

Drone RGB Image-Derived Vegetation Indices to Predict Yield in Humidicola (*Brachiaria humidicola*) grass

ALBERT F. ASTILLERO, MSc.

Department of Agriculture
Bureau of Animal Industry, Philippines

EN-CHUNG LIN, Ph.D.

National Taiwan University, Taiwan



Presented during the 75th EAAP 2024 Congress
September 1-5, 2024 in Florence, Italy

Outline of presentation

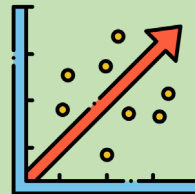
2



Introduction
Objective



Materials and
Methods



Results &
Discussion



Conclusion and
Perspective

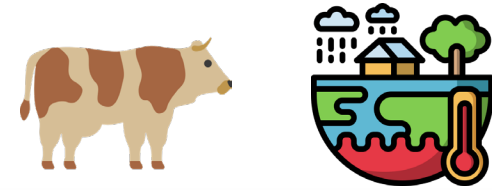
Importance of forage yield monitoring



pasture
productivity



pasture management
decision



animal productivity

environmental sustainability

climate change adaptability

Prevent overgrazing & pasture degradation

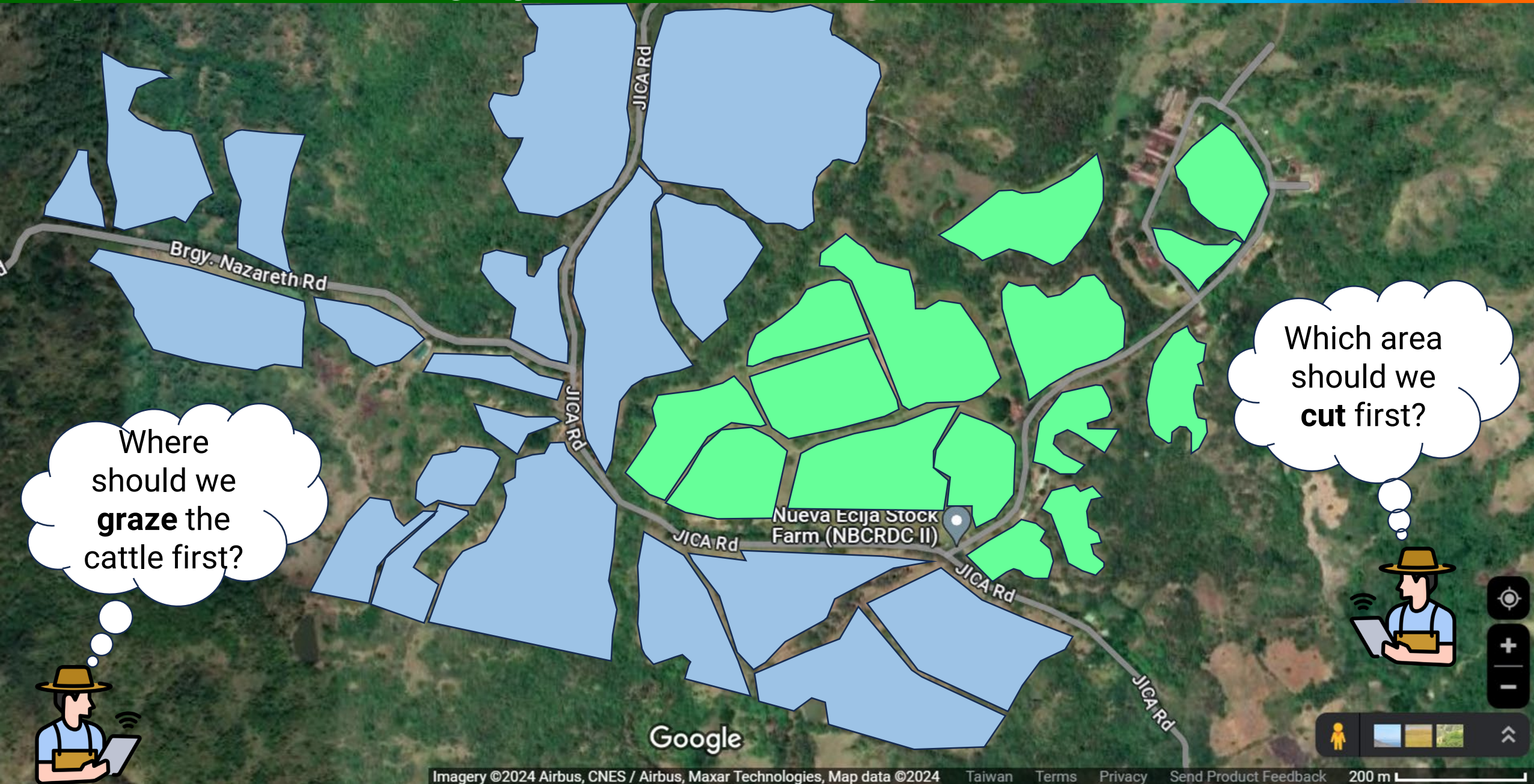


Photo: Albert F. Astillero



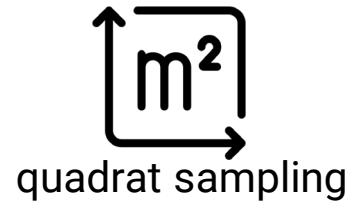
