NASA Western Water Applications Office

Award Number 1667355

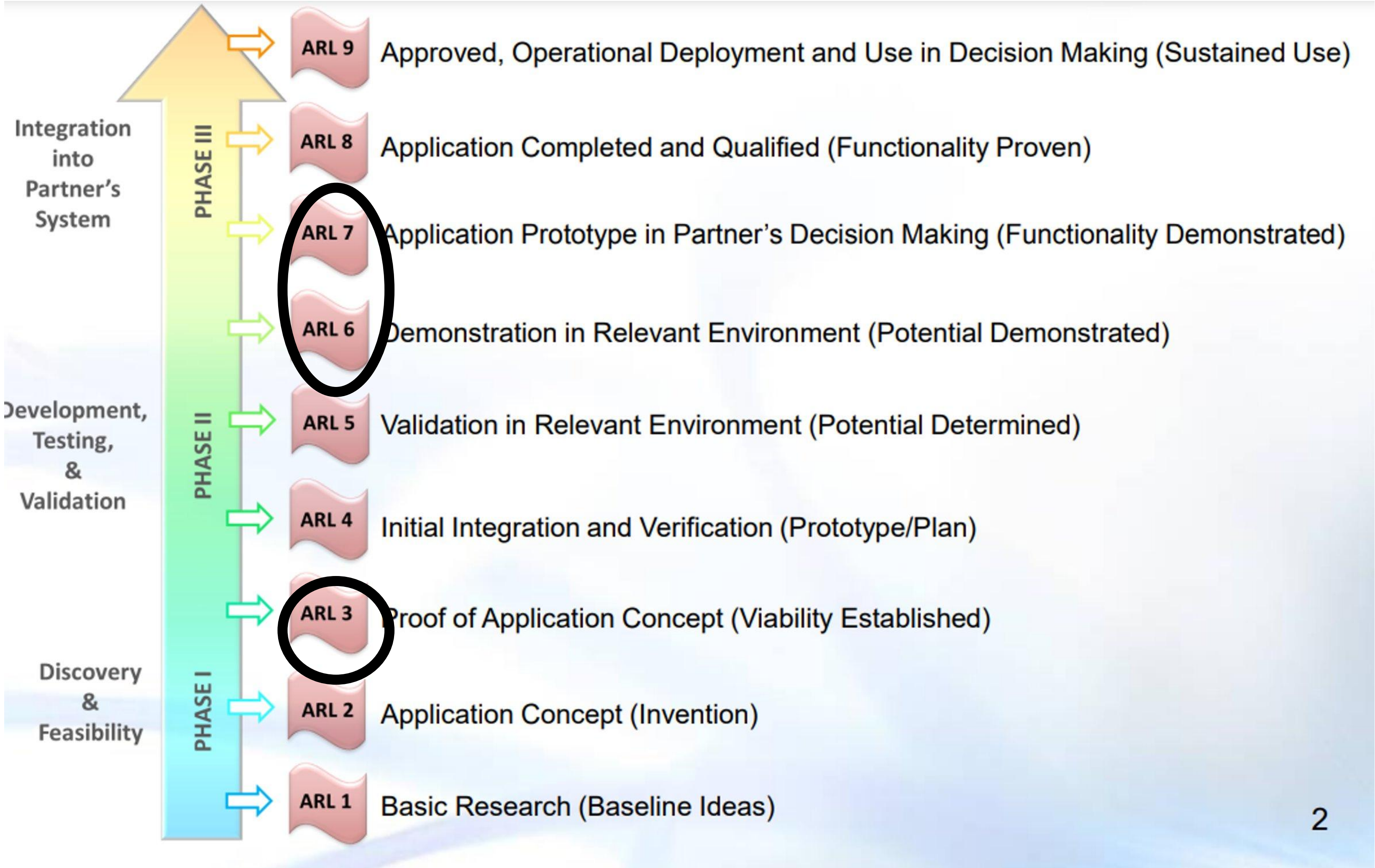
Final project presentation

Michael P. Brady, Kirti Rajagopalan, Hossein Noorazar, Supriya Savalkar, Perry Beale, and Mingliang Liu

Long-term mapping and trend analysis of double-cropping extent in the Columbia River Basin with Landsat and Google Earth Engine.

