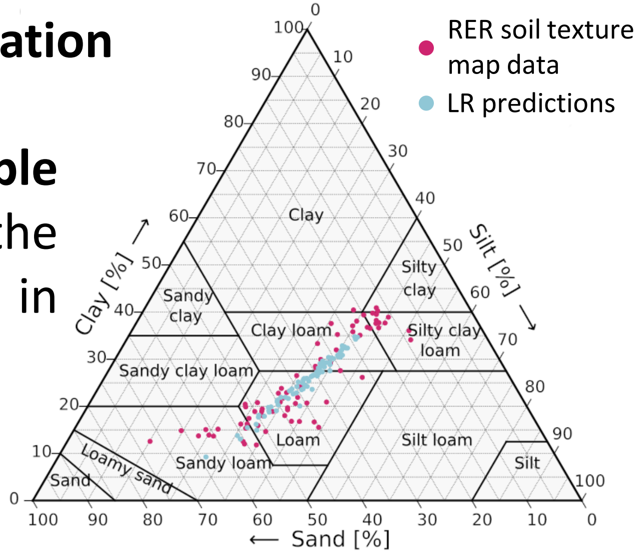
sand content predictions **CV(RMSE) ~ 0.28**
clay content predictions **CV(RMSE) ~ 0.25**

Advantages of Machine Learning

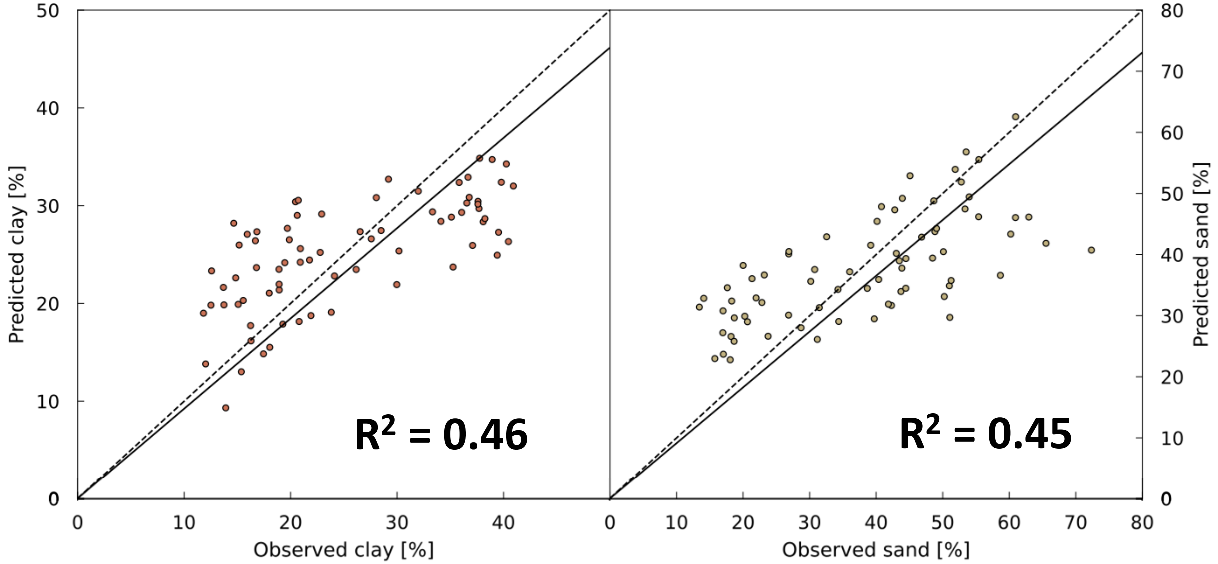
Linear Regression (LR)

- Bound to a **linear relation**

  Predictions are **unable to reproduce** the complexity of data in the textural triangle



