

Where to find the presentation

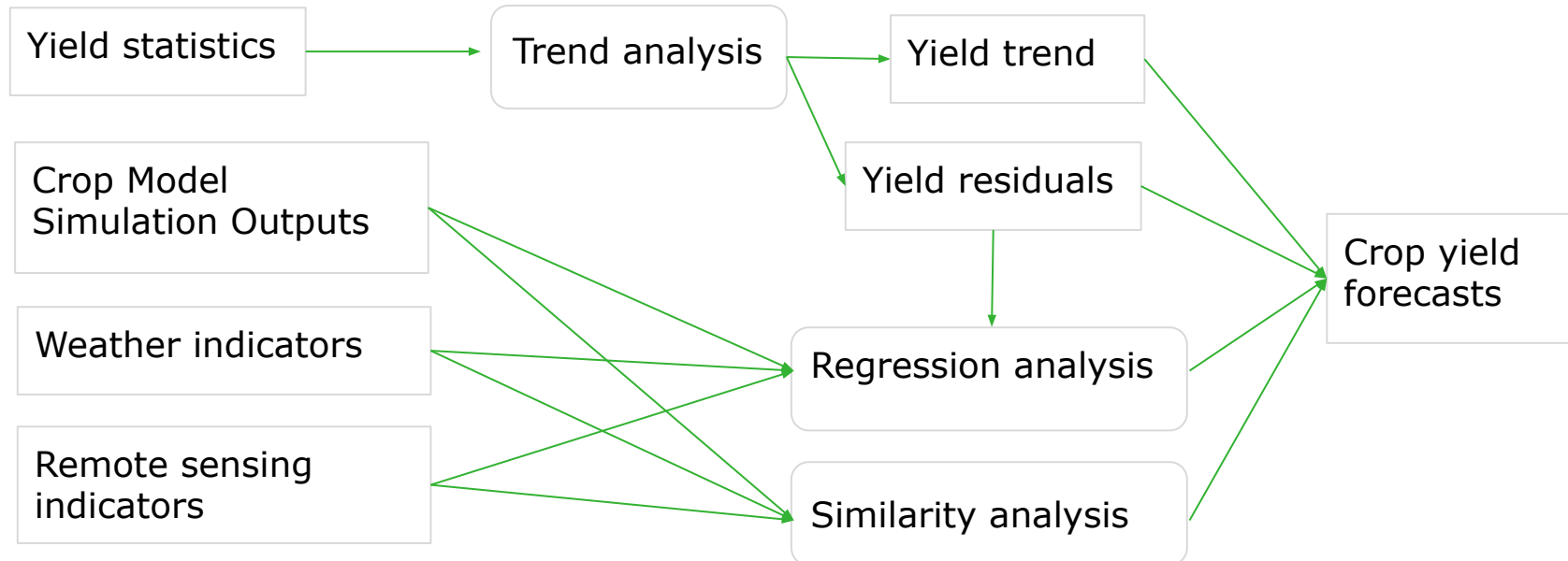
<https://drive.google.com/drive/folders/17UT5YDxyAZluaYj37Il6r7fyqwG12Prp?usp=sharing>

Background

Crop yield forecasts are useful to many stakeholders (e.g. farmers, commodity traders, agronomists, policymakers).

European Commission's Joint Research Centre uses the **MARS Crop Yield Forecasting System** - a suite of tools for analysts - to provide regular forecasts to policymakers at DG-AGRI and member states.

MCYFS forecasts are based on statistical regression analysis, similarity analysis and expert judgment.



Problem Definition

Predict crop yield using

- outputs of crop growth simulation models (e.g. WOFOST)
- weather observations
- remote sensing indicators
- soil
- regional yield statistics

Data from **MARS Crop Yield Forecasting System** and **Eurostat**.

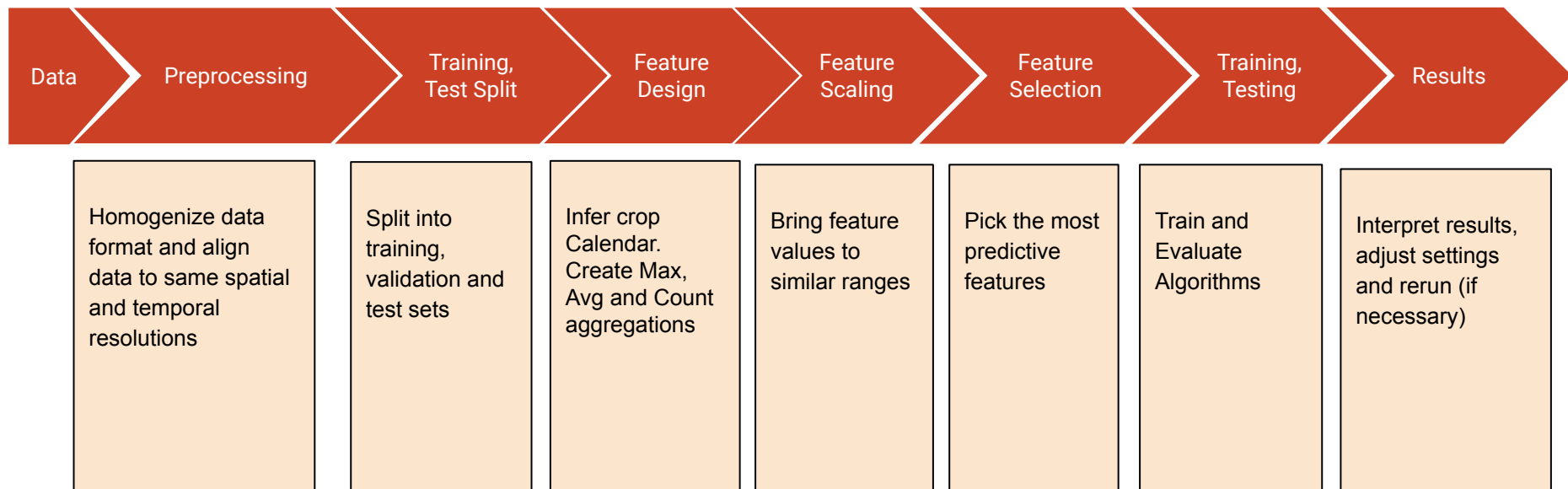
Design a workflow to explore the potential of **Machine Learning** in large-scale crop yield forecasting.

