Deep density estimation based on multi-spectral remote sensing data for in-field crop yield forecasting
Supplementary Material

Liana Baghdasaryan
Intelinair, Inc.
Yerevan, Am
liana@intelinair.com

Razmik Melikbekyan
Intelinair, Inc.
Yerevan, Am
razmik@intelinair.com

Arthur Dolmajain
Intelinair, Inc.
Yerevan, Am
arthur@intelinair.com

Jennifer Hobbs
Intelinair, Inc.
Chicago, IL, USA
jennifer@intelinair.com

1. Agronomic Indices

Agronomic indices are commonly featured in remote sensing and computational agriculture. These indices capture the reflectance ratio of different channels and have been demonstrated to correspond to properties such as biomass, photosynthetic output, stress, and others. The abbreviations, names, formulas, and brief descriptions of ones explored in this work are given in Table 1.

Index	Name	Formula	Property
NDVI	Normalized Difference Vegetation Index	$\frac{NIR-Red}{NIR+Red}$	Indicator of chlorophyll
NDWI	Normalized Difference Water Index	$\frac{Green-NIR}{NIR+Green}$	Water content
SAVI	Soil Adjusted Vegetation Index	$1.5*\frac{NIR-Green}{NIR+Green+0.5}$	Designed to minimize soil brightness
EVI	Enhanced Vegetation Index	$2.5*\frac{NIR-Red}{NIR+(6*Red)-(7.5*Blue+1}$	NDVI corrected for atmospheric and soil signals especially in the presence of dense canopy
GRNDVI	Green-Red Normalized Difference Vegetation Index	$\frac{NIR-(Green+Red)}{NIR+(Green+Red)}$	Reported better correlation to leaf area index.
GNDVI	Green Normalized Difference Vegetation Index	$\frac{NIR-Green}{NIR+Green}$	Estimates photosynthetic activity and nitrogen uptake in plant canopy

Table 1. Agronomic indices

In all fractional equations, a small factor of $\epsilon = 10^{-15}$ may be added to the denominator in practice for stability.

2. Tabular Model

2.1. Agronomic Feature Generation

Areas of low normalized difference vegetation index (NDVI) are known to correlate with low biomass or plant stress and therefore identifying such areas is a common approach in hand-crafted agronomic algorithms. To construct relevant features, we first run an anomaly detection algorithm based on NDVI. We generate an NDVI map from the original image according to the standard formulas in Table 1. After NDVI and GNDVI representations of the field are created, erosion and blurring is performed to remove noise. Next the NDVI and GNDVI "images" are thresholded at three levels to create three severity levels. These values along with the green and red differential (i.e. difference from the mean of the image) are used to describe each region on the field.

With the guidance of agronomists, we craft rules to classify each region as containing one or more of the following: (a) high stress, b) low biomass, c)low vigor, and d)low (relative) growth. Rules are based on the size and severity of the NDVI and GNDVI anomaly area, RGBN percentiles, and difference between the mean RGBN values in those areas vs. the overall field.

2.2. Hyperparameters

Hyperparameters for each tabular model were determined using the bayesian optimization algorithm in Scikit-Optimize and are shown in Table 2.

Model	Parameters
Lasso Regression	$\alpha = 0.5$
Random Forest	Max Depth = 10 Min Samples Per Leaf = 5 Min Samples Per Split = 3 Num Estimators = 133
LGBM	Boosting Type = Gradient Boosting Decision Tree Learning Rate = 0.17 Max Depth = 6 Num Estimators 54 Num Leaves = 50 Objective = *l2* Max Bins = 15

Table 2. Hyperparameters for tabular models

2.3. Additional Experiments on Input Feature Representation

Agronomic indices such as NDVI, SAVI, and are prominently featured in machine-learning applications for agricultural applications. These indices capture the reflectance ratio of different channels and have been demonstrated to correspond to properties such as biomass, photosynthetic output, stress, and others. In this set of experiments we look to see if incorporating these additional features (Table 1) can boost the performance of the tabular models.

For each of the three algorithms (Lasso, Random Forest, LightGBM), we incorporated different combinations of agronomic indices and features in addition to the raw RGBN channels. In several cases we also incorporated the latitude (lat) and longitude (long), or explicit number of growing degree days (GDD) into the model. Results are shown in Table 3.