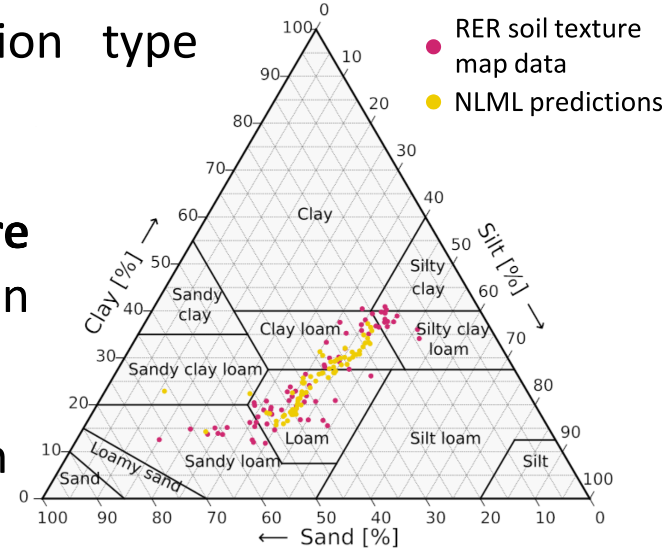
- **Worse R^2 scores** on predictions



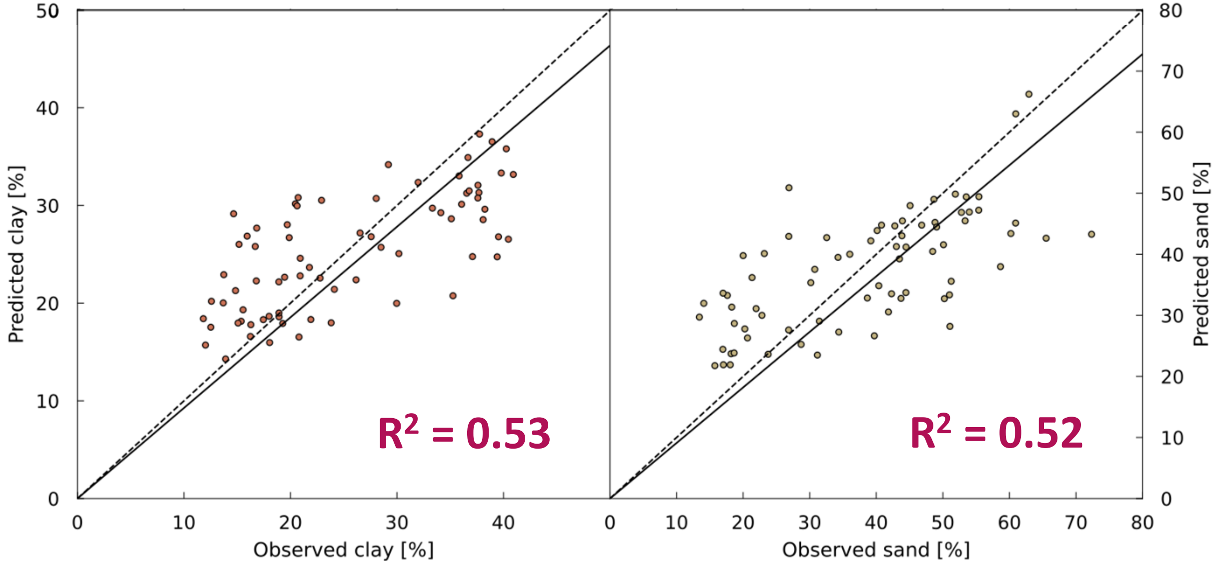
Non-Linear Machine Learning (NLML)

