Transforming real needs to research and business value

the digital agriculture case

presenters:
Alkiviadis Koukos, akoukos@noa.gr
Ilias Tsoumas, i.tsoumas@noa.gr
July 2022

Transforming real needs to research and business value

the digital agriculture case

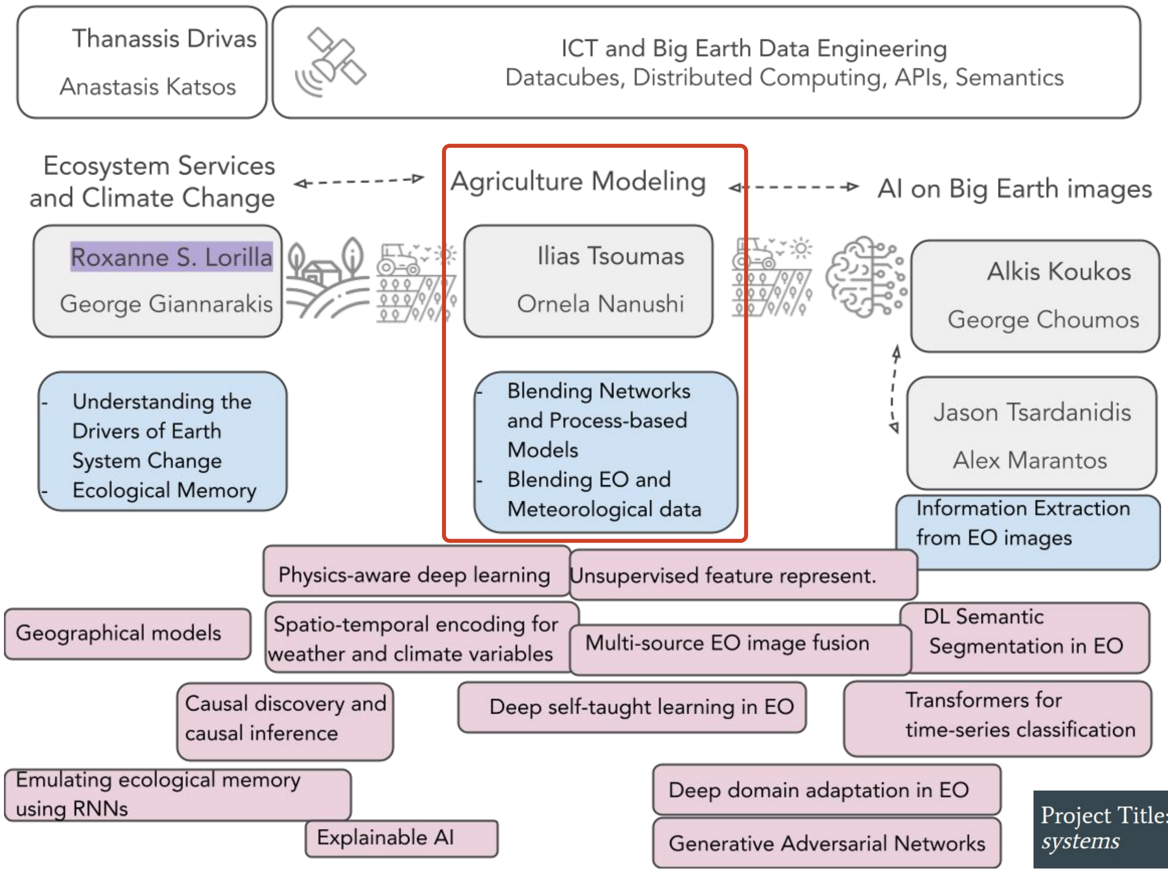
presenters:
Alkiviadis Koukos, akoukos@noa.gr
Ilias Tsoumas, i.tsoumas@noa.gr
July 2022

Transforming real needs to research and business value

the digital agriculture case

presenters:
Alkiviadis Koukos, akoukos@noa.gr
Ilias Tsoumas, i.tsoumas@noa.gr
July 2022

who, what, why

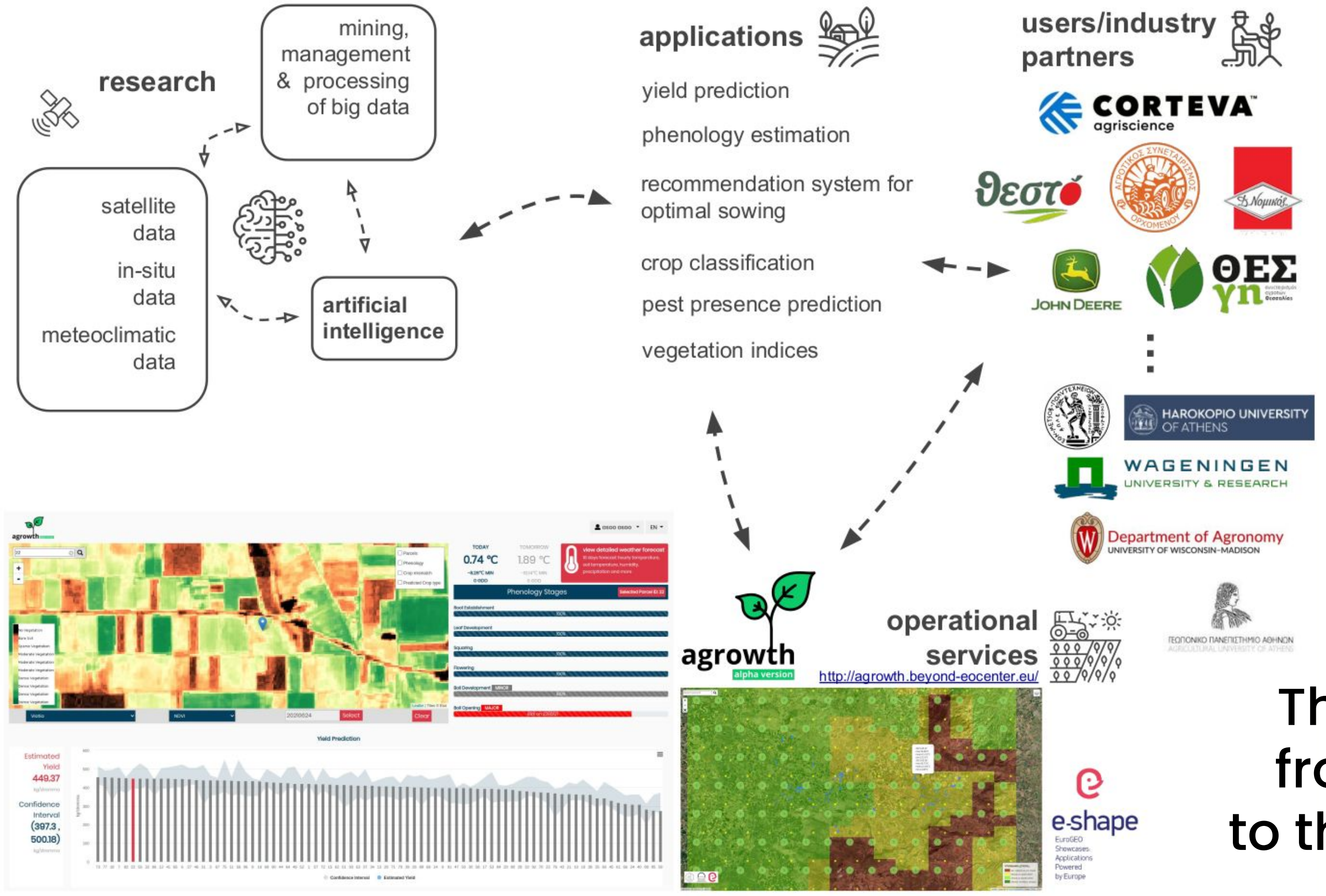


aims of scientific pillar II

➔ forming of **AI algorithms** and methodologies that can leverage this big skein of **observational data** and blend them with the **domain knowledge** in order to promote a sustainable *(profitable?)*, resilient and fair Agriculture.

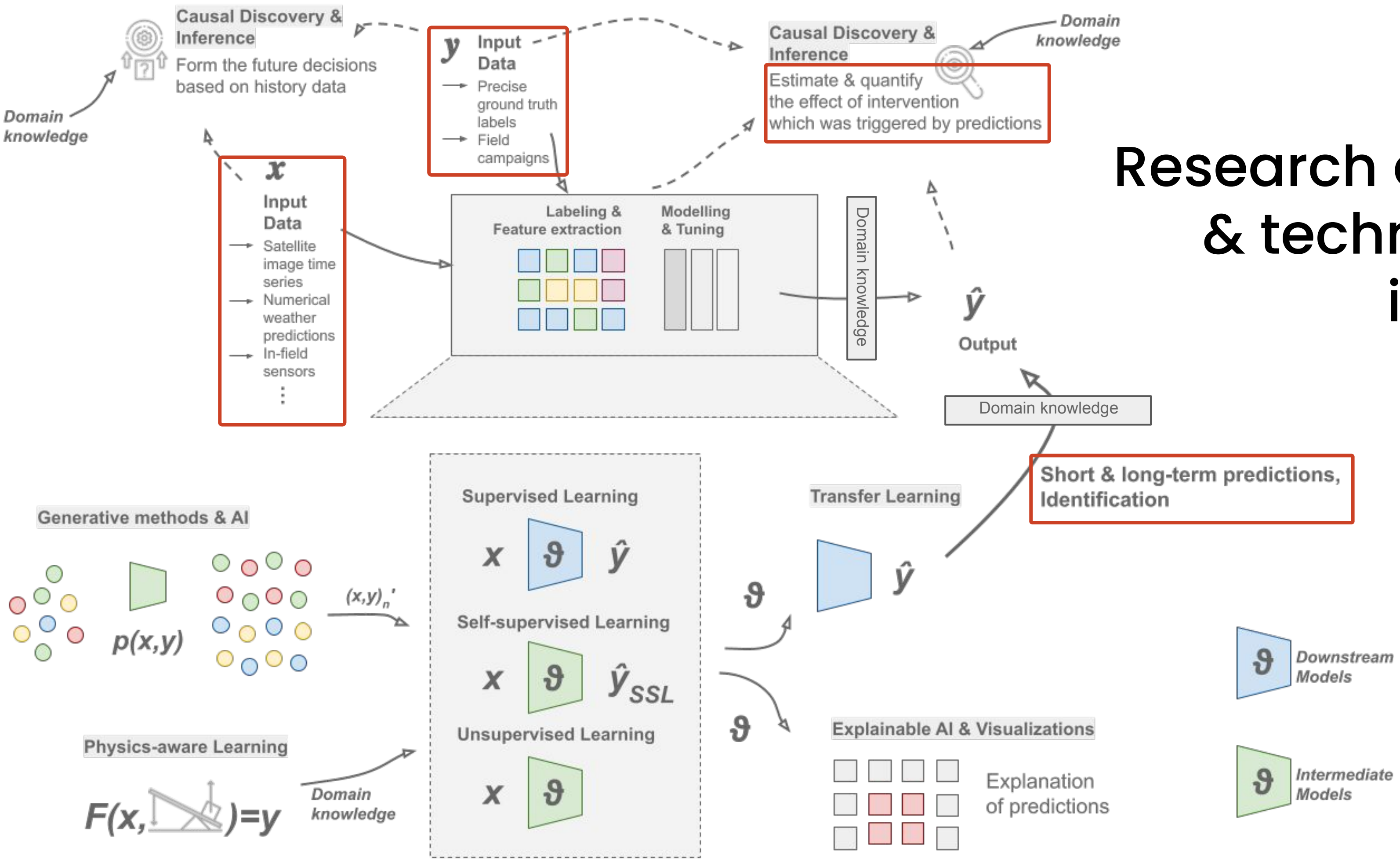
➔ equally important as the first, **serve your user** as your research.