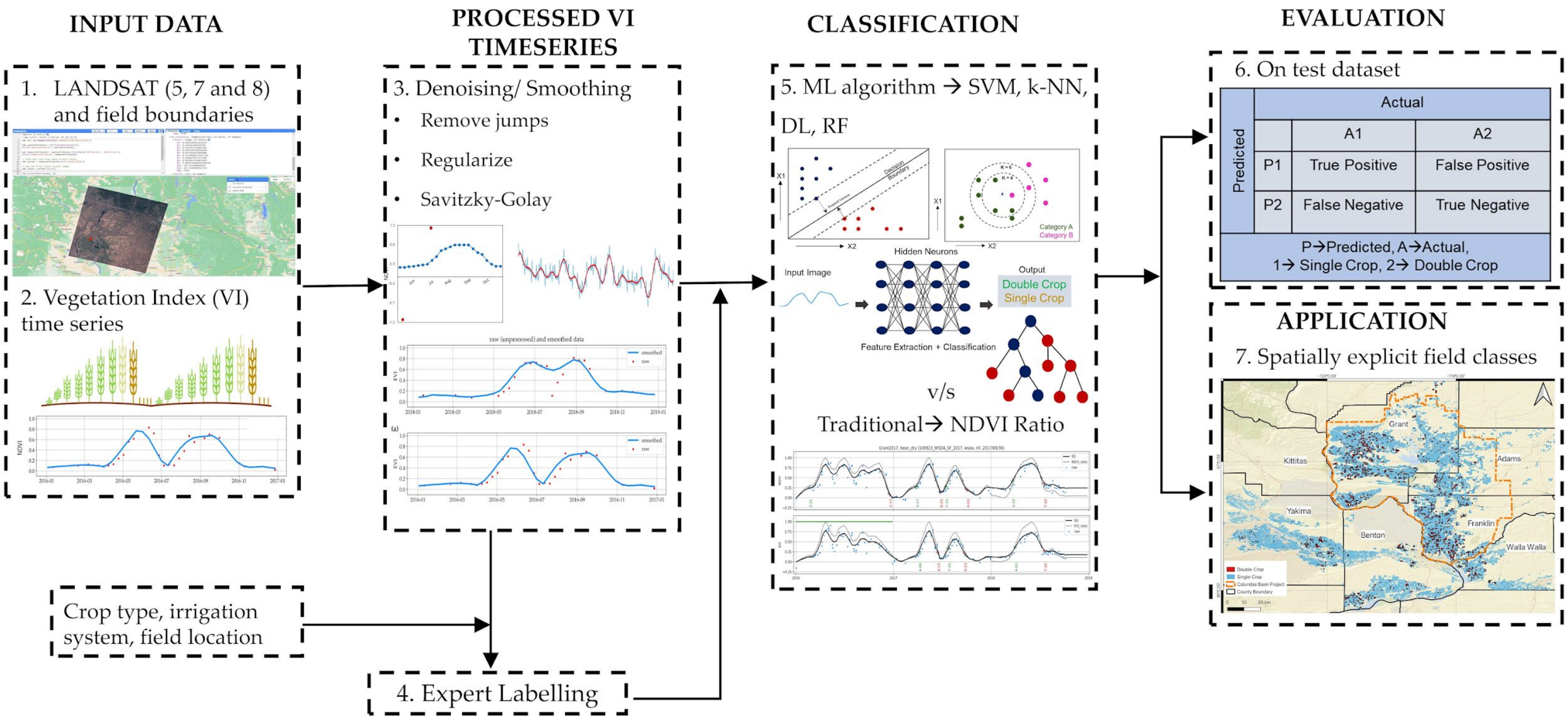
April 30, 2024

ARL 3 to 6/7



ARL 9
Approved, Operational Deployment and Use in Decision Making (Sustained Use)
ARL 8
Application Completed and Qualified (Functionality Proven)
ARL 7
Application Prototype in Partner's Decision Making (Functionality Demonstrated)
ARL 6
Demonstration in Relevant Environment (Potential Demonstrated)
ARL 5
Validation in Relevant Environment (Potential Determined)
ARL 4
Initial Integration and Verification (Prototype/Plan)
ARL 3
Proof of Application Concept (Viability Established)
ARL 2
Application Concept (Invention)
ARL 1
Basic Research (Baseline Ideas)
Integration into Partner's System
PHASE III
Development, Testing, & Validation
PHASE II
Discovery & Feasibility
PHASE I
2