Model	Input	Validation			Test		
		MSE	MAE	MAPE	MSE	MAE	MAPE
Lasso	NDVI, NDWI, SAVI, EVI, GRNDVI	788.37	21.69	10.17	597.48	19.16	8.84
Lasso	NDVI, NDWI, SAVI, EVI, GRNDVI, Stress, LowGrowth	744.59	21.30	9.93	565.80	18.7	8.61
Lasso	planter, NDVI, NDWI, SAVI, EVI, GRNDVI	707	20.7	9.80	578	19.0	8.78
Lasso	planter, NDVI, NDWI, SAVI, EVI, GRNDVI, Stress, LowGrowth	668	20.3	9.54	552	18.5	8.55
RF	red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI	757	21.4	10.02	537	18.1	8.36
RF	red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI, Stress, LowGrowth	732	21.1	9.85	514	17.7	8.18
RF	planter, red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI	705	20.5	9.70	526	17.9	8.27
RF	planter, red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI, GDD, lat, long	653	19.7	9.45	556	18.2	8.45
RF	planter, red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI, Stress, LowGrowth	681	20.2	9.53	508	17.6	8.16
RF	planter, red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI, GDD, lat, long, Stress, LowGrowth	644	19.4	9.21	493	17.5	8.08
RF	feature selection	636	19.5	9.27	488	17.5	8.09
LightGBM	red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI	777	21.7	10.17	535	18.0	8.31
LightGBM	red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI, Stress, LowGrowth	792	22.2	10.30	515	17.7	8.10
LightGBM	planter, red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI	689	20.1	9.60	555	18.5	8.57
LightGBM	planter, red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI, GDD, lat, long	583	18.4	8.80	495	17.2	7.97
LightGBM	planter, red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI, Stress, LowGrowth	636	19.4	9.23	497	17.3	8.00
LightGBM	planter, red, green, blue, nir, NDVI, NDWI, SAVI, EVI, GRNDVI, GDD, lat, long, Stress, LowGrowth	561	18.4	8.73	438	16.3	7.53
LightGBM	feature selection	576	18.5	8.77	423	16.0	7.36

Table 3. Impact of agronomic-indices as inputs to tabular model

3. Additional Experiments and Results

3.1. Tile-level Validation and Test Results

This table shows both validation and test results from the tile-level models. The test results mirror those in the main text.

	Tile-Level: Validation			Tile-Level: Test			Field-Level: Validation			Field-Level: Test		
	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE
Naive Baseline	1038.02	24.99	11.99	753.13	21.19	10.06	682.87	20.31	9.54	575.24	19.25	8.62
Lasso (best)	667.74	20.33	9.54	551.67	18.51	8.55	455.79	16.61	7.81	455.04	17.94	7.99
RF (best)	635.90	19.49	9.27	488.12	17.48	8.09	430.20	15.55	7.43	352.72	15.54	6.96
LGBM (best)	576.44	18.50	8.77	423.31	16.0	7.36	372.18	14.73	7.05	270.62	13.13	5.95
Lasso (image)	744.59	21.30	9.93	565.80	18.67	8.61	525.87	17.32	8.11	459.37	17.59	7.84
RF (image)	732.08	21.05	9.85	513.61	17.71	8.18	523.56	17.61	8.26	408.83	16.44	7.36
LGBM (image)	791.57	22.18	10.30	515.39	17.67	8.10	560.12	18.49	8.64	435.60	16.41	7.35
VGG16	594.78	19.28	8.76	374.21	15.21	7.08	511.46	17.78	7.96	378.28	15.45	6.85
ResNet-34	594.87	19.26	8.83	359.57	15.28	7.11	428.64	16.45	7.49	281.53	13.69	6.16
ResNet-50	618.56	19.34	8.80	370.01	14.47	6.75	389.24	15.58	7.06	251.57	12.99	5.78
RegnetY-040	573.32	18.73	8.57	347.26	14.53	6.77	388.48	15.75	7.12	251.39	13.08	5.82
Densenet-161	774.88	21.95	9.83	482.70	17.34	7.83	533.10	17.98	8.06	390.05	14.78	6.38

Table 4. Performance of Tile-Level Regression Models

3.2. Pixel Validation and Test Results

This table shows both validation and test results from the pixel-level model. The test results mirror those in the main text.