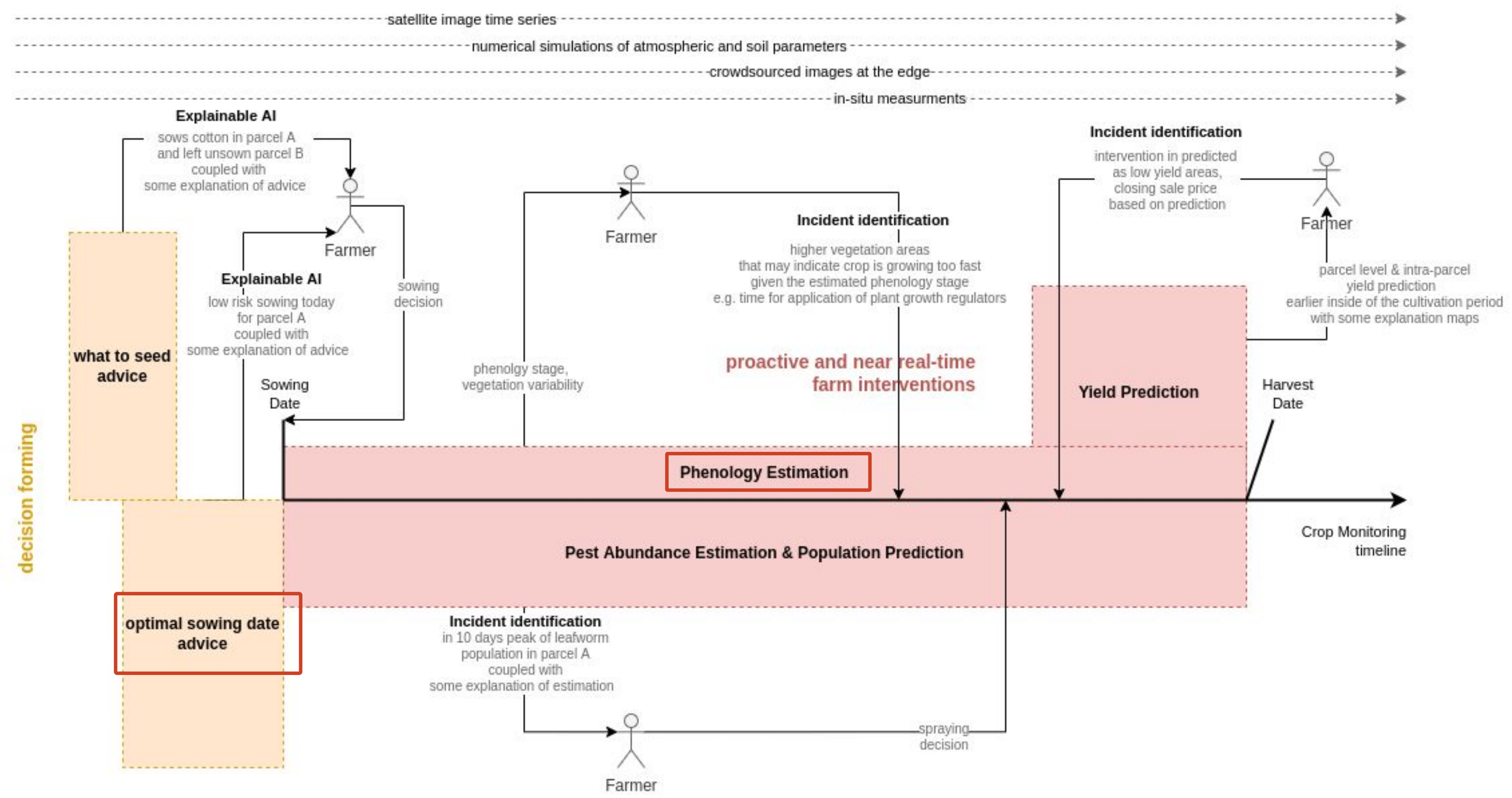
Project Title: *Using data-driven knowledge for profitable soybean management systems*



The narrative from the user to the research and back

Research concept & technologies involved

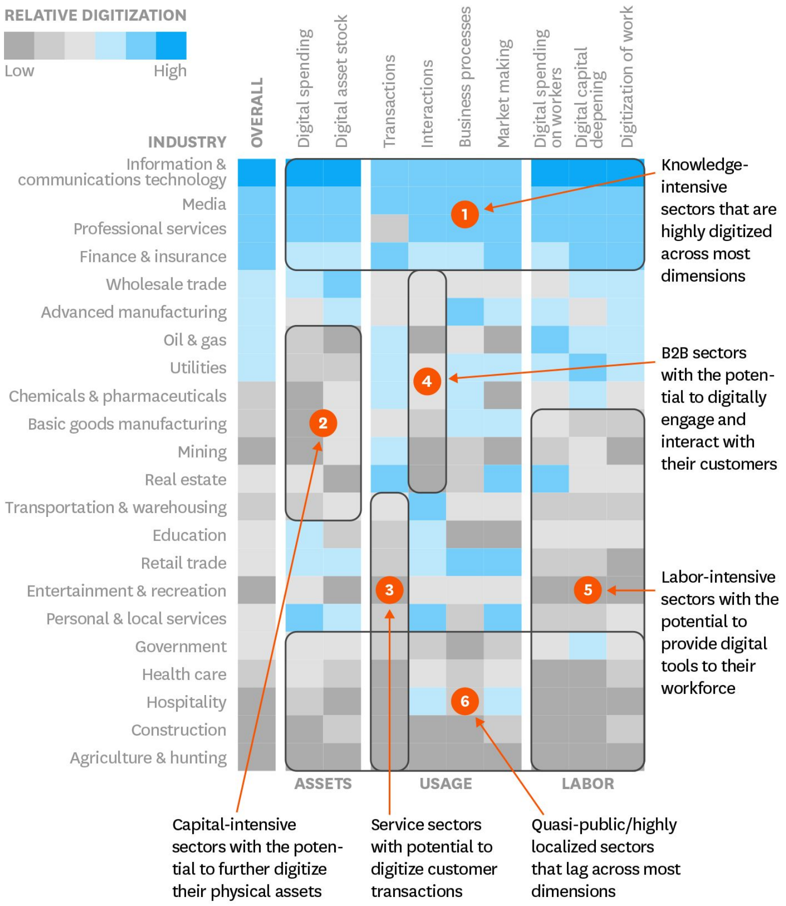




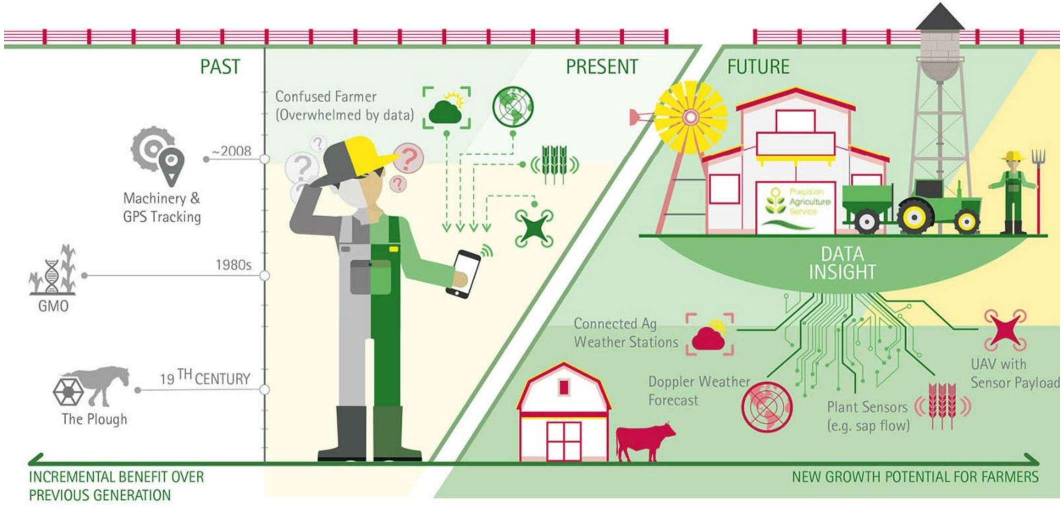
A full blown senario

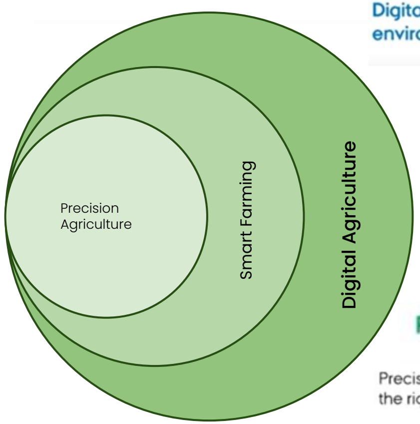
How Digitally Advanced Is Your Sector?

An analysis of digital assets, usage, and labor.



SOURCE DATA ANALYSIS AND EXPERT INTERVIEWS CONDUCTED BY THE MCKINSEY GLOBAL INSTITUTE





Digital Agriculture combines multiple data sources with advanced crop and environmental analyses to provide support for on-farm decision making

Fulton and Port, 2018

A farmer using Digital Agriculture will combine the latest technologies to increase the overall value of several areas of the farm (not the field!).

Data Insight?

Precision Agriculture deals with managing field's variability;

Precision Agriculture was defined as applying inputs at the right time, the right amount and the right place.

IEEE Spectrum Want a Really Hard Machine Learning Problem? Try Agricultu...

Type to search

NEWS | ROBOTICS

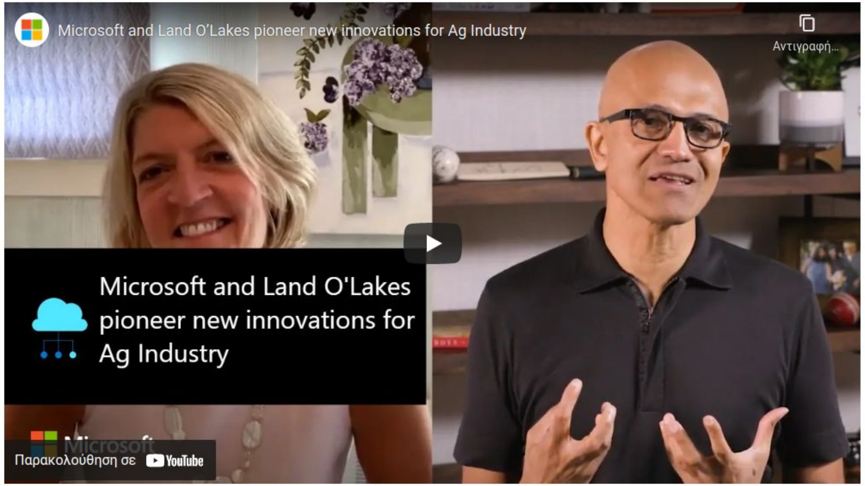
Want a Really Hard Machine Learning Problem? Try Agriculture, Says John Deere Labs › John Deere, the nearly 200-year-old tractor manufacturer, now considers itself a software company

BY TEKLA S. PERRY | 04 OCT 2019 | 5 MIN READ

Microsoft | Stories Our Company News and Stories Press Tools

Land O'Lakes and Microsoft form strategic alliance to pioneer new innovations in agriculture and support rural communities

July 15, 2020 | Microsoft News Center



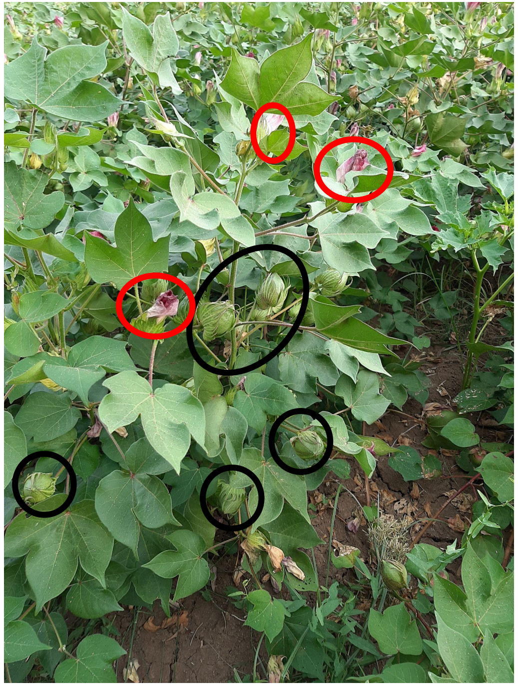
Together the companies aim to build tighter connections between consumers and farmers through innovative new technologies built on Microsoft's cloud