- **Free** of any relation type restriction

  Predictions **more closely** follow data in the textural triangle

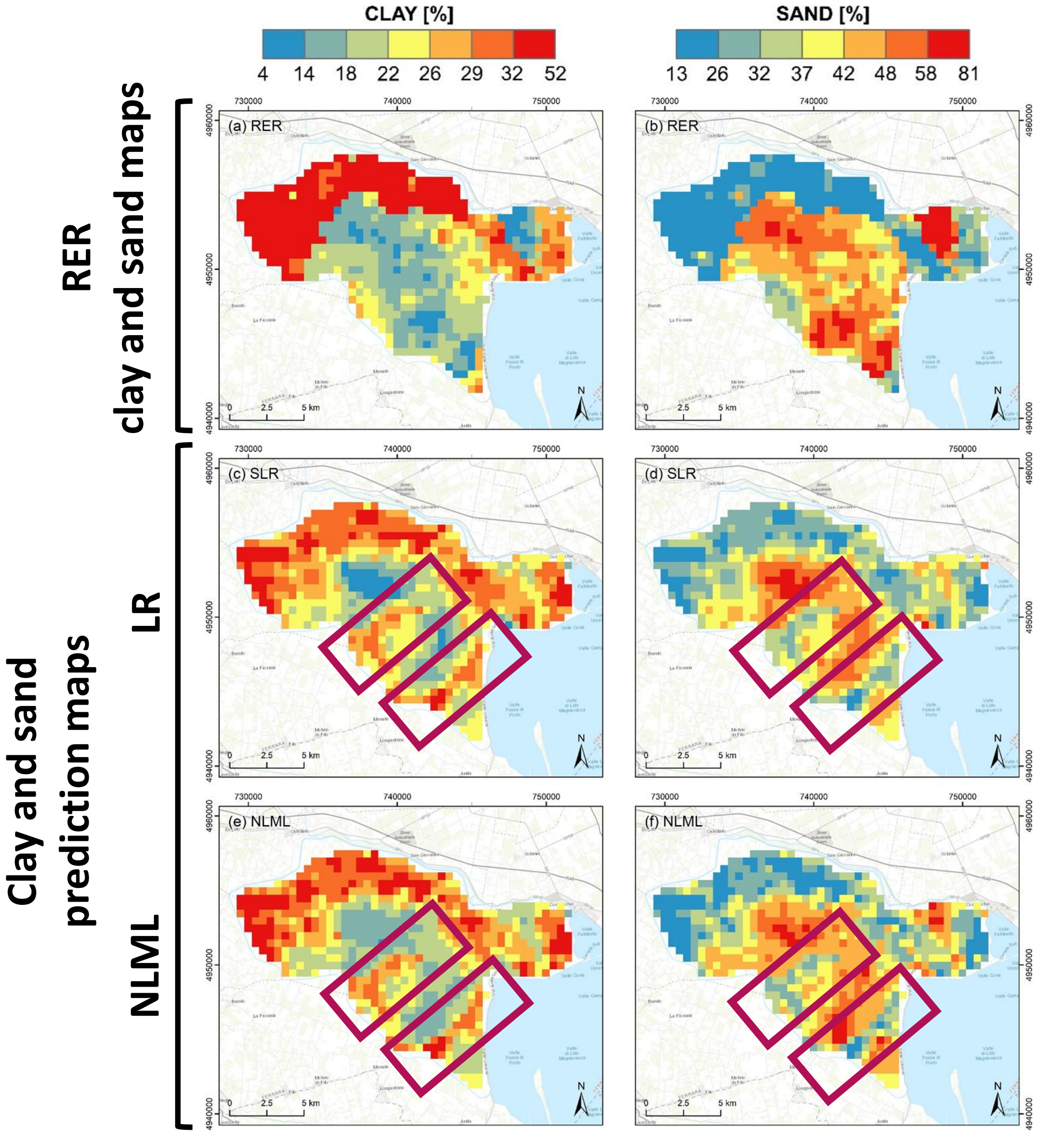


- **Better R^2 scores** on predictions

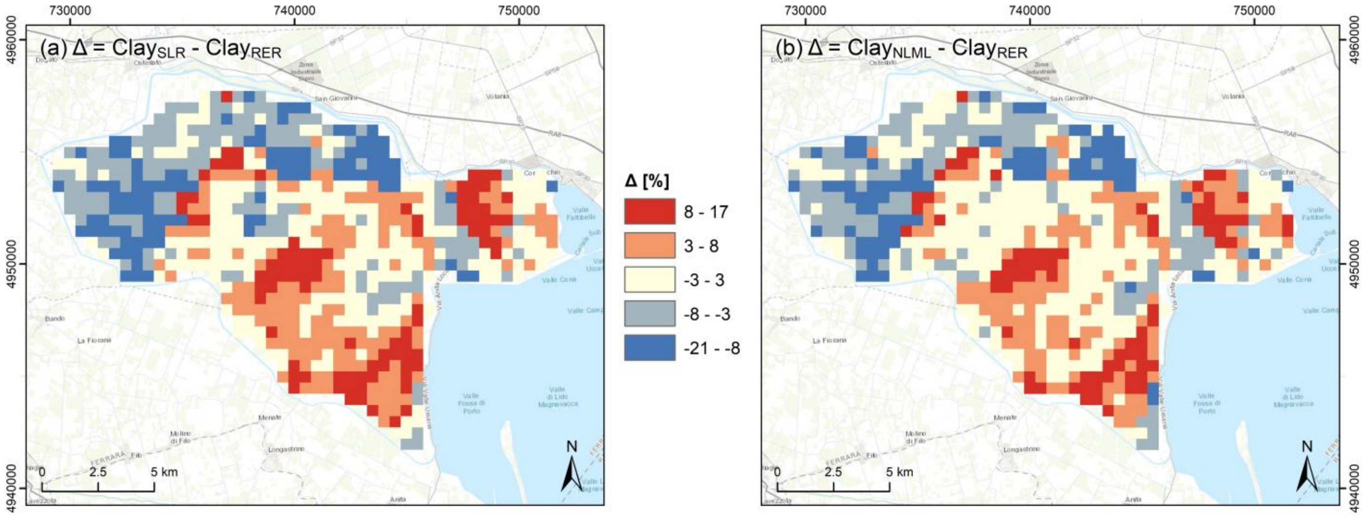


Comparing soil texture maps