The Process



INPUT DATA
1. LANDSAT (5, 7 and 8) and field boundaries
2. Vegetation Index (VI) time series
PROCESSED VI TIMESERIES
3. Denoising/ Smoothing
• Remove jumps
• Regularize
• Savitzky-Golay
raw (unprocessed) and smoothed data
Crop type, irrigation system, field location
4. Expert Labelling
CLASSIFICATION
5. ML algorithm → SVM, k-NN, DL, RF
Category A
Category B
Input Image
Hidden Neurons
Output
Double Crop
Single Crop
Feature Extraction + Classification
v/s
Traditional→ NDVI Ratio
EVALUATION
6. On test dataset
Actual
Predicted
A1
A2
P1
True Positive
False Positive
P2
False Negative
True Negative
P→Predicted, A→Actual,
1→ Single Crop, 2→ Double Crop
APPLICATION
7. Spatially explicit field classes
Grant
Kittitas
Adams
Yakima
Benton
Franklin
Walla Walla

Overall Accuracies

actual	predicted	SVM	DL	kNN	RF	NDVI-ratio
single	single	556 – 563	555 – 568	554 – 562	562 – 571	517-536
double	double	46 – 55	49 – 55	43 – 48	41 – 44	43-48
single	double	10 – 17	5 – 18	11 – 19	2 – 11	37-56
double	single	4 – 13	4 – 10	11 – 16	15 – 18	11-16
# errors		15 – 30	9 – 25	25 – 33	20 – 29	52-67
accuracy		95 – 98%	96 – 99%	95 – 96%	95 – 97%	89-91%
user acc.		73 – 84%	74 – 92%	70 – 80%	79 – 95%	46-54%
producer acc.		78 – 93%	83 – 93%	73 – 81%	70 – 75%	72-81%

Overall accuracies a good across all models

Even the rule-base NDVI method fares decent

But misleading, as majority data points are single crops

Focusing on just the double-cropped category

actual	predicted	SVM	DL	kNN	RF	NDVI-ratio
single	single	556 – 563	555 – 568	554 – 562	562 – 571	517-536
double	double	46 – 55	49 – 55	43 – 48	41 – 44	43-48
single	double	10 – 17	5 – 18	11 – 19	2 – 11	37-56
double	single	4 – 13	4 – 10	11 – 16	15 – 18	11-16
# errors		15 – 30	9 – 25	25 – 33	20 – 29	52-67
accuracy		95 – 98%	96 – 99%	95 – 96%	95 – 97%	89-91%
user acc.		73 – 84%	74 – 92%	70 – 80%	79 – 95%	46-54%
producer acc.		78 – 93%	83 – 93%	73 – 81%	70 – 75%	72-81%

Producer accuracy: what fraction of ground-truth double crop fields were correctly classified?

Deep Learning and SVM models -> better performance

Focusing on just the double-cropped category

actual	predicted	SVM	DL	kNN	RF	NDVI-ratio
single	single	556 – 563	555 – 568	554 – 562	562 – 571	517-536
double	double	46 – 55	49 – 55	43 – 48	41 – 44	43-48
single	double	10 – 17	5 – 18	11 – 19	2 – 11	37-56
double	single	4 – 13	4 – 10	11 – 16	15 – 18	11-16
# errors		15 – 30	9 – 25	25 – 33	20 – 29	52-67
accuracy		95 – 98%	96 – 99%	95 – 96%	95 – 97%	89-91%
user acc.		73 – 84%	74 – 92%	70 – 80%	79 – 95%	46-54%
producer acc.		78 – 93%	83 – 93%	73 – 81%	70 – 75%	72-81%

User accuracy: what fraction of mapped double crop fields were correctly classified?