Photo: BAI KPL Quintos

Importance of forage yield monitoring

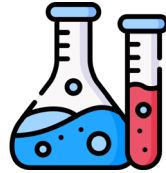


Ways to measure forage yield

manual



laboratory

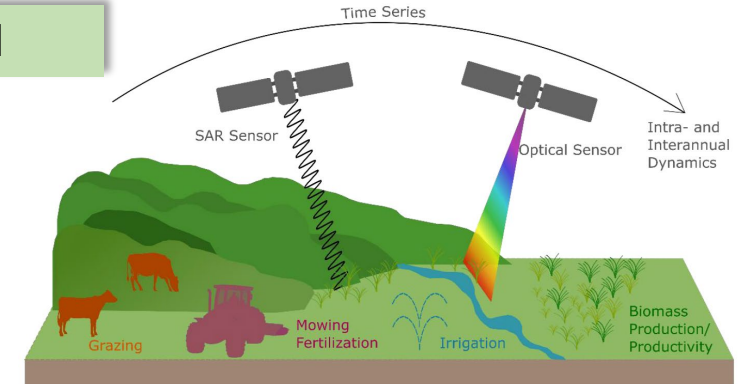


ground-based sensor

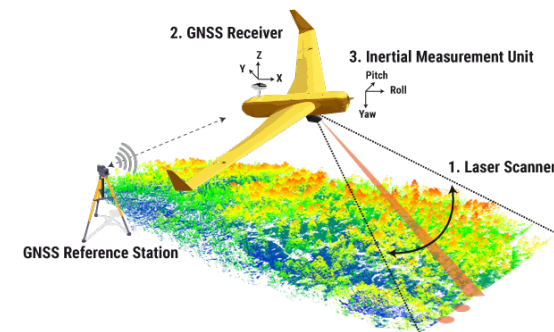


tractor-pulled multi-sensors

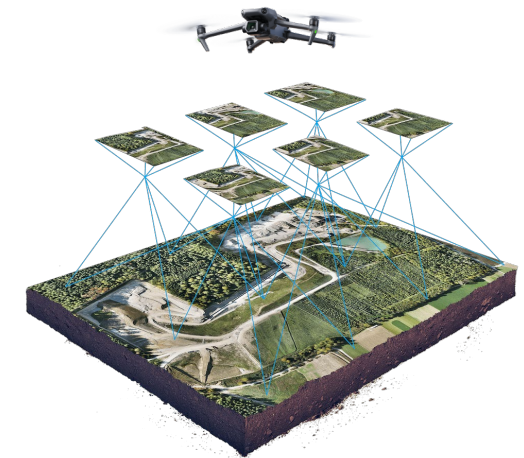
satellite-based



drone-based

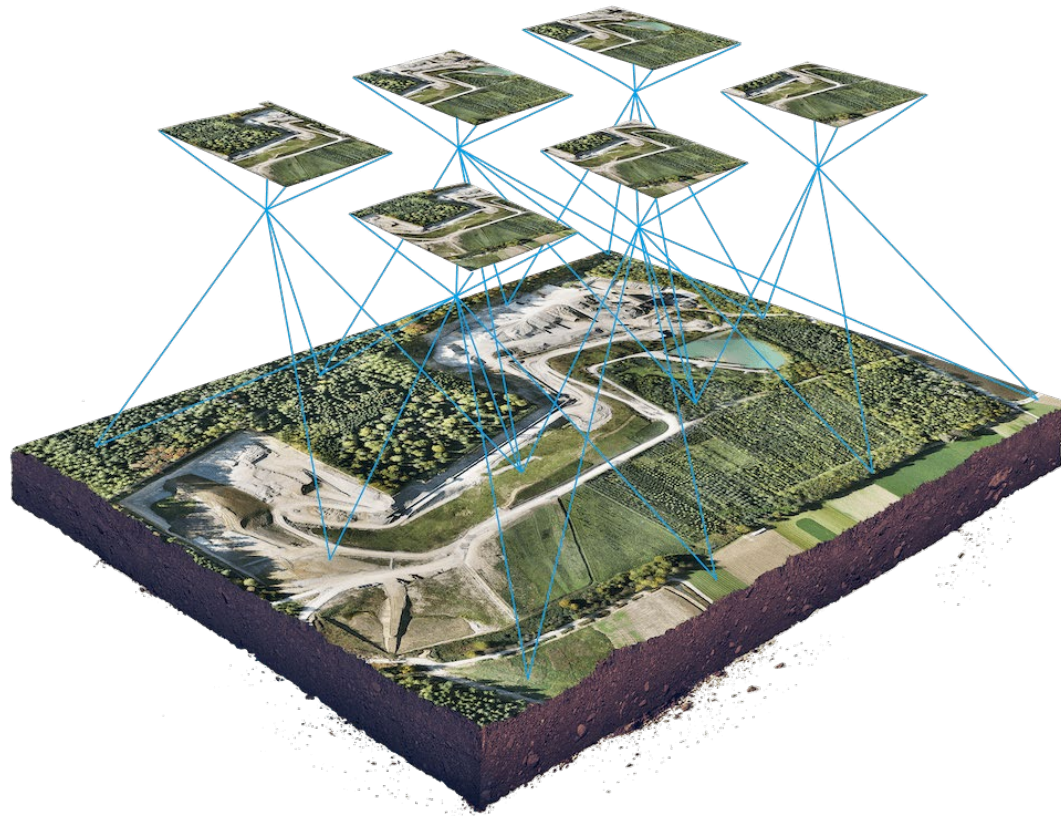


Light Detection and Ranging
(LiDAR)