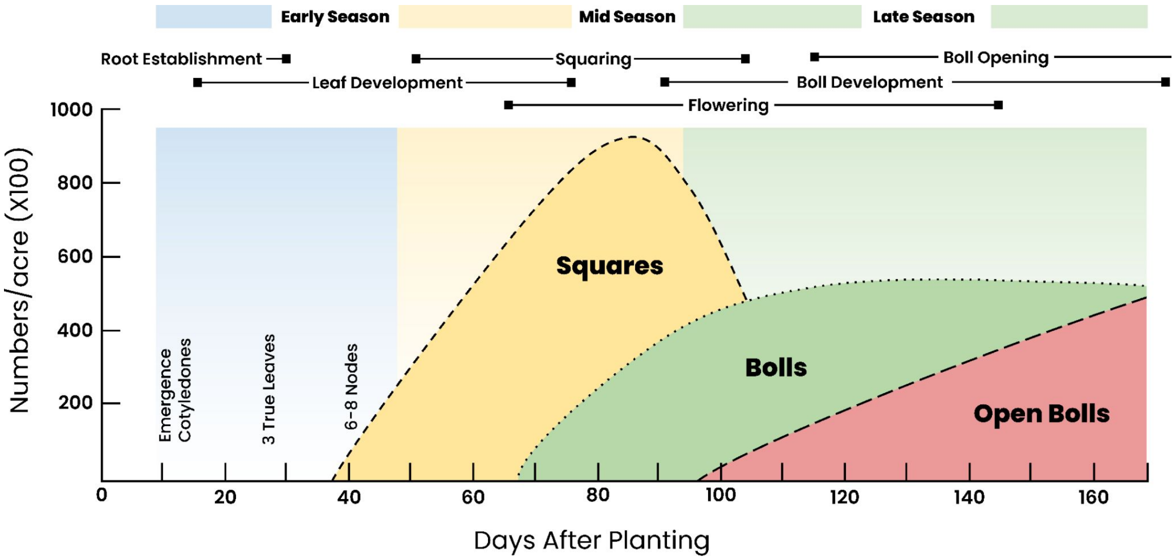
Phenology Estimation

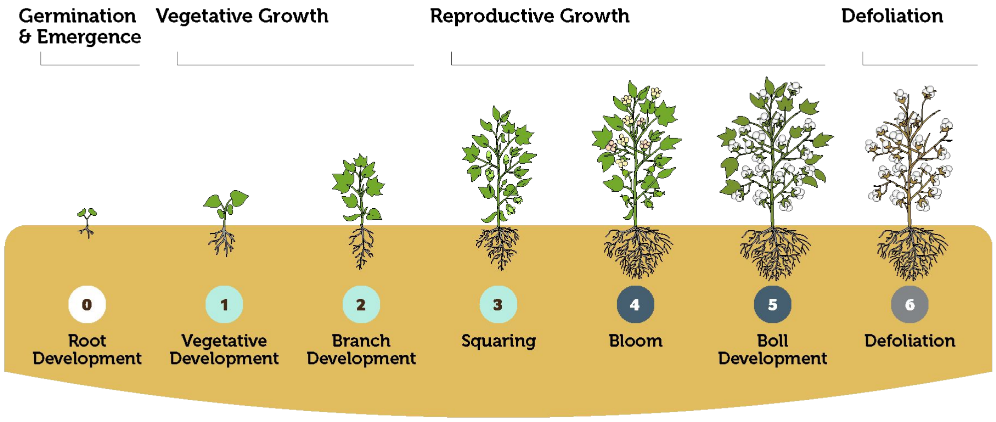
Remote crop monitoring

Case Study: Cotton

- ➔ Vital crop for the Greek economy
- ➔ Underrepresented in the literature



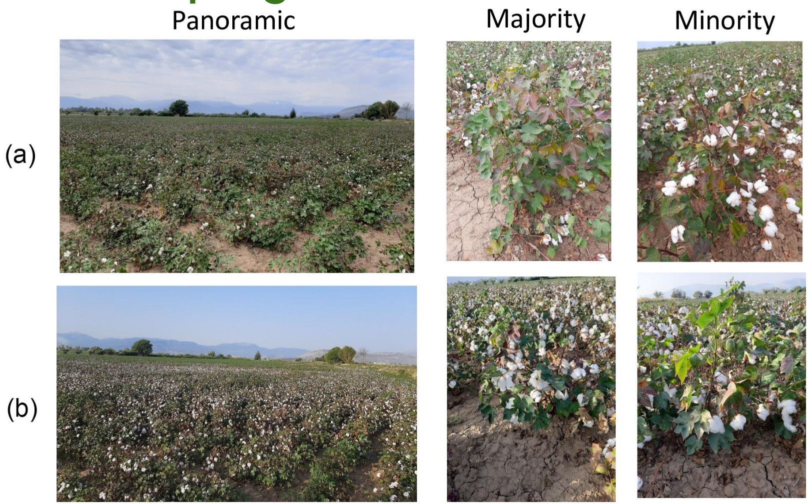




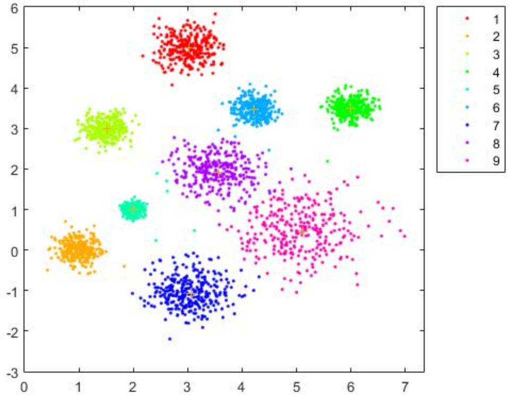
Case Study: Cotton

Lack of annotated data

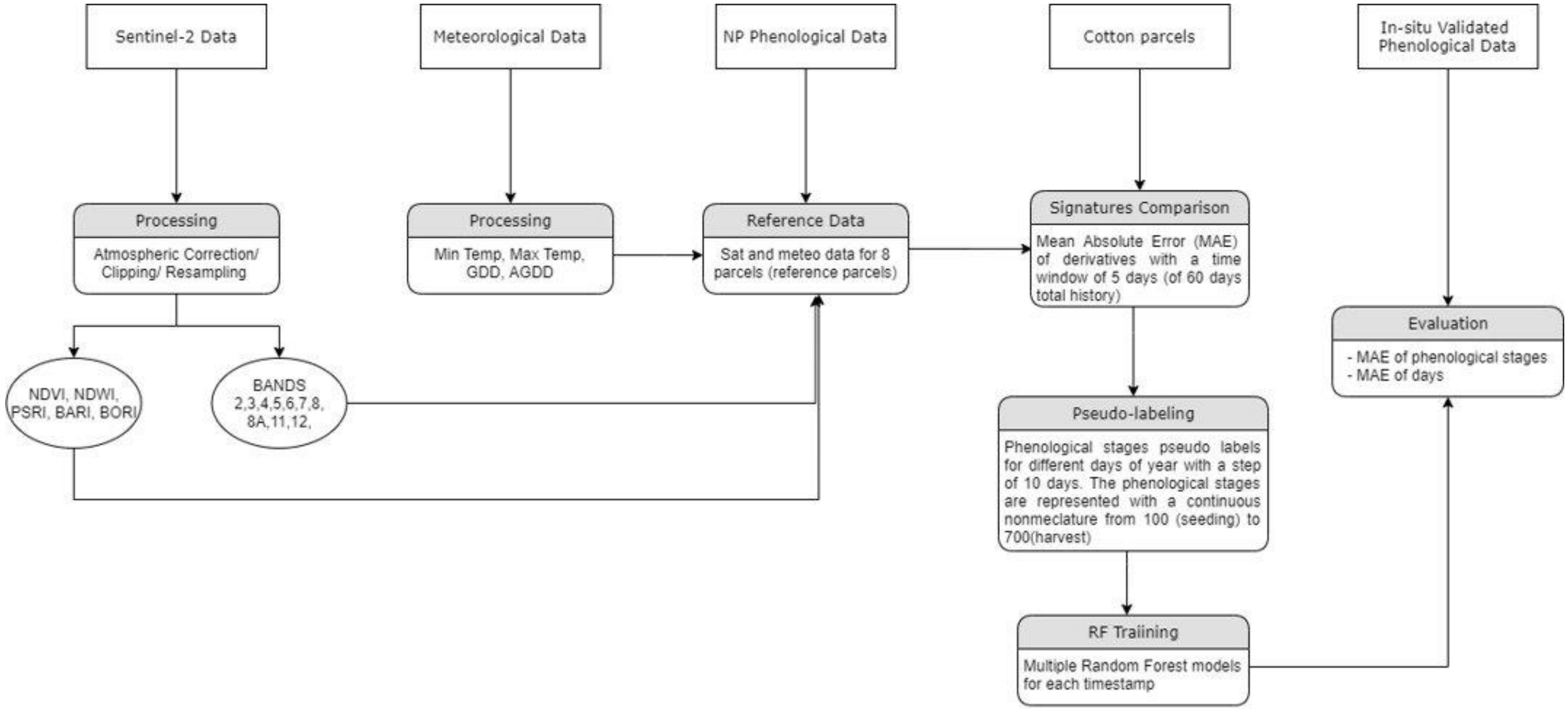
Phenology annotation campaigns

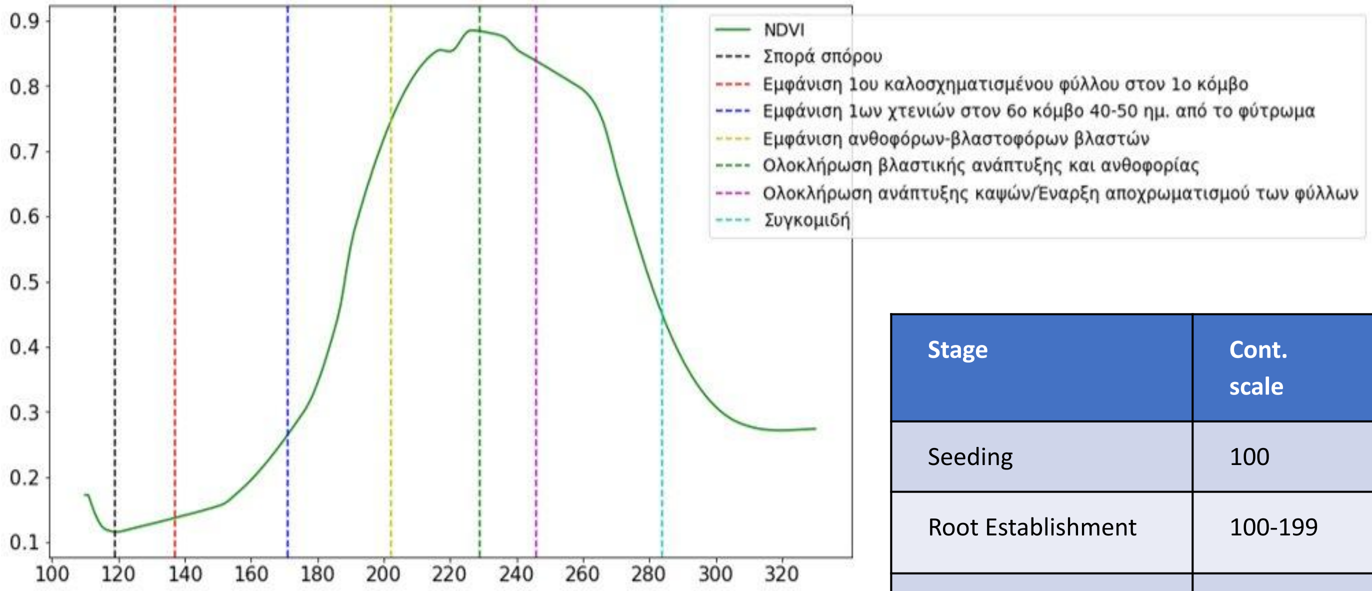


Unsupervised/Semi-supervised approaches



Semi-Supervised (heuristic) phenology estimation



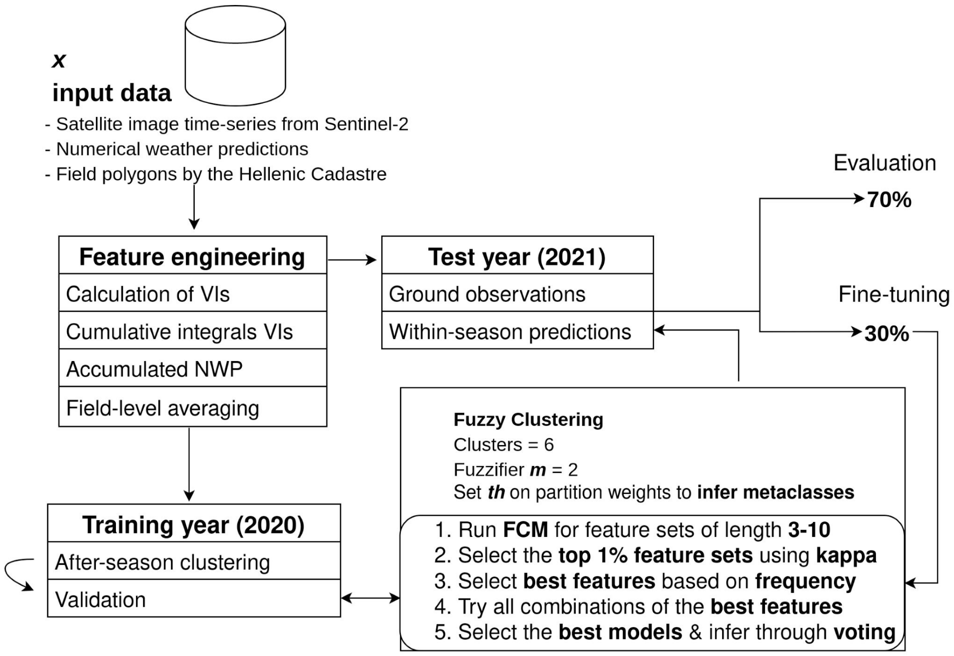


Semi-Supervised (heuristic) phenology estimation

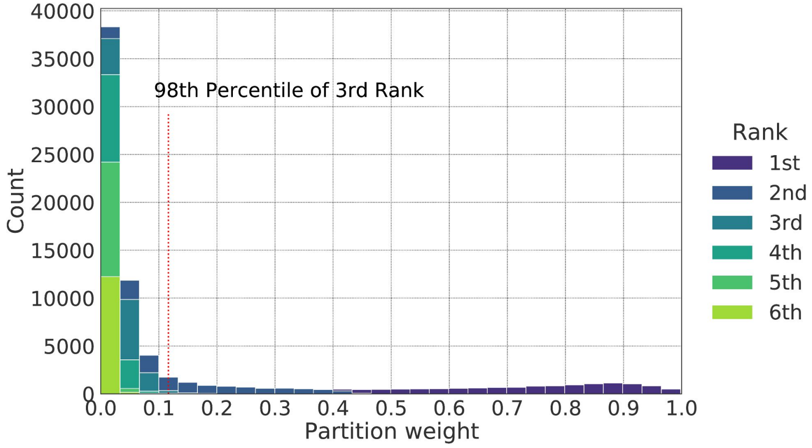
Stage	Cont. scale	DoY range	Duration (days)
Seeding	100	110-125	-
Root Establishment	100-199	110-150	15-25
Leaf Development	200-299	130-190	25-40
Square	300-399	160-215	20-25
Flowering	400-499	180-250	35-45
Boll Development	500-599	220-270	20-25
Boll Opening	600-699	240-315	25-45
Harvest	700	-	-

Unsupervised phenology estimation

Mixed stages ?? → fuzzy clustering !!



Prediction of primary and secondary stage



Phenology annotation campaigns

Supervised learning

Annotation Protocol

- ➔ At least 15 visits per field (approx. 3 per month) during the growing period
- ➔ Ideally, visit the fields in the days that Sentinel-2 passes over. Consult weather forecasts and decide if the inspection could be delayed for a few days.
- ➔ Walk with a zig-zag pattern for typical scouting through the field and inspect the growth status and how it varies in space.
- ➔ Decide on the phenological stage that best describes the majority of the plants in the field. If the field is in a transitioning phase between two phenological stages, mention both and decide which is the prevailing one, i.e., the primary stage.