Random Forest and Deep Learning models -> better performance

NDVI-ratio overestimating double cropped extent

Summary and comparison with a recent global-scale product

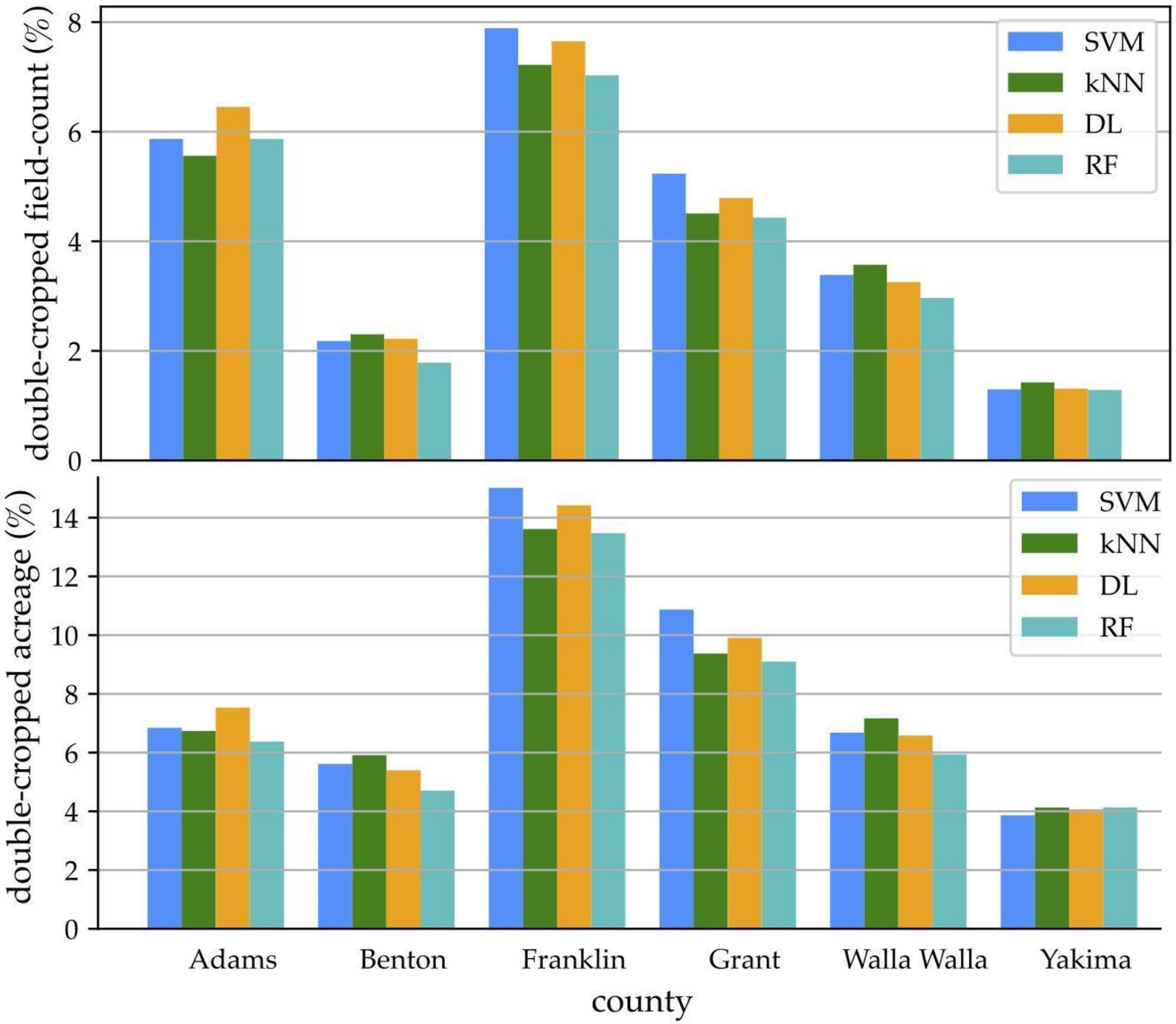
- Overall, our current estimates are that ~10% of the study region is double cropped