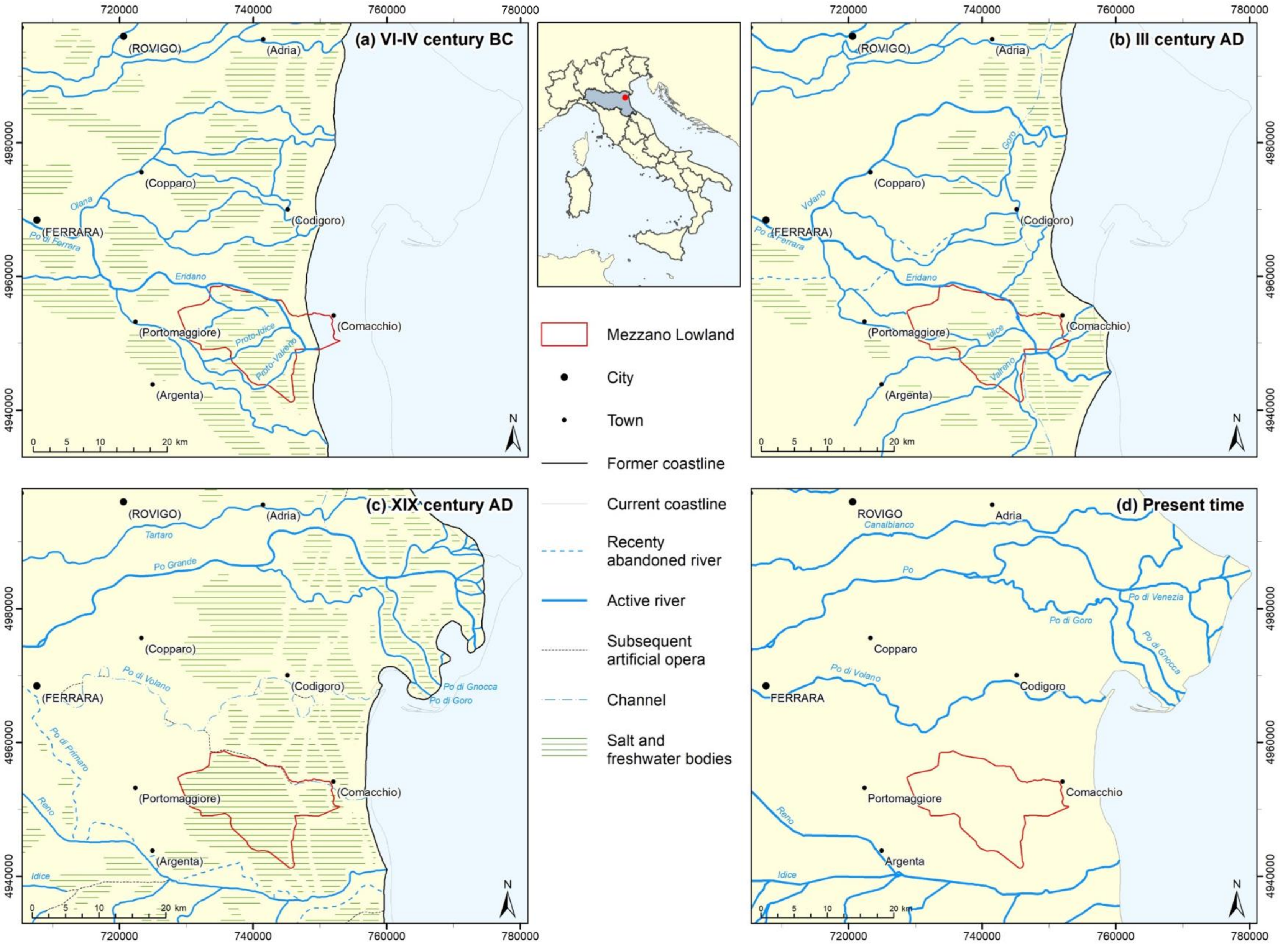
- Soil texture prediction maps show the **same macrostructures** present in the RER clay and sand maps.
- Main differences are two **high clay and low sand content narrow-shaped features** not shown by RER maps.