- ➔ Decide on the percentage that is explained by the primary and the secondary stage
- ➔ Take a panoramic photo of the entire field. Take two close-up photos of plants. The first one should be representative of the majority of the plants in the field. The second one should be representative of a minority of plants in the field.

Panoramic Majority Minority

(a)

(b)

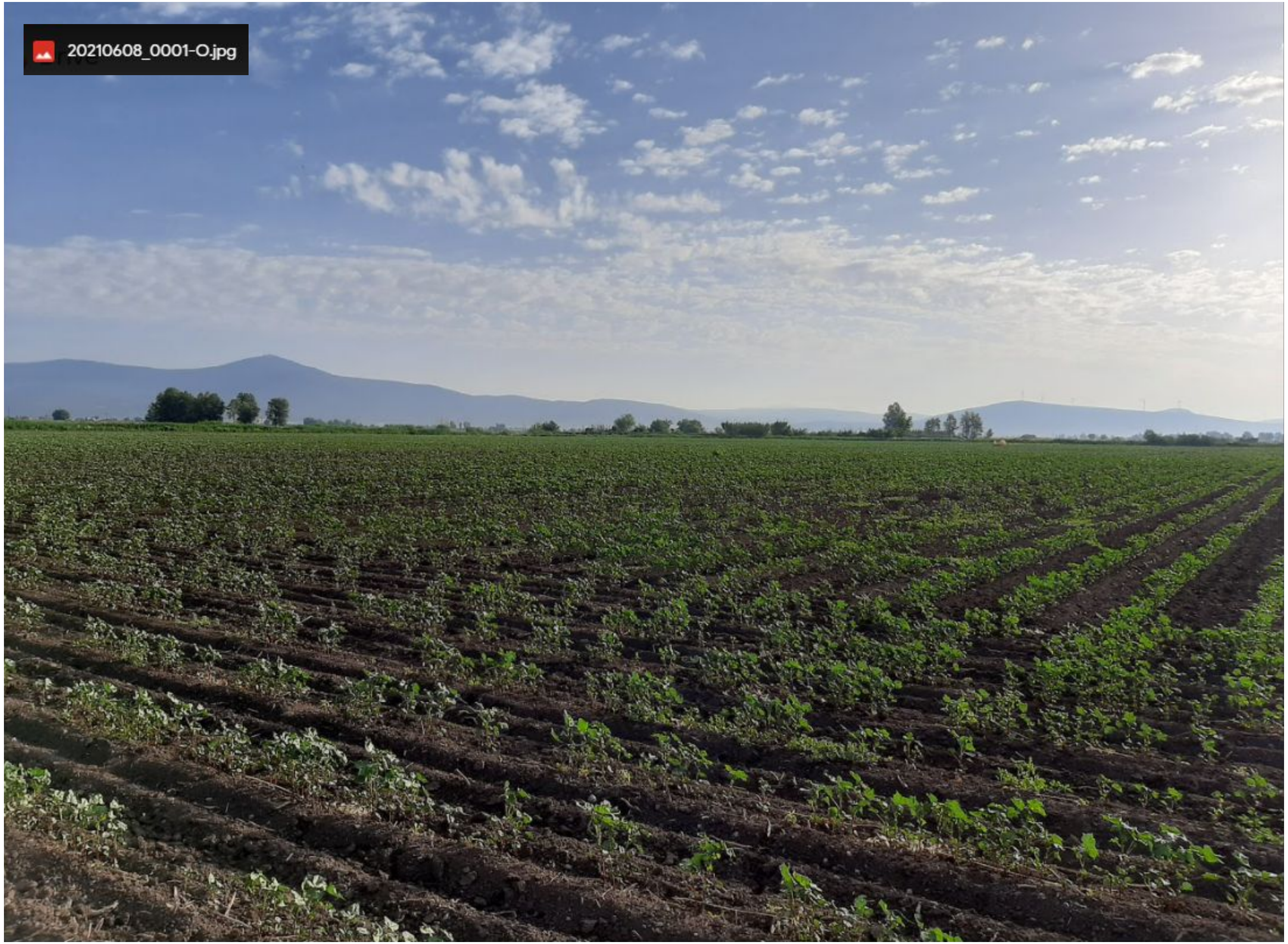
Case 1

Parcels
Phenology
Crop mismatch
Predicted Crop type
Selected Parcel: 22
Leaflet | Tiles © Esri
NDVI
20210607
Select
Clear
TODAY
26.89°C
16.37°C MIN
TOMORROW
27.08°C
16.75°C MIN
view detailed weather forecast
10 days forecast hourly temperature, soil temperature, humidity, precipitation and more
Phenology Stages
Selected Parcel ID: 22
Root Establishment
100%
Leaf Development
MAJOR
39% on 07/06/21
Squaring
Flowering
Boll Development
Boll Opening

20210604_0022-A1.jpg

Case 2


Parcels
Phenology
Crop mismatch
Predicted Crop type
Selected Parcel: 1
Leaflet | Tiles © Esri
NDVI
20210607
Select
Clear
TODAY
28.42°C
16.02°C MIN
TOMORROW
28.16°C
16.64°C MIN
view detailed weather forecast
10 days forecast hourly temperature, soil temperature, humidity, precipitation and more
Phenology Stages
Selected Parcel ID: 1
Root Establishment
100%
Leaf Development
MAJOR
79% on 07/06/21
Squaring
Flowering
Boll Development
Boll Opening



20210608_0001-A1.jpg

Future steps

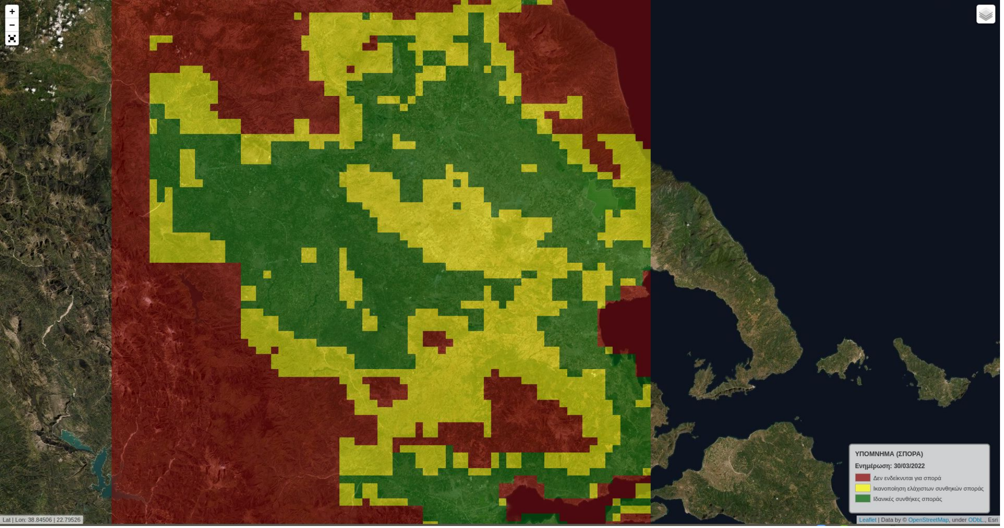
- Drone campaigns → Cotton phenology benchmark dataset
  - ➔ Phenological calendars
  - ➔ Field photos
  - ➔ UAV images

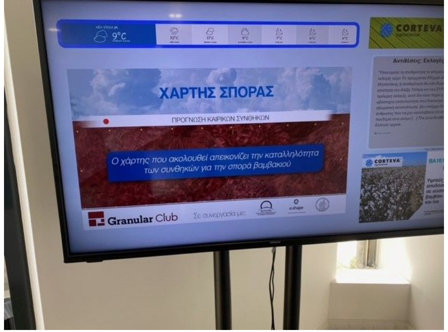
What else? 

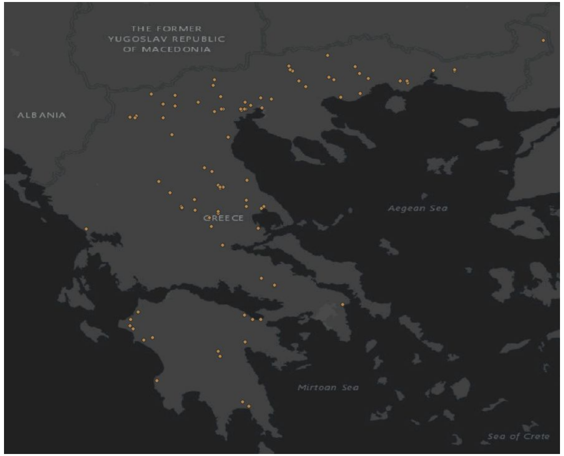
- Self supervised learning: generate latent representations
- Exploit the data!! How?
  - Fusion of satellite with UAV images
  - Phenology estimation on different data source
- Other Crop Types

To sow or not to sow?

