

Crop Yield Prediction

Aziza Kurbonova, Florjon Haxhija

Defining the Problem

Problem Space

- Informal: Assist US-based farmers in making an informed decision on what crop to grow and when to grow it
- Formal: Regression problem where the goal is to build a model that, given a set of input features (e.g. soil properties, climate conditions), predicts crop yield from which a farmer can base their farming strategy around

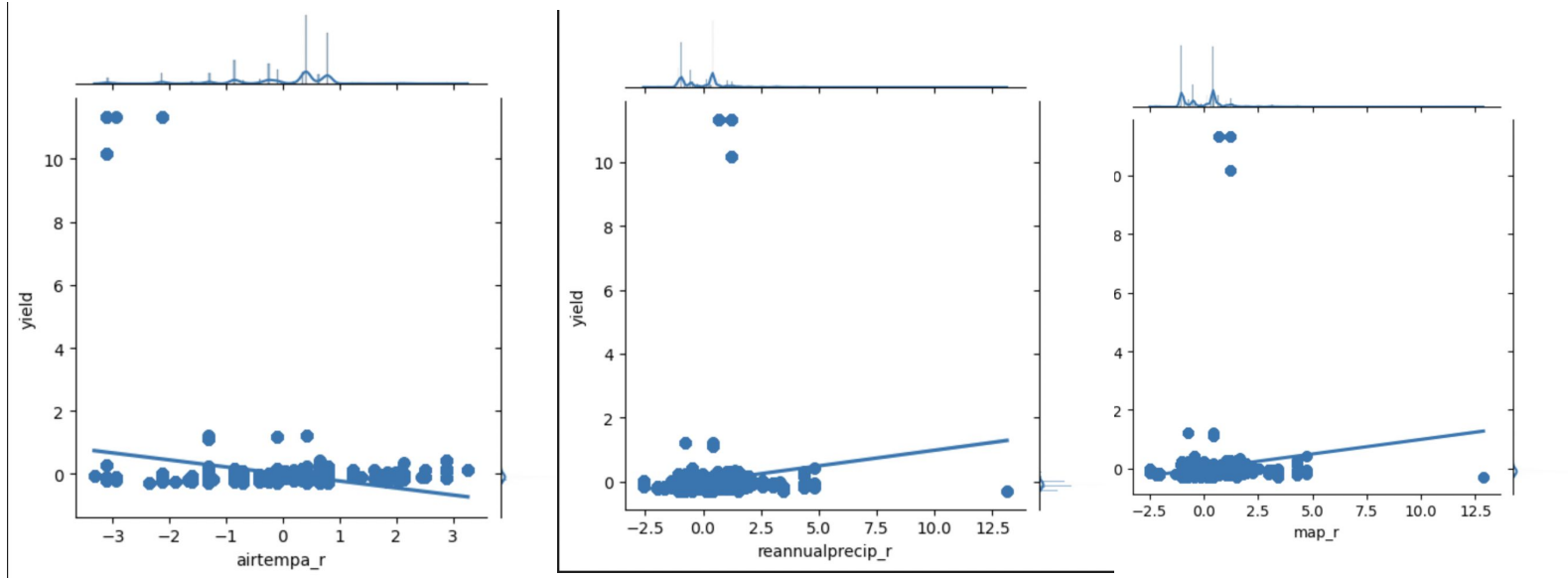
Need for ML:

- It's difficult for a farmer to achieve optimum crop yield as a result of changing environmental conditions that add unpredictability to crop yield estimations, against the backdrop of climate change
- Complex interrelated environmental, soil, and land management factors make crop yield prediction a difficult manual task, necessitating machine learning techniques