$\Delta = Clay_{SLR} - Clay_{RER}$ $\quad$ $\Delta = Clay_{NLML} - Clay_{RER}$

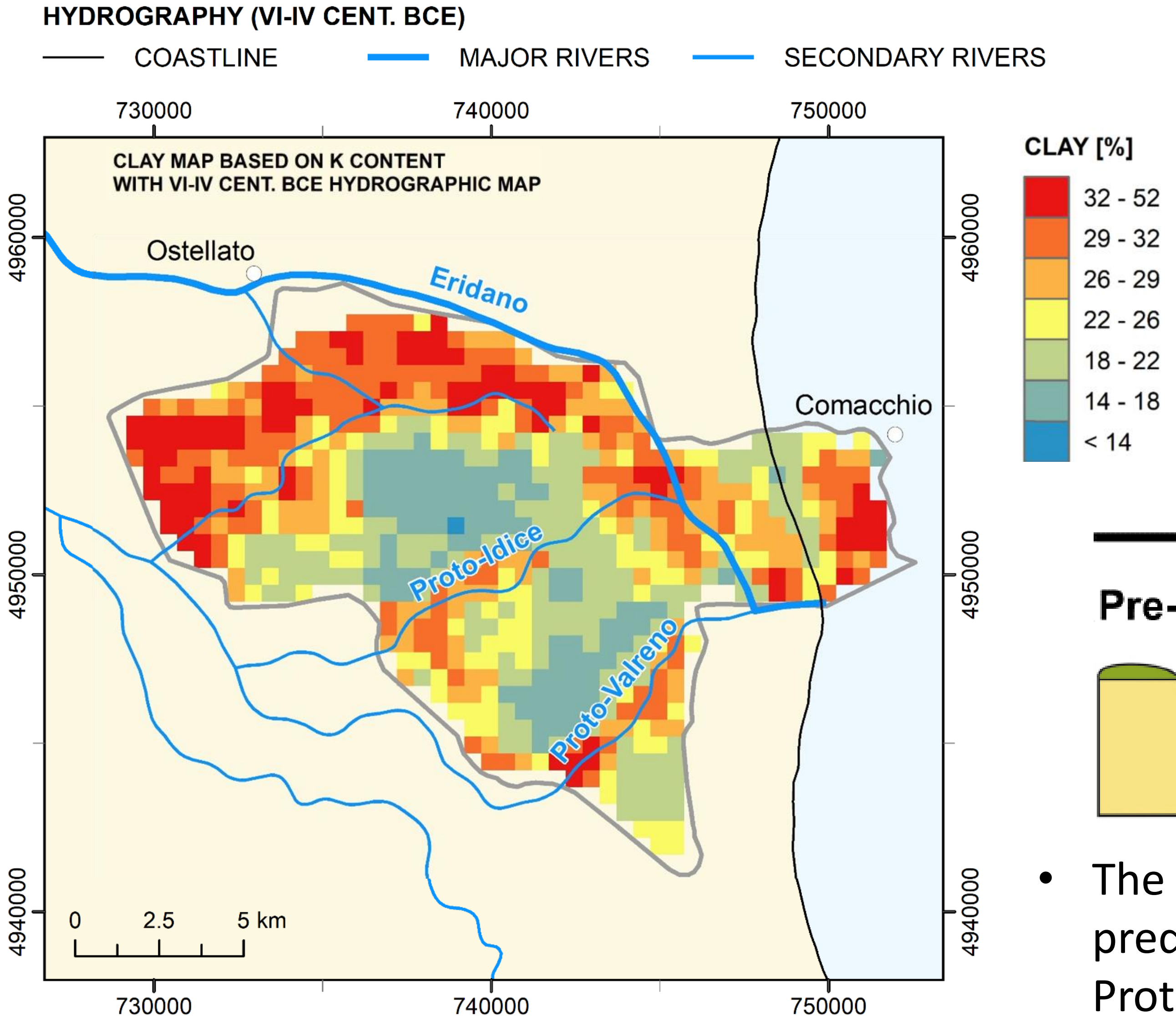


Hydrographic history of the area

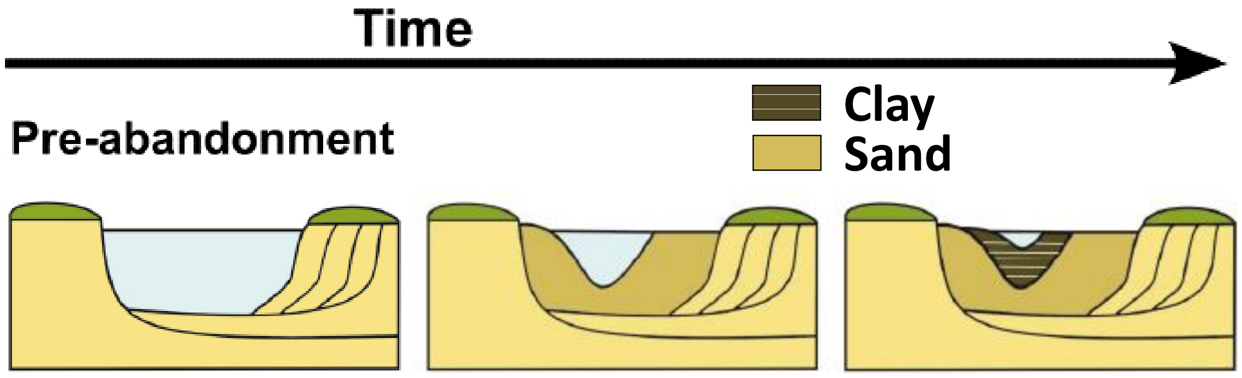


- The area was crossed until the III century BCE by the **Eridano, Idice** and **Valreno** channels branching off from the Po river.
- Subsequent hydrographic changes to the landscape led to the **abandonment** of the area by those river channels