A recommendation system for optimal sowing

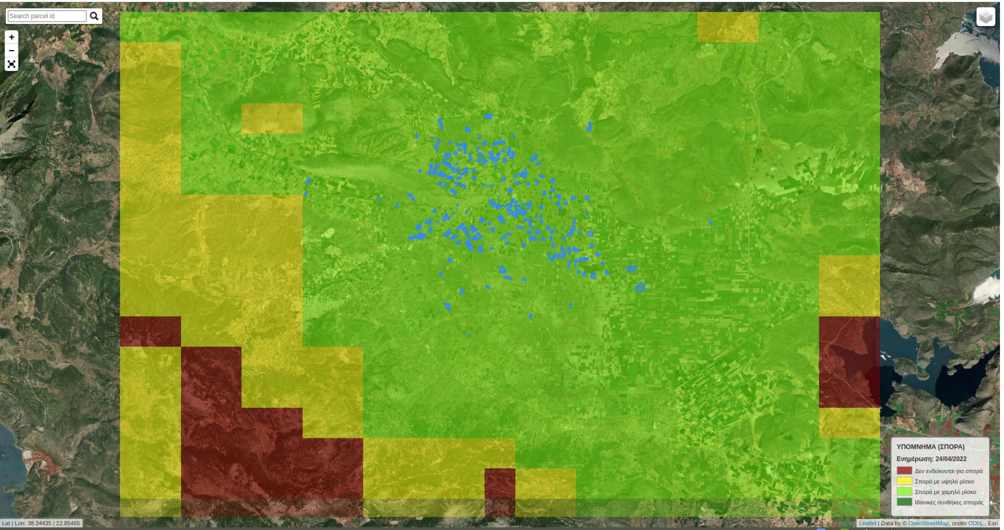


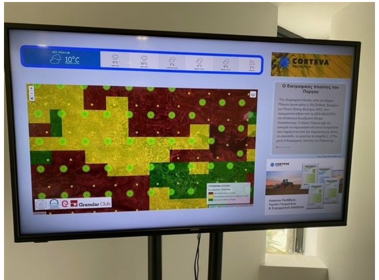




pilot of sowing map for cotton for cultivation period of 2021 in Orchomenos

commercial use of sowing map for cultivation period of 2022 in GR (sunflower, corn, cotton)

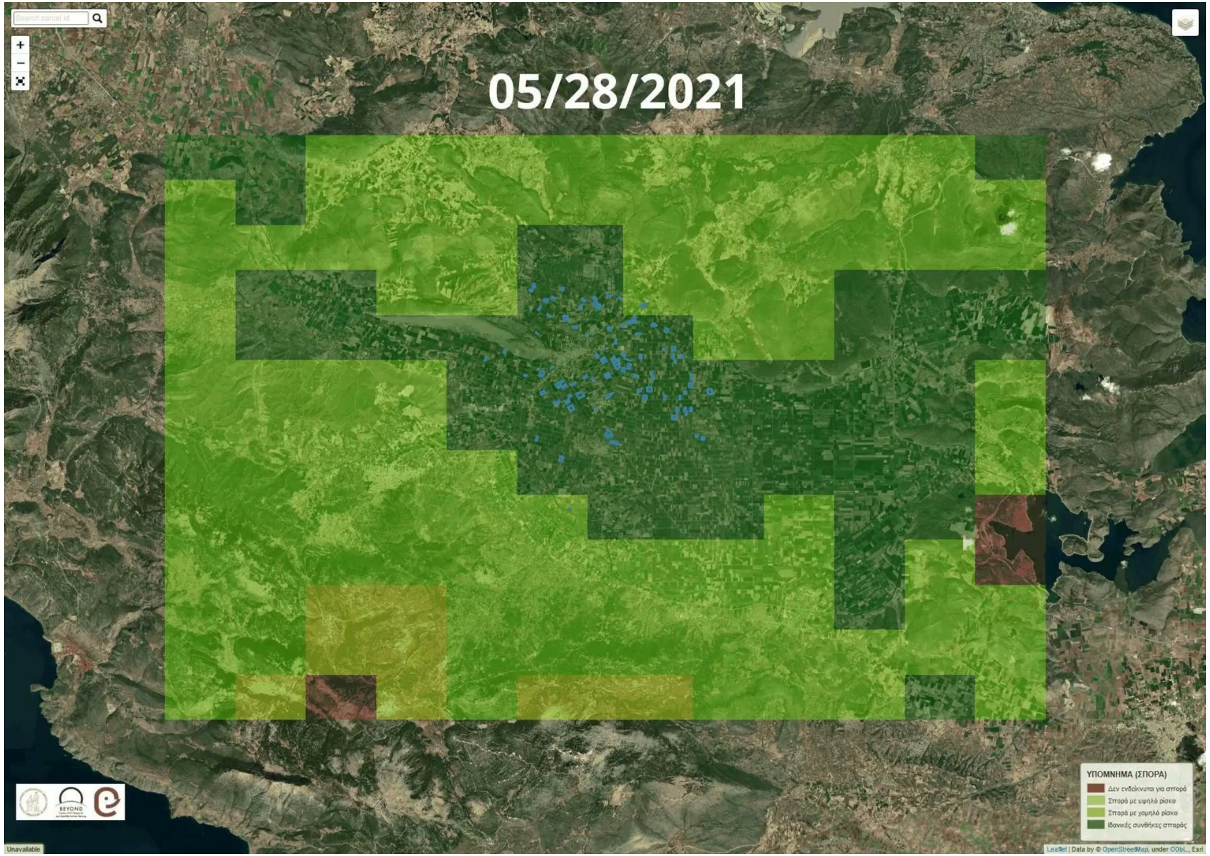




Serving another real need

Map

- ➔ Numerical weather predictions (2-day at 2km WRF , 10-day at 25km GFS, soil temp 0-10cm + ambient temp)
- ➔ Appropriate temp thresholds from Agricultural bibliography (credits: Dimitra Loka - cotton & George Zanakis - corn, sunflower)
- ➔ Basic time series analysis to generate an artificial 10-day at 2km



Knowledge-based Recommendation System

$$a_i = 1 - \frac{GFS_{day=1}}{GFS_{day=i}},\ i \in [2, 10]$$

$$for\ j = \{1, 2\}:$$

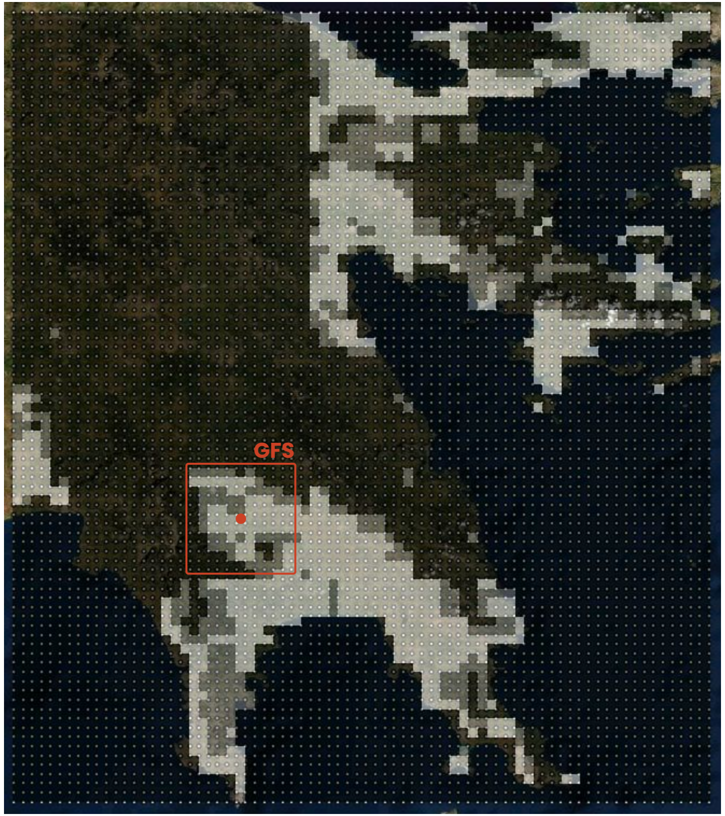
$$WRF_{artificial\ 10days_j} = WRF_j$$

$$for\ j \in [3, 10]:$$

$$WRF_{artificial\ 10days_j} = WRF_1 \times \left(1 + a_j\right)$$

Artificial 10-day
at 2km forecast
blending WRF & GFS

Spatial kNN



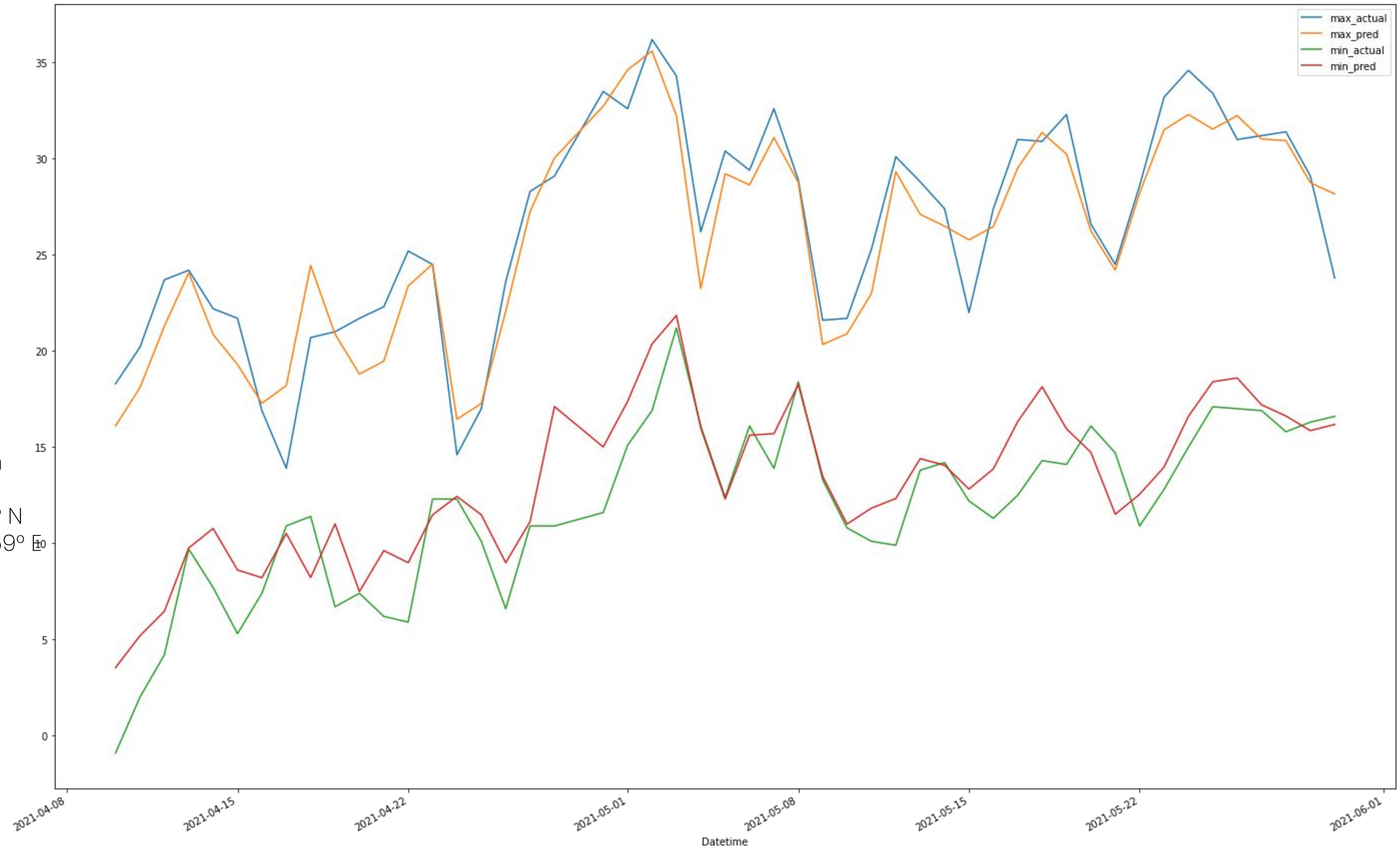
Knowledge-based rules

Type of Temperature	Statistic	Condition	Time Window(days)	Option	code
soil	mean	>18	10	optimum	opt1
ambient	max	>26	2~5	optimum	opt2
soil	mean	>15.56	5	mandatory	mand1
soil	min	>10	2~5	mandatory	mand2
ambient	min	>10	2~5	mandatory	mand3