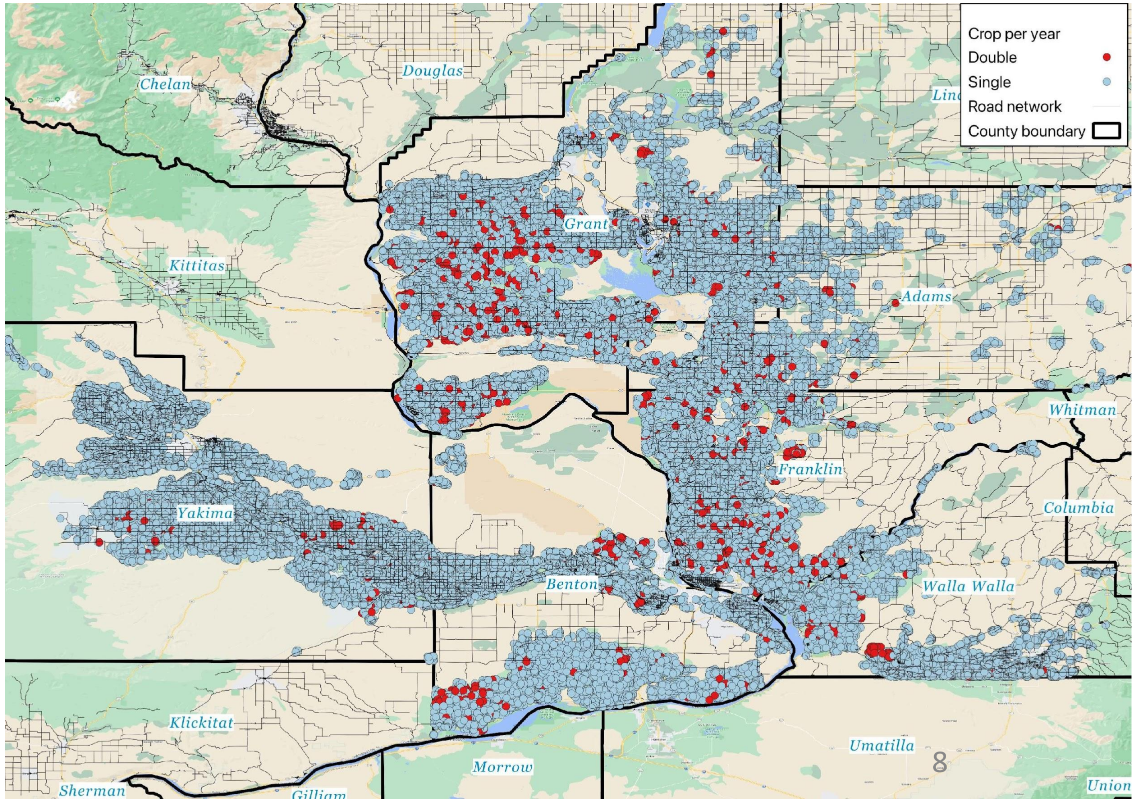
- The global product predicted ~2.5 times the double cropped extent as compared to our work