Sedimentary filling

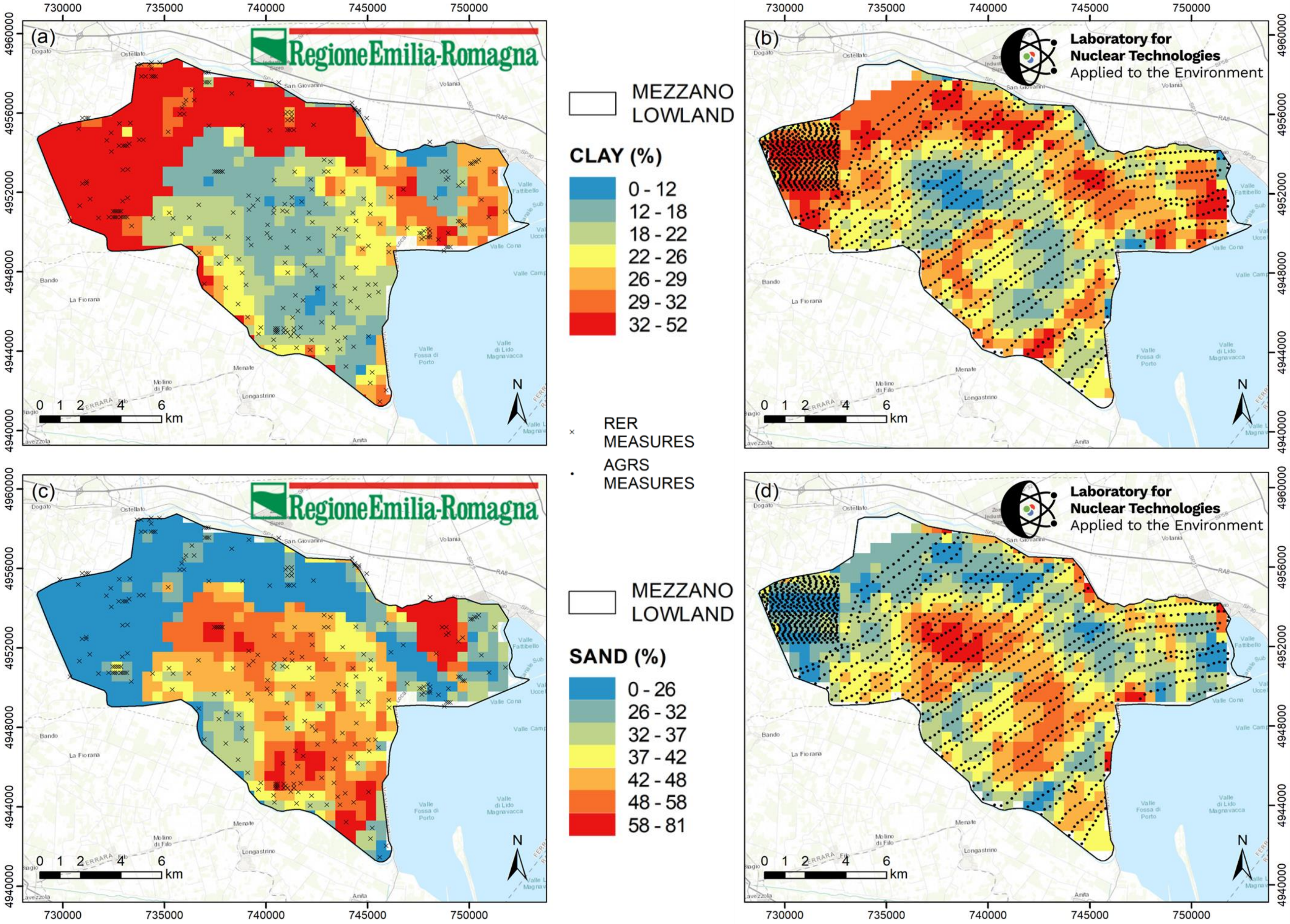


- The abandonment of the Eridano, Proto-Idice and Proto-Valreno channels led to the **sedimentary filling of their riverbeds**.
- During sedimentary filling, coarser grained particles like **sand** are deposited first, gradually followed by finer grained particles like **clay**.



- The **narrow-shaped features** present in the soil texture prediction maps **retrace the abandoned riverbeds** of Proto-Idice and Proto-Valreno.

Advantages of AGRS measurements



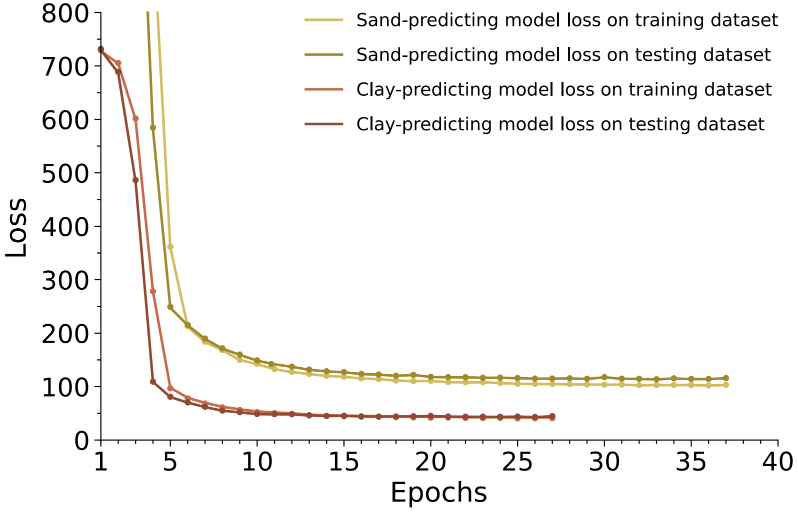
- Direct RER measurements are taken **unevenly** in the surveyed area.
- A total of **273** measurements were taken by RER.
- AGRS cover is much more **homogeneous**.
- **1469** gamma spectra were acquired, more than **5x** the amount of RER measurements.