WRF
1st day
vs
Real Temps

Real Weather measurements from Meteo station in Kopaida
Latitude : 38.45074° N
Longitude : 22.99469° E
Altitude: 98m

8/4/2021 ~ 1/6/2021

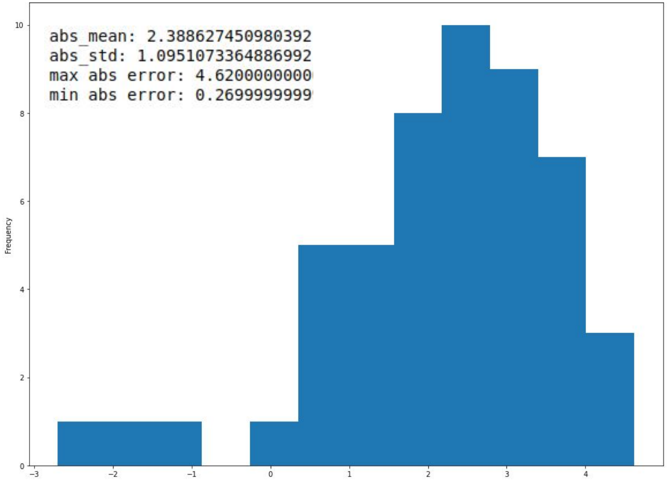


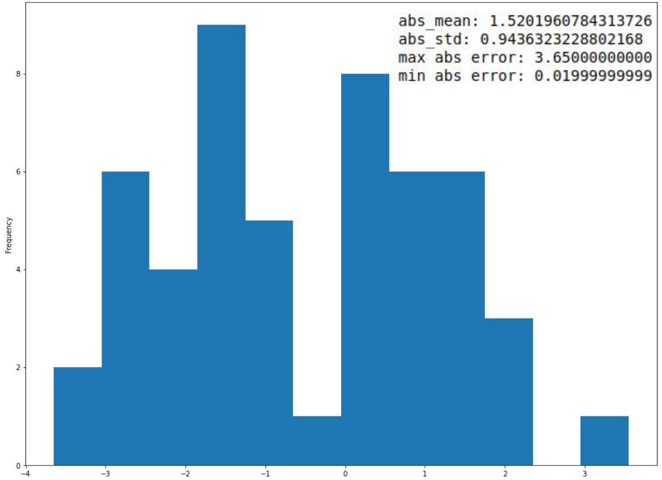
GFS vs our WRF

an embarrassingly simple approach

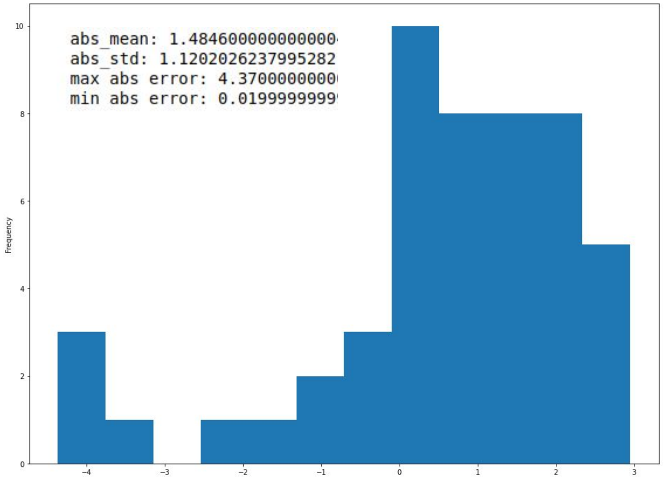
Real Weather measurements
from Meteo station in Kopaida
Latitude : 38.45074° N
Longitude : 22.99469° E
Altitude: 98m

Differences/errors of max/min
between Real Temps and **GFS** 1st day Temp at 2m
(10/4/21 - 30/5/21)





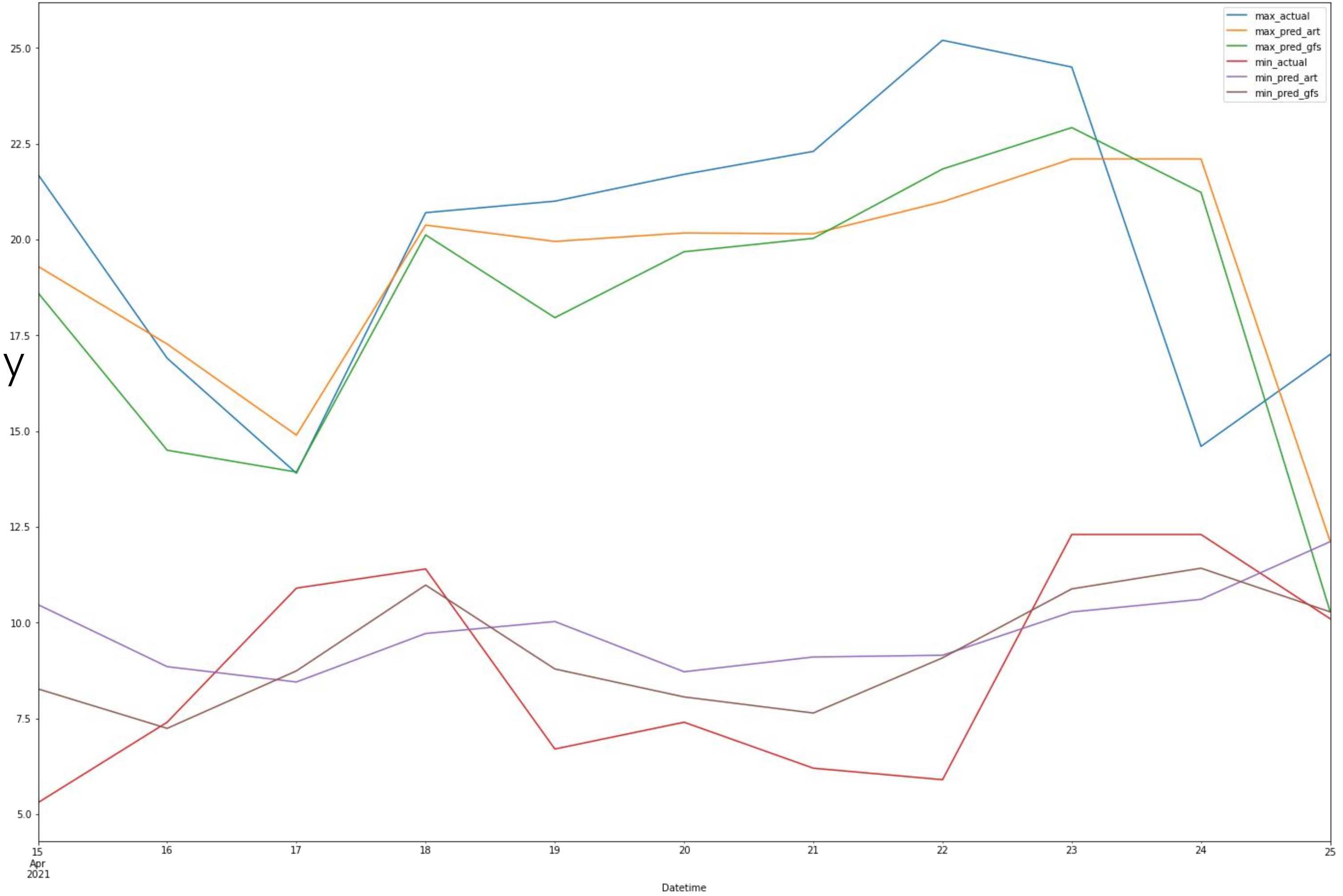
Differences/errors of max/min
between Real Temps and **WRF** 1st day Temp at 2m
(10/4/21 - 30/5/21)



GFS 10-day vs Artificial WRF-10day

an embarrassingly simple approach

Real Weather measurements from Meteo station in Kopaida
Latitude : 38.45074° N
Longitude : 22.99469° E
Altitude: 98m



Hmm,
but what is the
effectiveness of our
recommended
actions?

Evaluating agricultural recommendations

[BOOK] Evaluating decision support and expert systems

L Adelman - 1992 - dl.acm.org

Three approaches to evaluating decision support and expert systems are presented: subjective, technical, and empirical. Subjective evaluation assesses the decision support or expert system from the perspective of the system's users and sponsors. For subjective evaluation, the author presents several techniques including multiattribute utility technology, cost-benefit analysis, and decision analysis. Technical evaluation determines whether the delivered system is a good technical product. Technical evaluation techniques include …

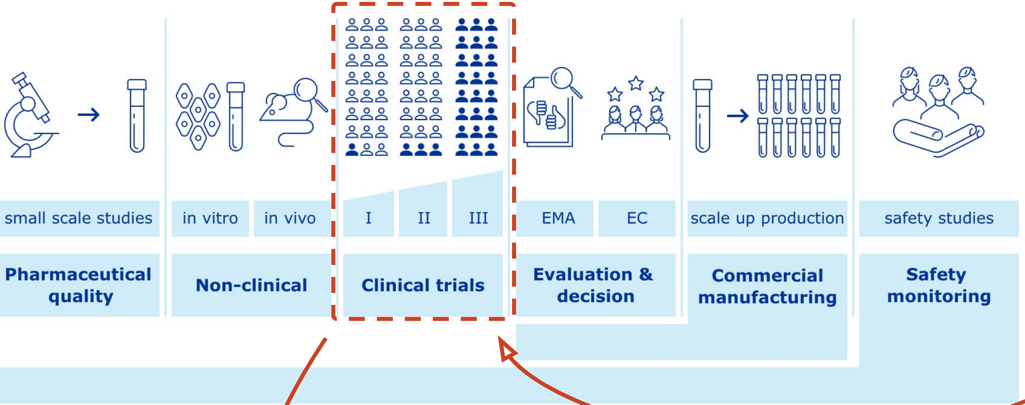


Table 1. *Evaluation methods overviewed herein*

Subjective evaluation methods for requirements validation and to obtain system performance and usability judgments

- Multi-Attribute Utility Assessment (MAUA)
- Task analysis
- Interviews and questionnaires
- Observation
- Human factors checklists
- User diaries

Technical evaluation methods

- Static and dynamic analysis to assess the logical consistency and completeness of the knowledge base
- Domain experts and the use of test cases to assess the functional completeness and predictive accuracy of the knowledge base
- Software testing methods to assess "service requirements"

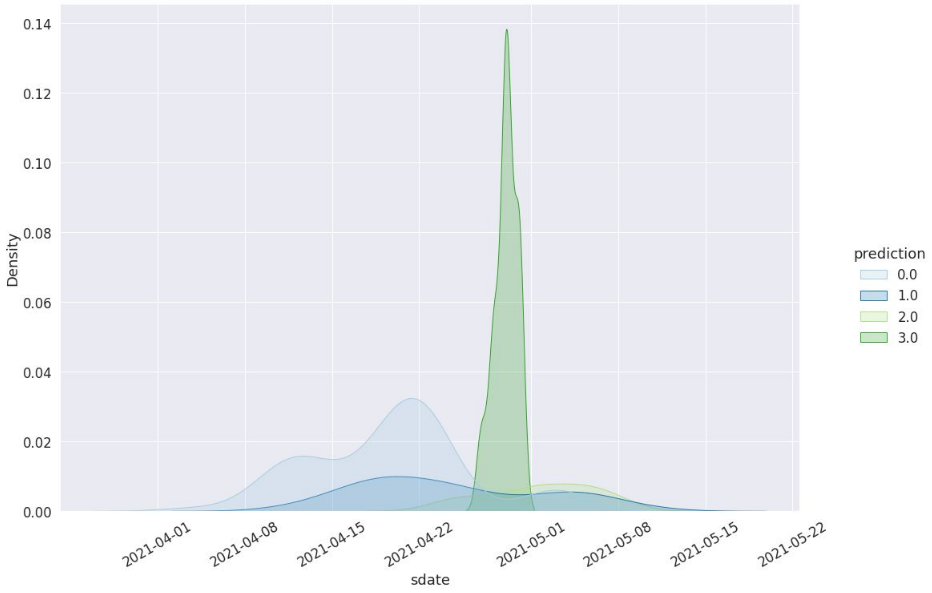
Empirical evaluation methods to obtain objective measures of system performance