- Run the same workflow across Europe for several crops.
- Predict at different points during the crop growing season.
- Make yield predictions at regional level and aggregate to national level to compare with MCYFS predictions.
- Analyze regional patterns seen in regional forecasts.

Machine Learning Baseline

Baseline: A machine learning **workflow** focused on **correctness**, **modularity**, **reusability**.

Easy to run: Something like **Auto ML** to get **baseline results**.

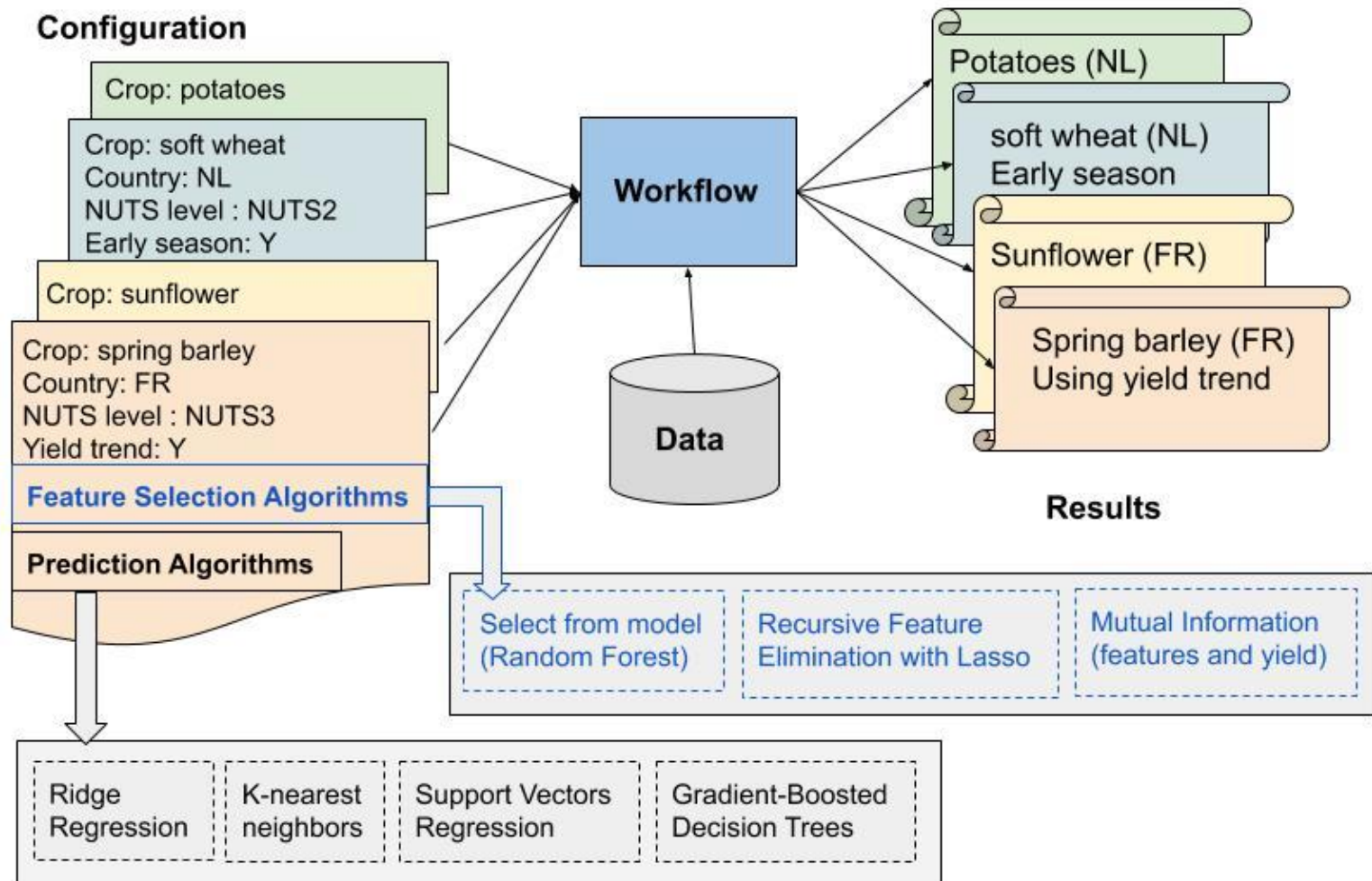


Feature Design

Using WOFOST data, we inferred crop calendar periods for each crop. For each period, we identified the indicators that affect or capture the state of crop growth and development.