Take home messages

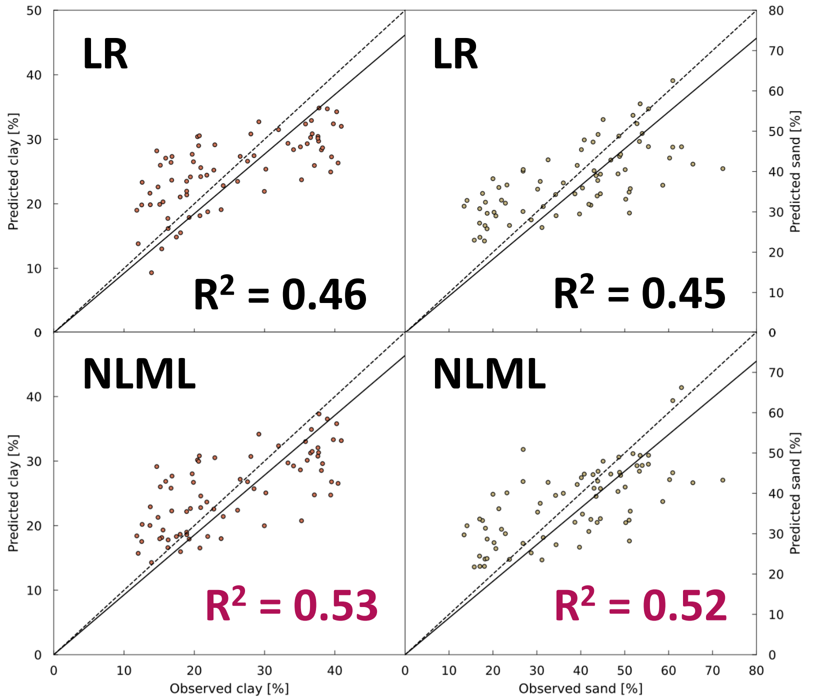
Hyperparameter optimization is key in Deep Learning to ensure **fast convergence and high-quality results**

↓

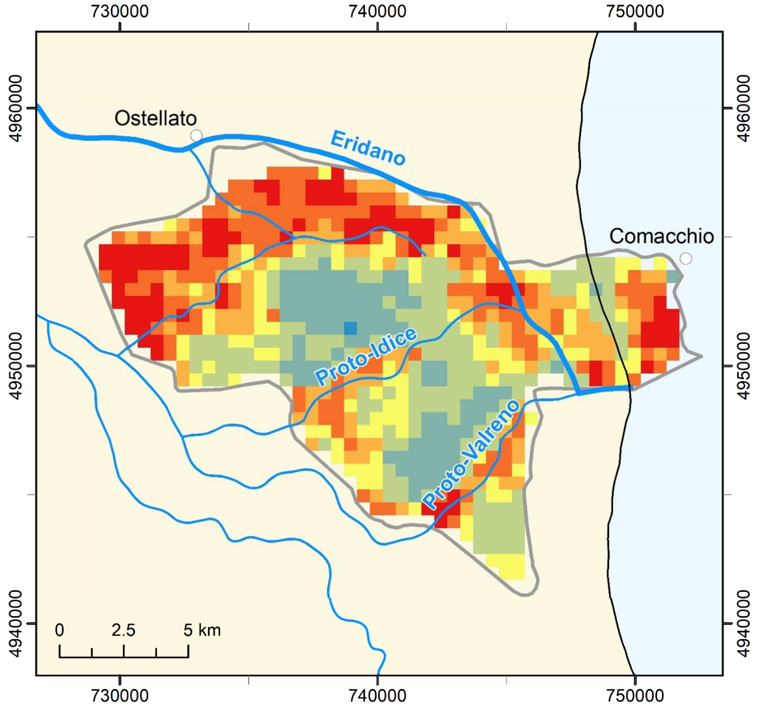
I **prevented** both **underfitting and overfitting**



Soil texture and radioelements abundances are **non-linearly correlated**, and are therefore best studied with **Deep Learning**



Gamma rays can unveil data about **soil texture** and reveal **ancient traces** left by the hydrographic history of the soil



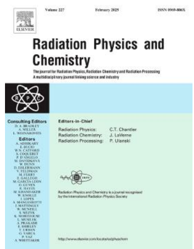
Thesis summary

- The **electromagnetic spectrum can be fully exploited** to obtain valuable agricultural data related to:
  - **soil properties**
  - **plant health status**
  - **water levels**

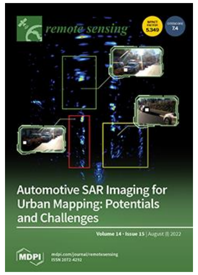
- Sensing technologies can **accurately measure** agricultural indicators, providing farmers and organization with the necessary knowledge to **increase production quality and yield while reducing resource wastes**