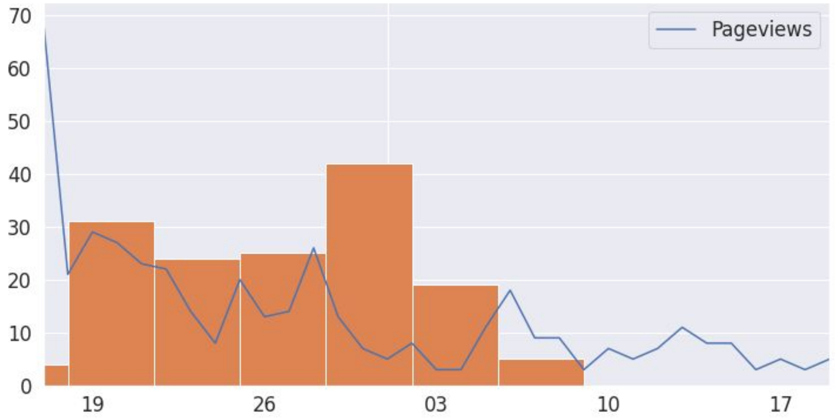
- Experiments
- Quasi-experiments
- Case studies (i.e., field tests)

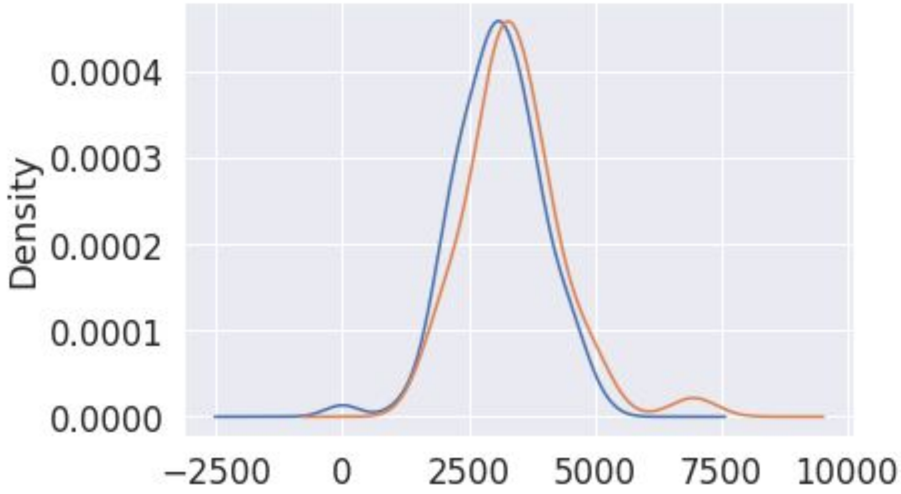
Ok, lets run our experiments!

Context info

	id	ha	variety	sdate	hdate	yield21	lat	lon
0	80	1.8	ARMONIA	2021-04-04	2021-09-27	2400.0	38.523236	22.959435
1	228	0.2	ST 402	2021-04-10	2021-09-27	3200.0	38.531556	22.961721
2	233	0.55	ST 402	2021-04-10	2021-09-27	2800.0	38.532873	22.961703
3	230	0.55	ST 402	2021-04-10	2021-09-27	2880.0	38.530406	22.961877
4	3	2.48	FIDEL	2021-04-11	2021-09-29	3060.0	38.517362	22.994160
...	...	...	...	...	...	...	...	...
166	66	0.26	FIDEL	2021-05-06	2021-09-20	3000.0	38.495017	22.999464
167	206	0.97	ST 402	2021-05-07	2021-09-10	2000.0	38.522107	22.996148
168	207	0.96	FIDEL	2021-05-07	2021-09-15	3300.0	38.521262	22.976469
169	204	1.82	FIDEL	2021-05-07	2021-09-15	2700.0	38.497463	22.969979
170	82	1.94	ARMONIA	2021-05-09	2021-09-27	2800.0	38.521021	22.957991







But, is ok to simply run a independent t-test? What are the assumptions under the hood?

1. the scale of measurement applied to the data collected follows a continuous or ordinal scale, such as the scores for an IQ test. ✅
2. The data, when plotted, results in a normal distribution, bell-shaped distribution curve. ✅
3. There is a reasonably large sample size is used. A larger sample size means the distribution of results should approach a normal bell-shaped curve. ✅
4. Homogeneity of variance. Homogeneous, or equal, variance exists when the standard deviations of samples are approximately equal. ✅
5. Data is collected from a representative, randomly selected portion of the total population. ❌

```
from scipy.stats import ttest_ind
ttest_ind(df_asoo[df_asoo.prediction<3].yield21,
         df_asoo[df_asoo.prediction==3].yield21)
```

```
Ttest_indResult(statistic=-1.8055164512525876, pvalue=0.0727
```

The Average Treatment Effect (ATE)

like as we have runned a randomized expirement

```
df_asoo_updated.query("prediction<3")["yield21"].
```

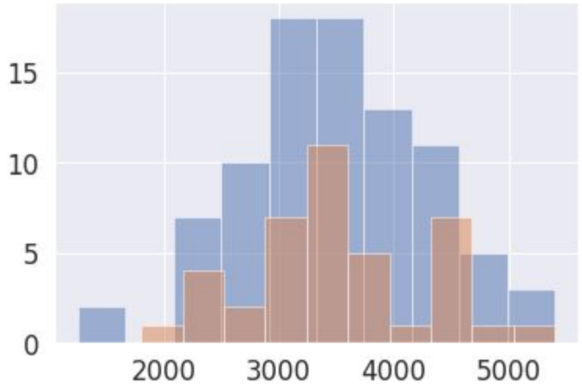
```
836.0874132137353
```

```
df_asoo_updated.query("prediction==3")["yield21"].std()
```

```
891.496540449066
```

```
df_asoo_updated.query("prediction==3")["yield21"].mean() - df_asoo_updated.query("prediction<3")["yield21"].mean()
```

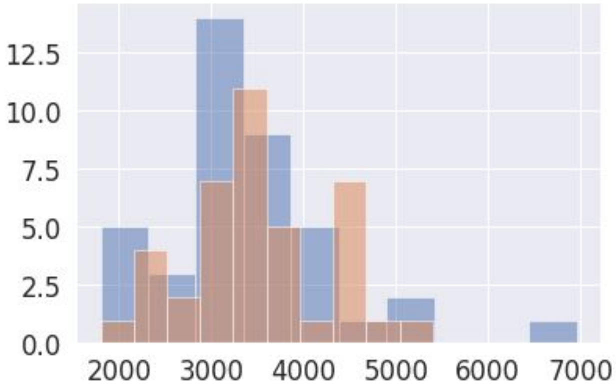
```
258.77851239669417
```



t-test between two groups - 2020

```
# only parcels with available yields for 2020
from scipy.stats import ttest_ind
ttest_ind(df_asoo_yield20[df_asoo_yield20.prediction<3].yield20,
          df_asoo_yield20[df_asoo_yield20.prediction==3].yield20)
```

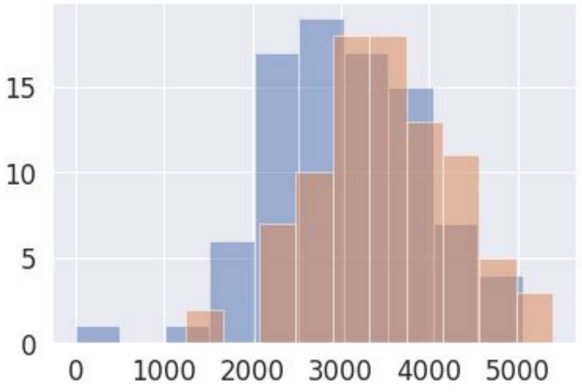
```
Ttest_indResult(statistic=-0.11261261771291173, pvalue=0.91051832
```



t-test between treated of 2021 and their yields of 2020

```
from scipy.stats import ttest_ind
ttest_ind(df_asoo_yield20[df_asoo_yield20.prediction==3].yield21,
          df_asoo_yield20[df_asoo_yield20.prediction==3].yield20)
```

```
Ttest_indResult(statistic=-0.4869871939773592, pvalue=0.62763314
```



t-test between controls of 2021 and their yields of 2020

```
from scipy.stats import ttest_ind
ttest_ind(df_asoo_yield20[df_asoo_yield20.prediction<3].yield21,
          df_asoo_yield20[df_asoo_yield20.prediction<3].yield20)
```

```
Ttest_indResult(statistic=-3.2332272996540423, pvalue=0.00146742
```

About skills... how many of them sown in both cases

```
len(set(farmers_parce
```

```
0.6111111111111112
```

Hmm, interesting insights! Something good happens there!

```
mean(df_asoo_yield20[df_asoo_yield20.prediction<3].yield21 -
```

```
-406.13793103448273
```

```
).std(df_asoo_yield20[df_asoo_yield20.prediction<3].yield21 -
```

```
718.1424981448293
```

```
mean(df_asoo_yield20[df_asoo_yield20.prediction==3].yield21
```

```
-95.0
```

```
std(df_asoo_yield20[df_asoo_yield20.prediction==3].yield21 -
```

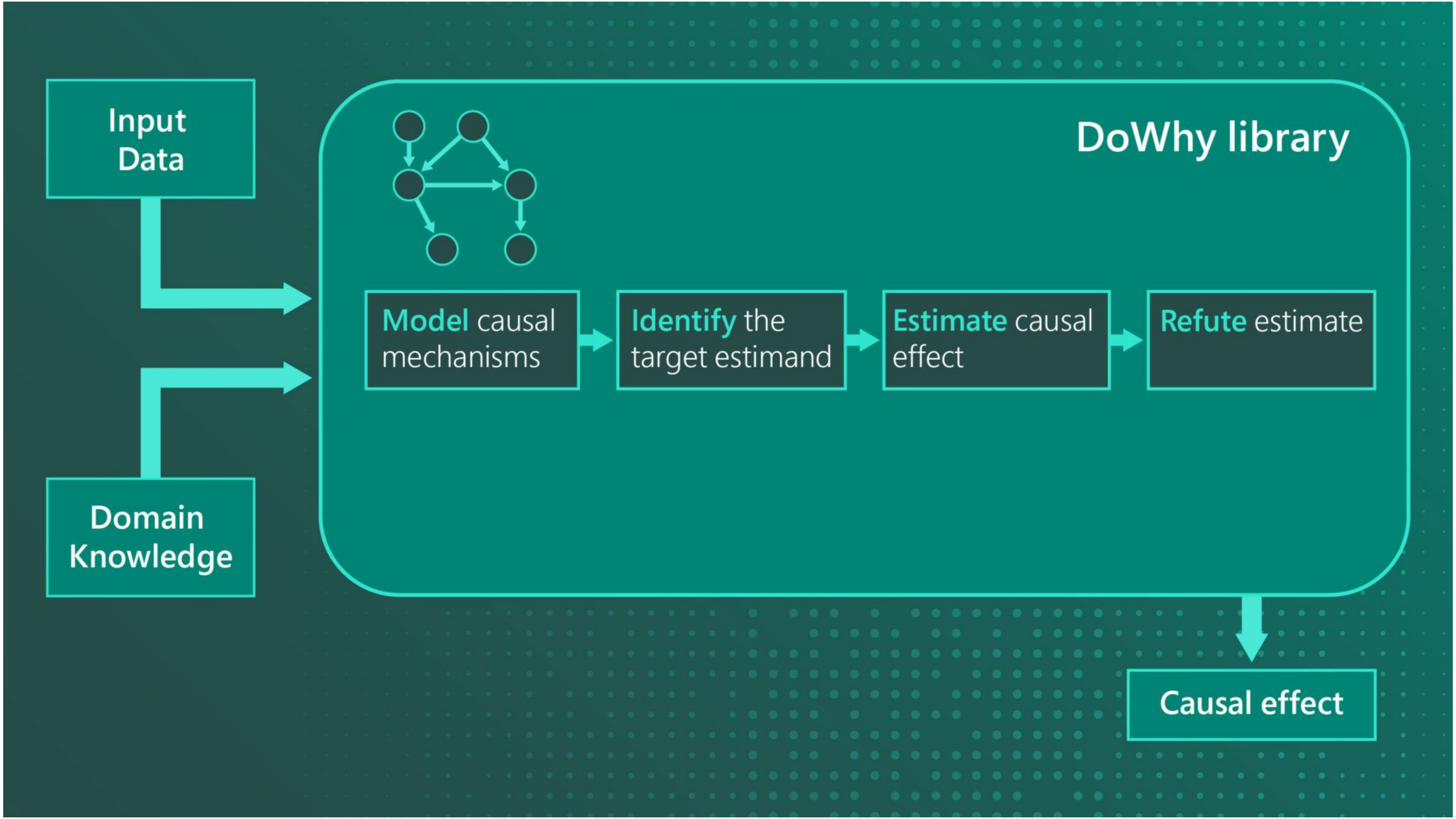
```
757.8093427769282
```

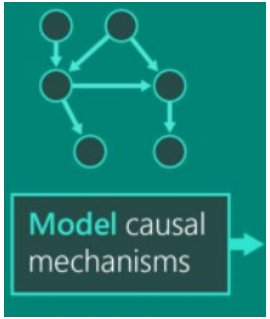
Evaluating agricultural recommendations using **causal inference ftw**