  -> Consistent with our observation of overestimation

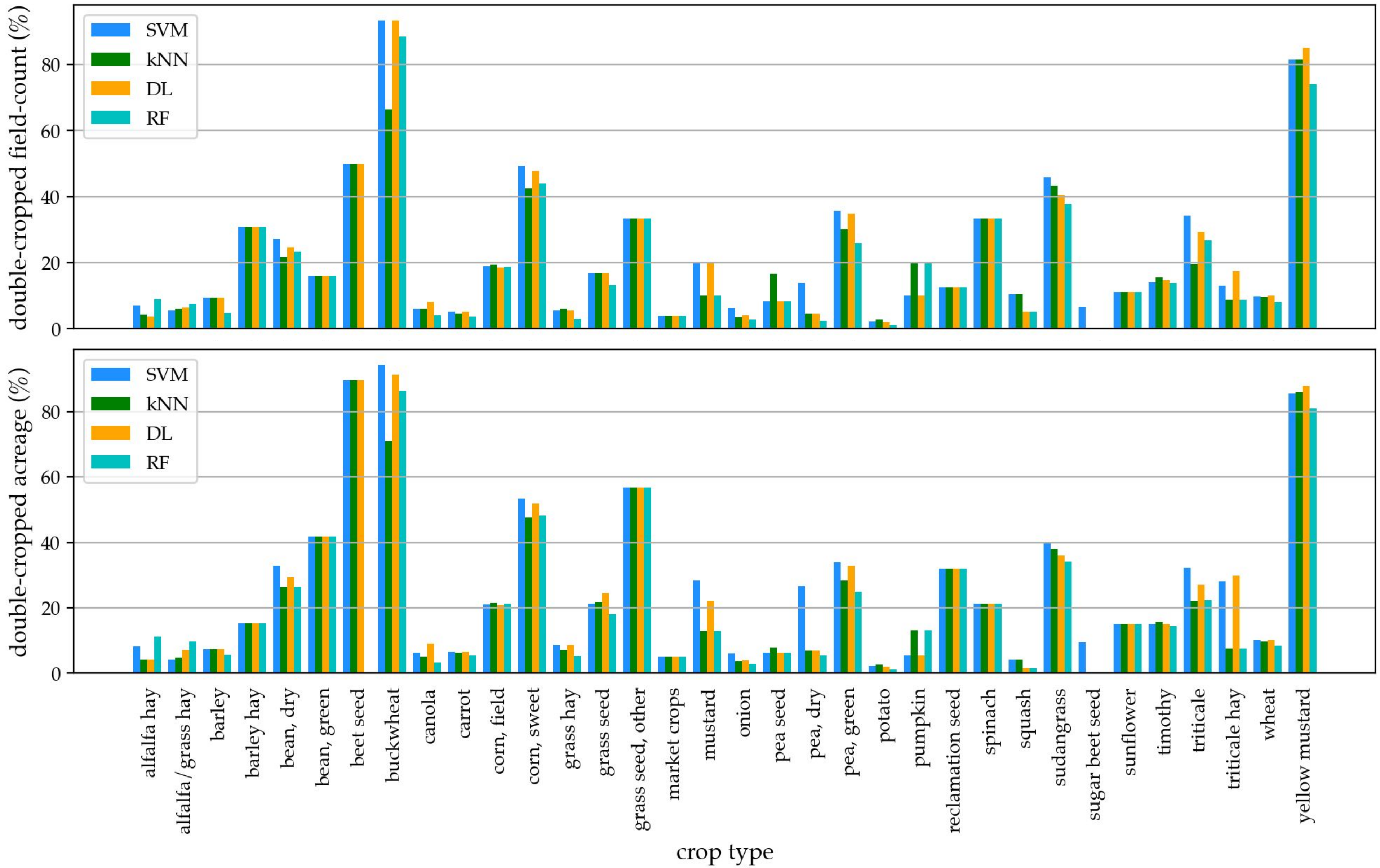
  ->Perennial fields incorrectly coded as double cropped

Qualitative agreement: regions

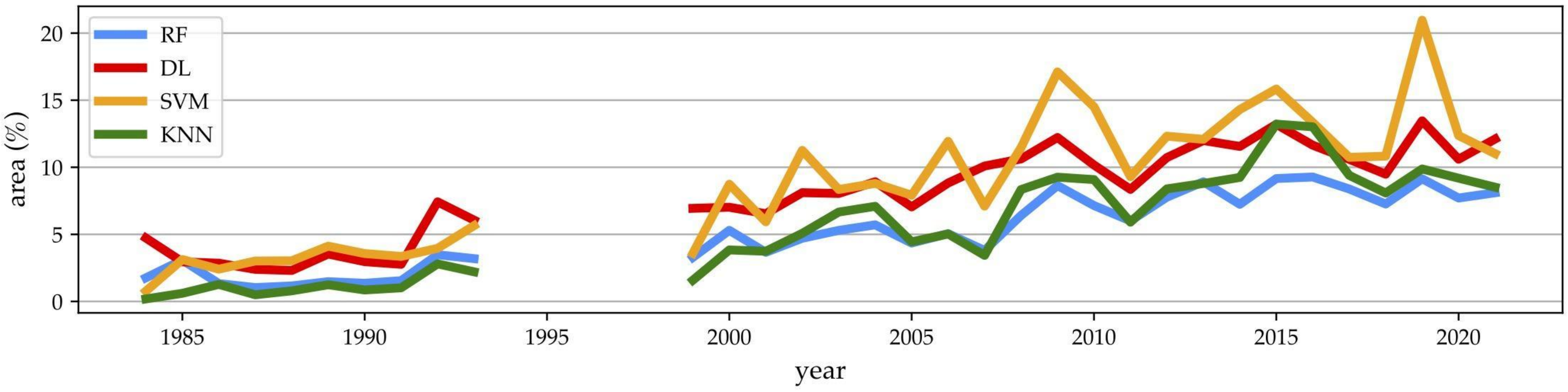




Qualitative agreement: crops



Long-term trend



Workflow for WSDA

- Python workflow and documentation as a Google CoLab notebook
- Shared with the end user
- Google payments to be explored
- Initial few applications will be in collaboration with WSU as training continues
- Interest in expanding to Oregon and California

Implementation: CoLab Notebook

Mount Google Drive and import my Python modules

Here we are importing the Python functions that are written by Hossein Noorazar and are needed; `NASA_core` and `NASA_plot_core`.