Photogrammetry + GIS

Ways to measure forage yield



SfM-MVS Photogrammetry

UAV

UAV type

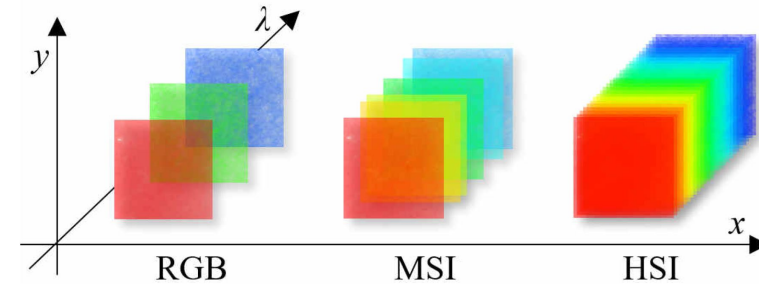


multi-rotor

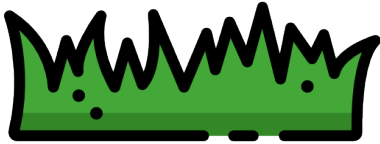


fixed-wing

sensor/camera



ENVIRONMENT



forage species & pasture type



sample size

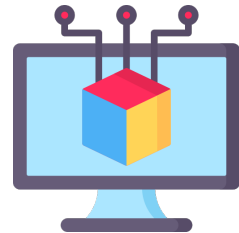


sampling frequency

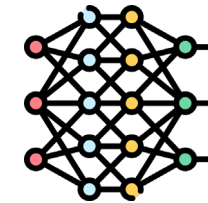
TOOLS and METHODS



cheaper drone + RGB
vs. others



photogrammetry vs.
simple workflow



statistical vs
machine learning model



statistical prediction model

Fresh Matter (FM)



Dry Matter (DM)



Study site: location

National Beef Cattle Research and Development Center II, Nueva Ecija, Philippines
15°22'45.0"N 121°09'46.7"E