- **Remote sensing** plays a major role in increasing survey efficiency in terms of **time, costs** and **data quality**

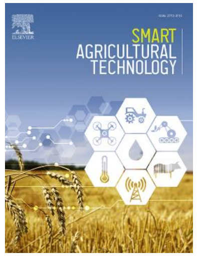
- **Artificial Intelligence** can enhance the study of **complex systems** and will be a **powerful tool** for smart agriculture going forward



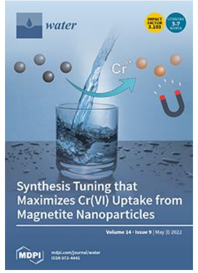
A deep neural network for predicting soil texture using airborne radiometric data



Airborne Radiometric Surveys and Machine Learning Algorithms for Revealing Soil Texture



Moving forward the automatic detection of Flavescence Dorée in vineyards with airborne imaging



Combining Precision Viticulture Technologies and Economic Indices to Sustainable Water Use Management

My publications

Peer reviewed papers:

- **Maino, A.** et al. *A deep neural network for predicting soil texture using airborne radiometric data.* Radiation Physics and Chemistry 2024, 221, 111767 DOI 10.1016/j.radphyschem.2024.111767
- Coppi, A. et al. *Mass testing of the JUNO experiment 20-inch PMT readout electronics.* Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment 2023, 1052, 168255 DOI 10.1016/j.nima.2023.168255
- Cerrone, V. et al. *Validation and integration tests of the JUNO 20-inch PMT readout electronics.* Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment 2023, 1053, 168322 DOI 10.1016/j.nima.2023.168322
- Triozzi, R. et al. *Implementation and performances of the IPbus protocol for the JUNO Large-PMT readout electronics.* Nuclear Instruments & Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment 2023, 1053, 168339 DOI 10.1016/j.nima.2023.168339
- **Maino, A.** et al. *Airborne Radiometric Surveys and Machine Learning Algorithms for Revealing Soil Texture.* Remote Sensing 2022, 14, 3814 DOI 10.3390/rs14153814
- Raptis, K.G.C. et al. *External effective dose from natural radiation for the Umbria region (Italy).* Journal of Maps 2022 DOI 10.1080/17445647.2022.2093659
- Finco, A. et al. *Combining Precision Viticulture Technologies and Economic Indices to Sustainable Water Use Management.* Water 2022, 14, 1493 DOI [illegible]
- Serafini, A. et al. *Proximal Gamma-Ray Spectroscopy: an Effective Tool to Discern Rain from Irrigation.* Remote Sensing 2021, 13(20) DOI 10.3390/rs13204103
- Torri, M.D.C. et al. *Predictions of Ultra-High Energy Cosmic Ray Propagation in the Context of Homogeneously Modified Special Relativity.* Symmetry 2020, 12(12), 1961 DOI 10.3390/sym12121961

Conference proceedings:

- Strati, V. et al. *Airborne surveys for the detection of Flavescence Dorée in vineyards.* EGU24-10773. EGU General Assembly (2024).
- Francheschi, M. et al. *Proximal Gamma Ray Spectroscopy for monitoring Soil Water Content in vineyards.* EGU24-908. EGU General Assembly (2024).
- Natali, M. et al. *Irrigation Estimation from Soil Water Balance and the Water Cloud Model by leveraging Sentinel-1 and Sentinel-2 observations.* EGU24-18745. EGU General Assembly (2024).
- Lopane, N. et al. *Geoelectric joint inversion: a novel approach for grape vineyards investigation.* Agrogeo24, Zürich (Switzerland), 1-2 February 2024.
- Natali, M. et al. *A Simple Framework to Calibrate a Soil Water Balance Model With Sentinel-1 and Sentinel-2 Observations Over Irrigated Fields.* 2023 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), Pisa (Italy), 6-8 November 2023.
- **Maino, A.** et al. *Soil texture predictions through Machine Learning from airborne radiometric data.* IRRMA11-8320. IRRMA11 (2023).
- Strati, V. et al. *BlueSky: a system for in-situ identification of 137Cs in industrial waste.* EGU[illegible]2800. EGU General Assembly (2023).
- **Maino, A.** et al. *Mapping soil texture with airborne gamma ray spectroscopy.* EGU22-361. EGU General Assembly (2022).
- Albéri, M. et al. *RadHawk: a smart UAV for hunting radioactivity.* EGU22-11835. EGU General Assembly (2022).
- Mo[illegible]uschi, M. et al. *A Web GIS tool for 3D visualization of bathymetric data.* EGU22-11828. EGU General Assembly (2022).
- Serafini, A. et al. *Mapping the outdoor effective dose: the case study of the Umbria region (Italy).* EGU2021-7284, EGU General Assembly (2021).

Thank you!