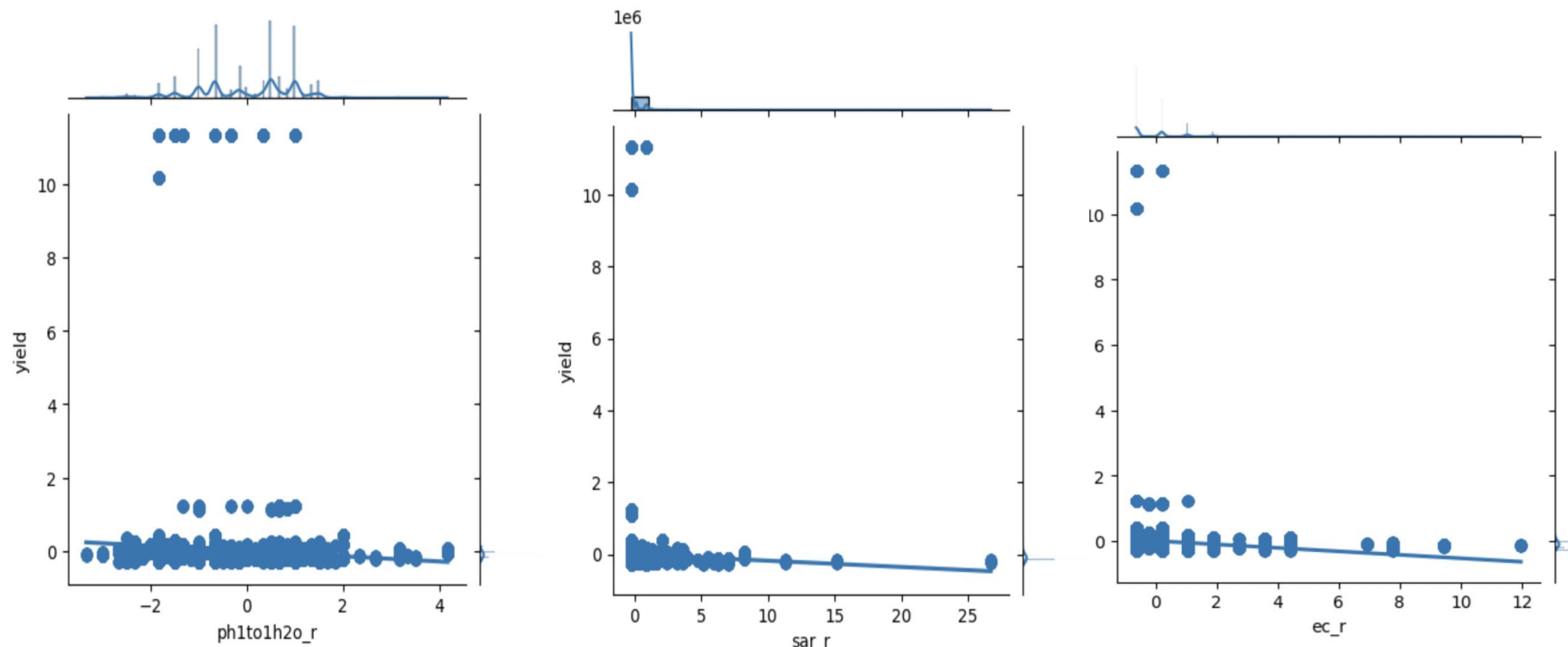
Soil Survey Geographic (SSURGO) Database

Climate / Weather	Mean Annual Air Temperature	Mean Annual Precipitation	Moisture Availability for Plant Use and/or Soil Forming Processes
	Total Silt Composition	Total Clay Composition	Soil Horizon Depth
Soil Physical Properties	Available Water Capacity		
	Gypsum (Calcium Sulfate) Content	Calcium Carbonate Content	Cation-Exchange Capacity (CEC-7)
Soil Chemical Properties	Sodium Adsorption Ratio (SAR)	Saturated Hydraulic Conductivity (KSAT) - Rate at which water flows through soil	Electrical Conductivity
	pH		
	Crop Name	Crop Yield (Irrigated Yield, Non Irrigated Yield)	Crop Yield Units (e.g. BU, Ton, AUM)
Crop	Month		

Data Visualization



Data Visualization



Model Comparisons