Source: Google Map imagery 2024
 Date retrieved: January 2, 2024

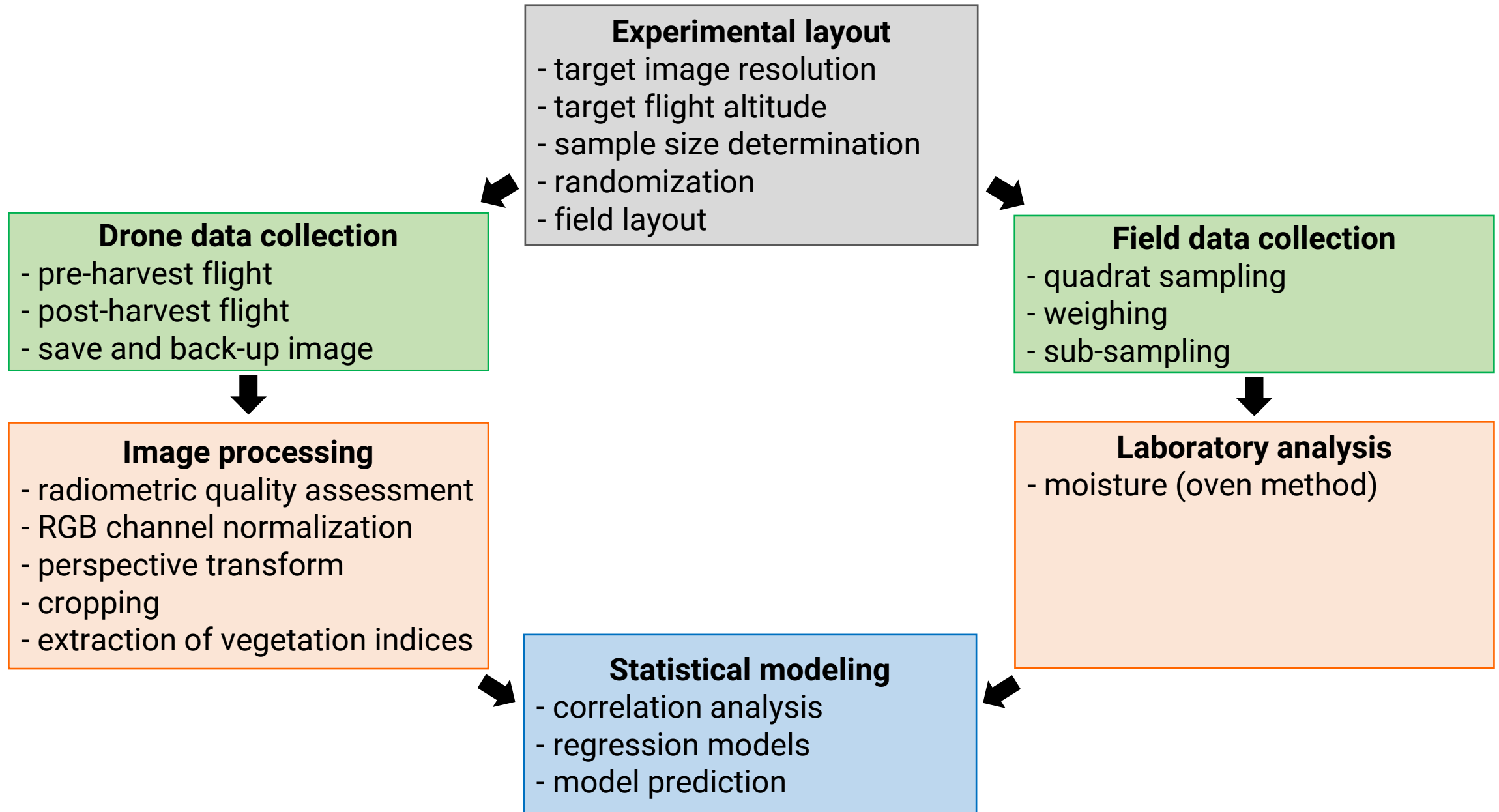
Study site: selection criteria

10



1. pure *B. humidicola* grass
2. relatively flat terrain
3. minimal drone hazards (trees)
4. areas are closer to each other
5. with fence
6. approved by the farm manager





Methodology: Experimental layout

1. Ground sampling distance (GSD) or resolution and flight altitude



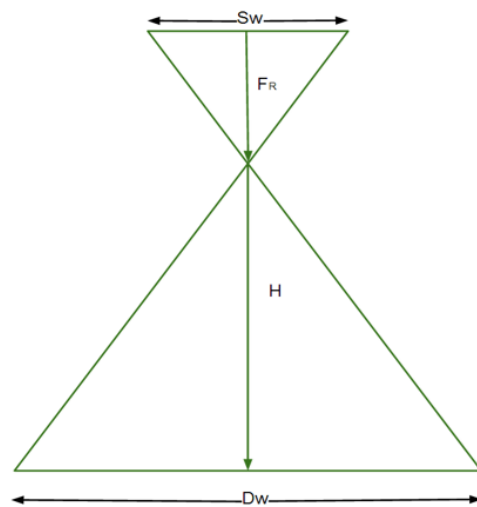
GROUND SAMPLING DISTANCE CALCULATOR

Instructions

1. Enter the Sensor Width (millimeters) in cell **C14**
2. Enter the Focal Length (millimeters) in cell **C15** (real focal length, not 35 mm equivalent)
3. Enter the Flight Height (meters) in cell **C16**
4. Enter the Image Width (pixels) in cell **C17**
5. [Optional] Enter the Image Height (pixels) in cell **C18**
6. Hit Enter