Period	Maximum Values	Averages, *Cumulative Sums, **from previous year	z-scores based on long term avg, std
Pre-planting window		TAVG, PREC, CWB*	
Planting window		TAVG, PREC	RSM, TMIN, PREC
Vegetative phase	WLIM_YB, TWC, WLAI	RSM, TAVG, CWB*, FAPAR, RAD	RSM
Flowering phase		PREC	RSM, PREC, TMAX
Yield Formation phase	WLIM_YB, WLIM_YS, TWC, WLAI	RSM, CWB*, FAPAR, RAD	RSM
Harvest window		PREC**	PREC**

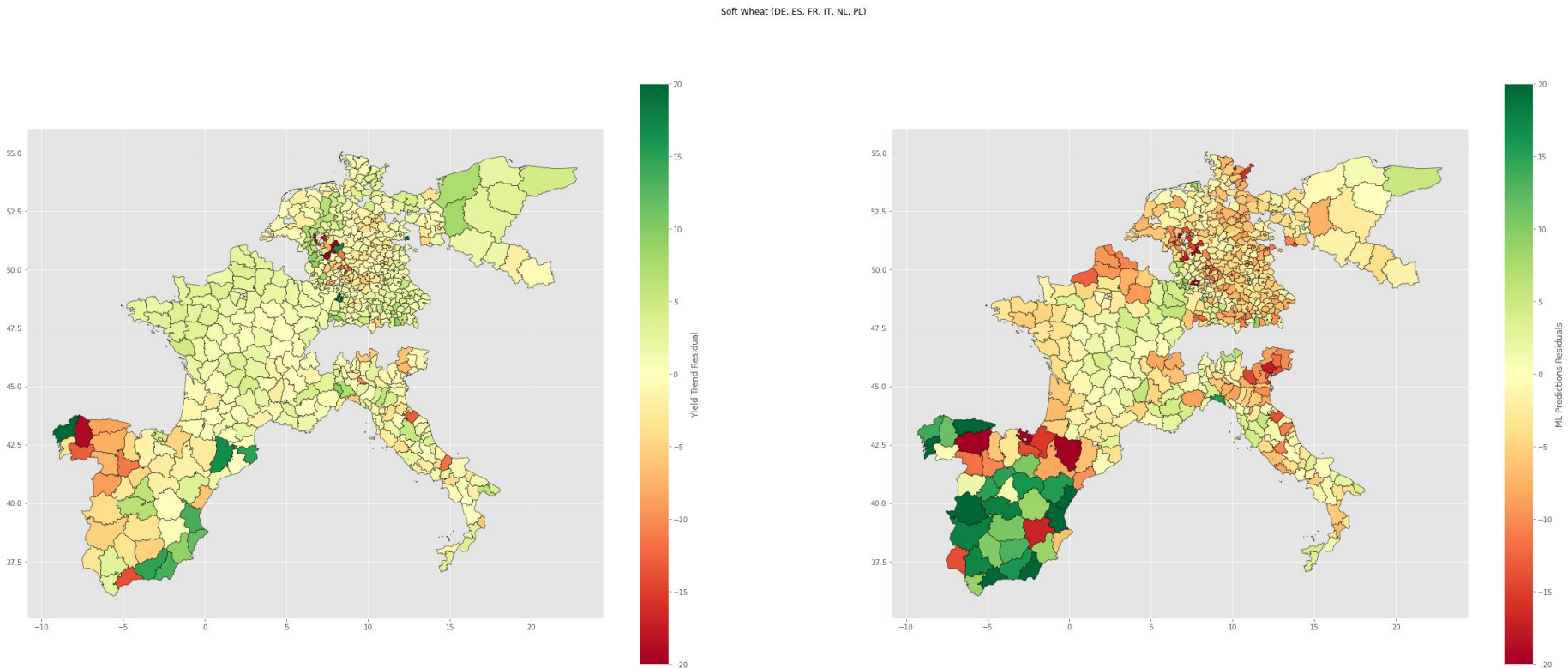
Modularity and Reusability



Visualizations for regional data, predictions

Analyzing predictions and residuals

<https://colab.research.google.com/drive/1fuCvXoGASnaBCoEdP53luc30veGJmH6G#scrollTo=YK7UzM50xQ9k>



For interested participants

- Change crop, end of season dekad
- Compare performance of an additional algorithm (use `mlbaseline.ipynb`)
- Use Deep Learning. Use `crop_yield_DL_performance.ipynb` from `dlinterpret` branch)