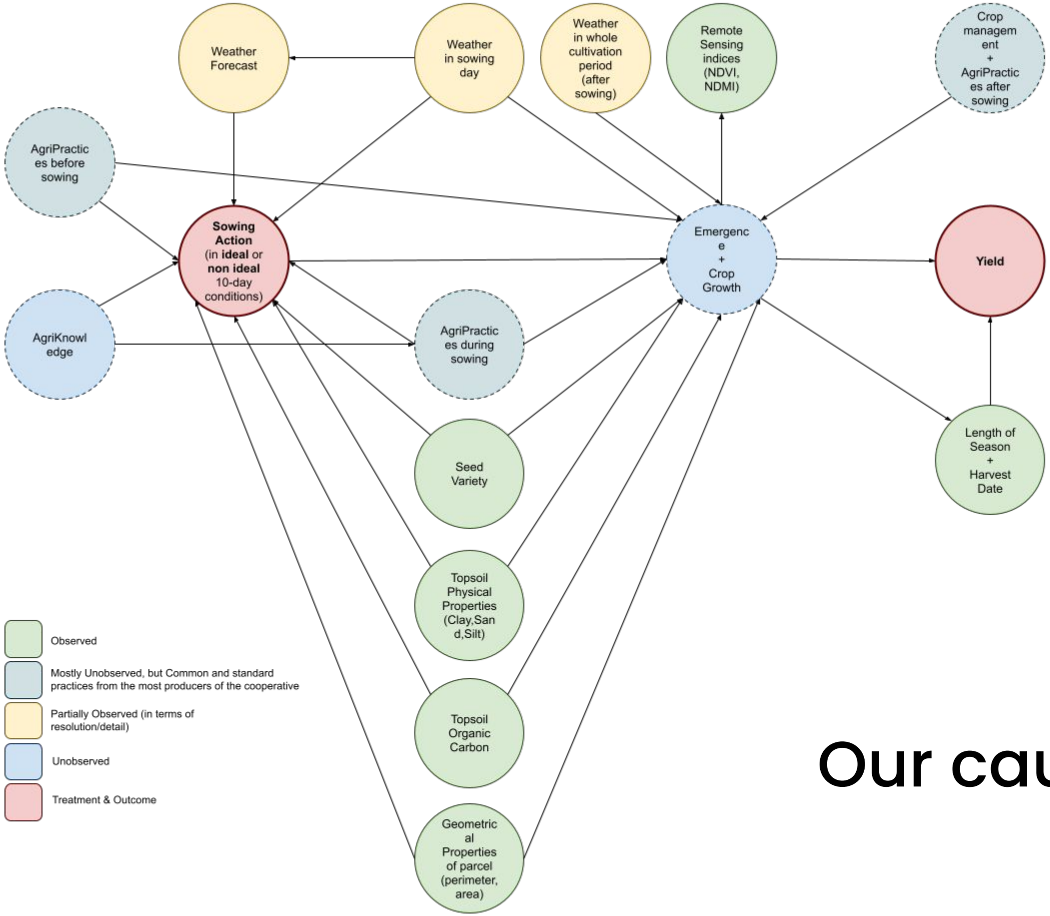
Experiment	identical		Control and treatment are **identical** and their behavior is **deterministic**. Causal effect of treatment is directly the difference between observations for the two groups. *Physics, Biology, ~~Social sciences~~*
Statistical Experiment			Control and treatment are not identical but divided at random. This makes it possible to build a precise estimate of the causal effect of treatment. *A/B testing, Central Limit Theorem, Bayesian Statistics*
Quasi-experiment			Control and treatment are not identical and divided by a "natural" criterion. Depending on "internal" and "external" quality of the criterion, it is possible to build a good estimate of the causal effect of treatment. *Differences-in-differences, Regression Discontinuity, Instrumental variables, Matching, Controlled Regression*
Counterfactuals		?	Control group does not exist, instead its behaviour is estimated with a predictive model of what would have happened without the treatment (= counterfactual). *Synthetic Differences-in-Differences, Athey & Imbens, CausalImpact*

Stronger evidence

Levels of evidence ladder for causal inference methods


Input Data
Domain Knowledge
DoWhy library
Model causal mechanisms
Identify the target estimand
Estimate causal effect
Refute estimate
Causal effect





Our causal graph

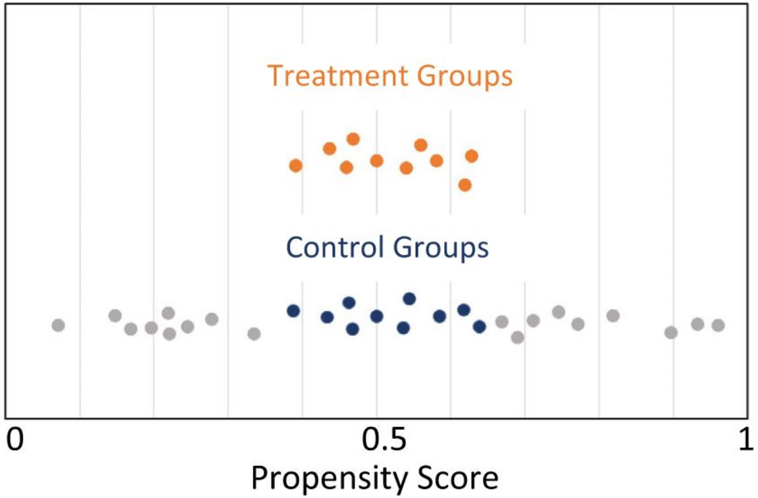
Identify the target estimand

```
Estimand type: nonparametric-ate

### Estimand : 1
Estimand name: backdoor
Estimand expression:
      d
─────────────(Expectation(yield21|clay_mean,HIGH,occont_mean,LOW,silt_mean,san
d[prediction]


d_mean,var_code,ratio))

Estimand assumption 1, Unconfoundedness: If U→{prediction} and U→yield21 then P(yield21|prediction,clay_mean,HIGH,oc
cont_mean,LOW,silt_mean,sand_mean,var_code,ratio,U) = P(yield21|prediction,clay_mean,HIGH,occont_mean,LOW,silt_mean,
sand_mean,var_code,ratio)

### Estimand : 2
Estimand name: iv
No such variable(s) found!

### Estimand : 3
Estimand name: frontdoor
Estimand expression:
Expectation(Derivative(yield21, [trapezoidal_ndvi_sow2harvest])*Derivative([tr
apezoidal_ndvi_sow2harvest], [prediction]))
Estimand assumption 1, Full-mediation: trapezoidal_ndvi_sow2harvest intercepts (blocks) all directed paths from pred
iction to y,i,e,l,d,2,1.
Estimand assumption 2, First-stage-unconfoundedness: If U→{prediction} and U→{trapezoidal_ndvi_sow2harvest} then P(t
rapezoidal_ndvi_sow2harvest|prediction,U) = P(trapezoidal_ndvi_sow2harvest|prediction)
Estimand assumption 3, Second-stage-unconfoundedness: If U→{trapezoidal_ndvi_sow2harvest} and U→yield21 then P(yield
21|trapezoidal_ndvi_sow2harvest, prediction, U) = P(yield21|trapezoidal_ndvi_sow2harvest, prediction)
```

Estimate causal effect

$$ATE = E[Y_1 - Y_0]$$

(a)

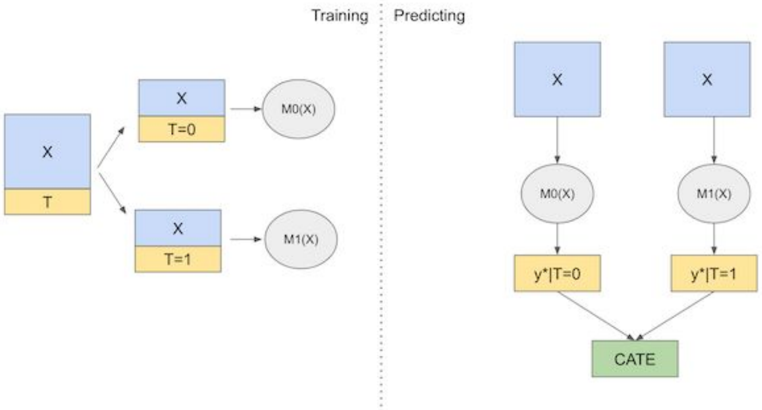


Matching Methods:

- Matching methods (e.g., propensity score analysis) selects control and treatment groups that are similar across selected covariates—based on confounding variables—to reduce confounding bias

Benefit of Causal Diagrams:

- Allows appropriate selection of variables to enter the propensity score, reducing confounding, overcontrol, and collider bias



Causal Effefct Estimation Methods	ATE	CI	p-value
Linear Regression (like S-Learner)	**457.02**	**(117.43, 796.61)**	**0.0086**
Matching	**397.84**	**(42.40, 789.82)**	**0.0160**
T-Learner (RF)	**365.88**	**(−590.96, 1261.82)**	**-**
Matching with trimmed PS (0.20,0.80)	310.58	(-63.72, 603.95)	0.062
IPS weighting trimmed (0.10,0.95)	**532.77**	**(150.53, 906.01)**	**0.0010**
IPS weighting trimmed (0.2,0.8)	**447.39**	**(96.69, 749.71)**	**0.0060**

Refute estimate

Refutation Methods	Causal Effefct Estimation Methods	Estimated Effect	New Effect	p-value
Placebo	Linear Regression	457.02	21.38	0.39
	Matching	397.84	25.1	0.45
	T-Learner (RF)	365.88	-27.37	0.41
	Matching with PS	301.17	-0.67	0.49
	Matching with trimmed PS (0.15,0.95)	787.83	-27.68	0.46
	Matching with trimmed PS (0.20,0.80)	310.58	40.10	0.41
	IPS weighting trimmed (0.10,0.95)	532.77	15.51	0.48
	IPS weighting trimmed (0.2,0.8)	447.39	-15.67	0.47
Random Common Cause	Linear Regression	457.02	455.74	0.5
	Matching	397.84	405.65	0.46
	T-Learner (RF)	365.88	362.15	0.47
	Matching with PS	301.17	447.84	0.5
	Matching with trimmed PS (0.15,0.95)	787.83	457.76	0.01
	Matching with trimmed PS (0.20,0.80)	310.58	488.16	0.06
	IPS weighting trimmed (0.10,0.95)	532.77	533.39	0.47
	IPS weighting trimmed (0.2,0.8)	447.39	444.62	0.38
Removing Random Subset	Linear Regression	457.02	447.67	0.50
	Matching	397.84	397.06	0.46
	T-Learner (RF)	365.88	345.73	0.43
	Matching with PS	301.17	588.01	0.48
	Matching with trimmed PS (0.15,0.95)	787.83	526.23	0.16
	Matching with trimmed PS (0.20,0.80)	310.58	488.54	0.46
	IPS weighting trimmed (0.10,0.95)	532.77	533.036	0.45
	IPS weighting trimmed (0.2,0.8)	447.39	440.09	0.49
Unobserved Common Cause	Linear Regression	457.02	(-188.54, 509.53)	
	Matching	397.84	(-308.35, 452.68)	
	T-Learner (RF)	365.88	(-236.59, 396.54)	
	Matching with PS	301.17	(-359.41, 532.39)	
	Matching with trimmed PS (0.15,0.95)	787.83		
	Matching with trimmed PS (0.20,0.80)	310.58		
	IPS weighting trimmed (0.10,0.95)	532.77	(-245.06, 604.94)	
	IPS weighting trimmed (0.2,0.8)	447.39	(-244.82, 495.38)	

Let’s take our drinks and discuss it