Note: These are on Google Drive now. Perhaps we can import them from GitHub.

```
# Mount YOUR google drive in Colab
from google.colab import drive
drive.mount('/content/drive')
import sys
sys.path.insert(0, "/content/drive/My Drive/Colab Notebooks/")
import NASA_core as nc
import NASA_plot_core as ncp
import GEE_Python_core as gpc
```

```
Mounted at /content/drive
```

Please tell me where to look for the shapefile and other directories

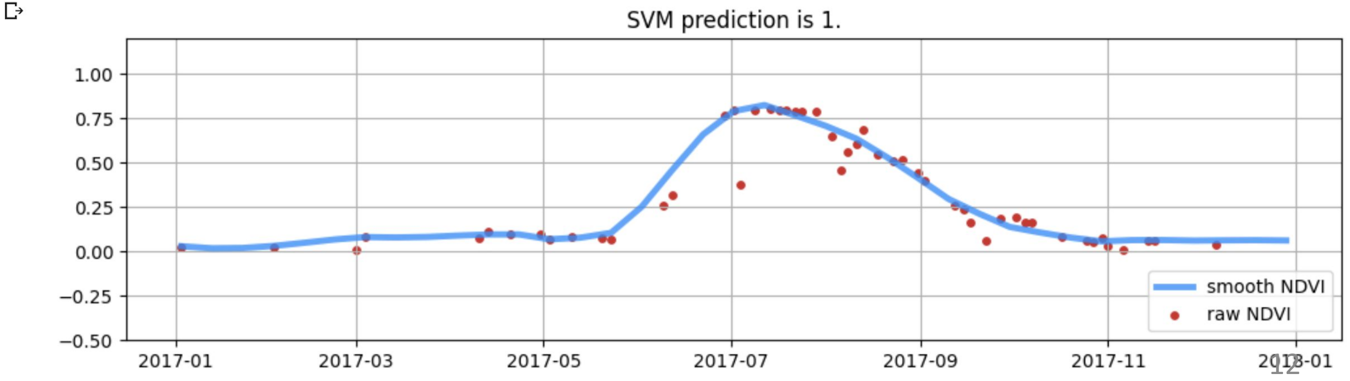
```
[ ] drive_pre     = "/content/drive/My Drive/"
    shp_path_base = drive_pre + "NASA_trends/shapefiles/"
    out_dir_base  = drive_pre + "colab_outputs/"
    Colab_NB_dir  = drive_pre + "Colab Notebooks/"

    import os
    os.chdir(Colab_NB_dir)
    # !ls
```

```
#  Pick a field
an_ID = IDs[3]
a_field = regular_df[regular_df.ID==an_ID].copy()
a_field.sort_values(by='human_system_start_time', axis=0, ascending=True, inplace=True)

# Plot
fig, ax = plt.subplots(1, 1, figsize=(12, 3),
                       sharex='col', sharey='row',
                       gridspec_kw={'hspace': 0.2, 'wspace': .05});
ax.grid(True);
ax.plot(a_field['human_system_start_time'], a_field[VI_idx],
        linestyle='-',  linewidth=3.5, color="dodgerblue", alpha=0.8,
        label=f"smooth {VI_idx}")

# Raw data where we started from
raw = reduced[reduced.ID==an_ID].copy()
raw.sort_values(by='human_system_start_time', axis=0, ascending=True, inplace=True)
ax.scatter(raw['human_system_start_time'], raw[VI_idx], s=15, c='#d62728', label=f"raw {VI_idx}");
label_ = list(predictions.loc[predictions.ID==an_ID, "SVM_NDVI_preds"])[0]
ax.set_title(f"SVM prediction is {label_}.")
ax.legend(loc="lower right");
plt.ylim([-0.5, 1.2]);
```



Future Applications and Extensions

- Water supply and demand forecasts – Office of the Columbia River (WA DOE)
- Regional crop mix shifts and climate change
- Trial methodology in Oregon – Oregon Water Resources Department