	Tile-Level: Validation			Tile-Level: Test			Field-Level: Validation			Field-Level: Test		
	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE
U-Net VGG16	596.20	19.48	8.92	504.82	17.46	7.82	412.31	16.48	7.66	392.38	15.96	7.06
U-Net ResNet-34	552.59	18.49	8.39	395.24	15.61	7.03	332.04	14.35	6.64	274.49	13.26	5.91
U-Net ResNet-50	519.77	17.68	8.13	365.36	15.19	6.88	338.67	14.11	6.61	296.47	13.42	5.98
U-Net RegnetY-040	550.36	18.08	8.26	394.96	15.72	7.11	326.26	13.71	6.41	256.30	13.12	5.89
U-Net DenseNet-161	525.79	17.63	8.06	379.54	15.42	6.98	305.53	12.91	6.04	255.23	12.88	5.78
FPN VGG16	646.57	20.15	9.07	525.50	17.74	7.83	441.66	16.63	7.58	436.88	12.24	7.09
FPN ResNet-34	587.01	18.92	8.57	371.26	15.18	6.84	345.91	14.77	6.85	269.58	13.33	5.91
FPN ResNet-50	591.89	18.95	8.62	395.23	15.68	7.06	357.36	15.04	6.99	264.34	12.92	5.73
FPN RegnetY-040	601.12	18.95	8.67	421.49	16.38	7.38	341.71	14.15	6.62	261.72	13.04	5.80
FPN DenseNet-161	560.25	18.34	8.30	347.01	14.77	6.59	318.65	13.93	6.44	234.97	12.29	5.41

Table 5. Performance of Pixel-Level Regression Models

A plot of the actual vs. predicted values and actual vs. residuals for the U-Net DenseNet-161 model are shown in Figure 1

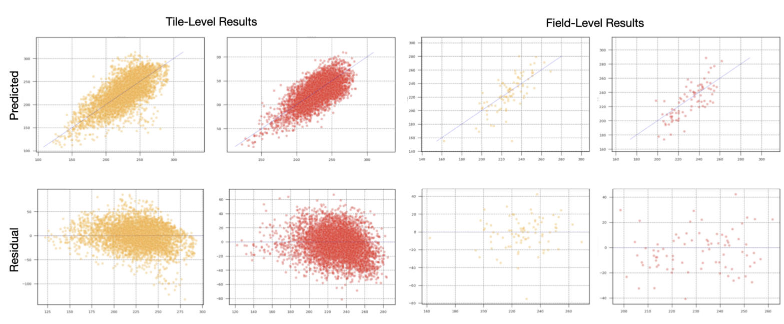

Figure 1. Actual vs. Predicted and Residuals of the U-Net DenseNet-161 model plotted for validation (yellow) and test set (red) per tile (top two rows) and per field (bottom two rows).

3.3. Validation and Test for Additional Inputs to Pixel-Model

Validation results corresponding to Section 6.3 in the main text. Test results mirror those reported there.

	Tile-Level: Validation			Tile-Level:Test			Field-Level: Validation			Field-Level:Test		
UNet DenseNet-161	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE
RGBN	525.79	17.63	8.06	379.54	15.42	6.98	305.53	12.91	6.04	255.23	12.88	5.78
RGBN+NDWI+SAVI	544.20	17.95	8.24	389.64	15.50	7.04	307.05	13.10	6.18	263.61	13.09	5.89
RGBN+Planter	489.40	17.17	7.89	381.76	15.30	6.93	281.63	12.61	5.92	268.61	13.32	5.98
RGBN+Stress	558.52	18.2	8.30	356.21	14.90	6.72	327.68	13.63	6.33	234.89	12.06	5.40
RGBN+All Outliers	530.13	17.70	8.07	394.99	15.68	7.11	311.99	13.38	6.24	267.32	13.42	6.04

Table 6. Impact of additional feature channels