Calculator

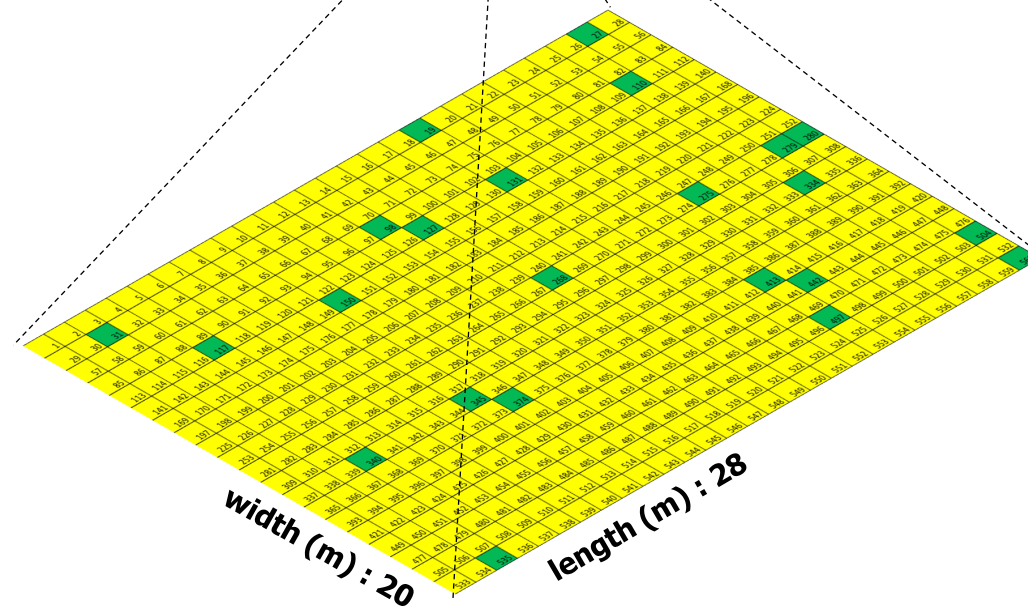
Sw	18	= the sensor width of the camera (millimeters)
Fr	12.49	= the focal length of the camera (millimeters)
H	30	= the flight height (meters)
imW	5280	= the image width (pixels)
imH	3956	= the image height (pixels)
GSD	0.82	= Ground Sampling Distance (centimeters/pixel)
Dw	43	= width of single image footprint on the ground (meters)
DH	32	= height of single image footprint on the ground (meters)



Dw = The footprint width / distance covered on the ground by one image in width direction

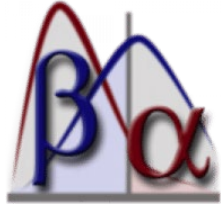
www.pix4d.com

Altitude (m): 20

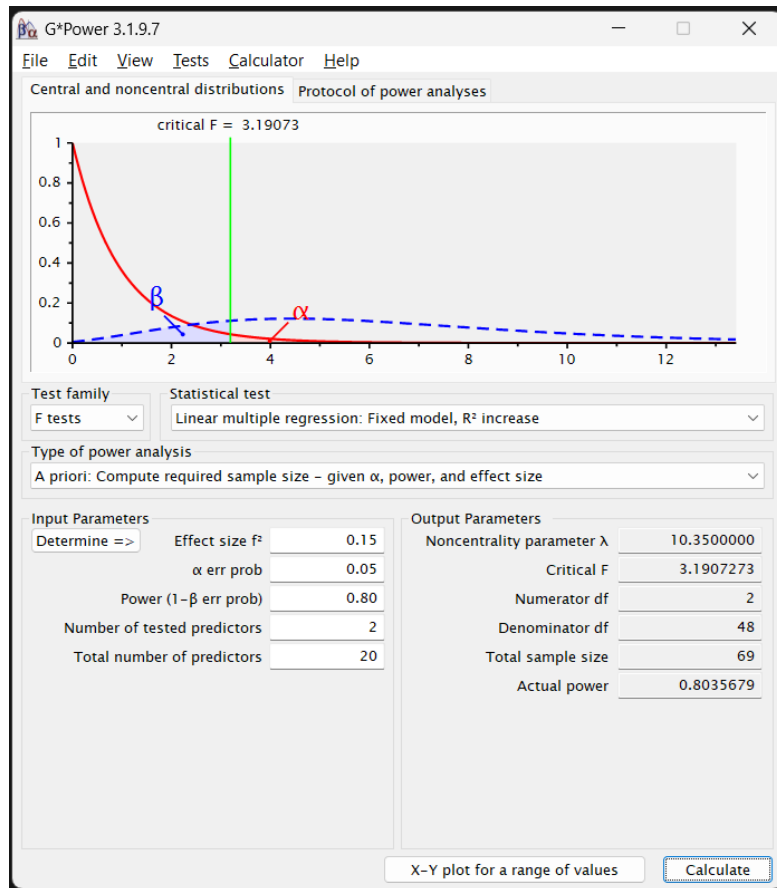


Methodology: Experimental layout

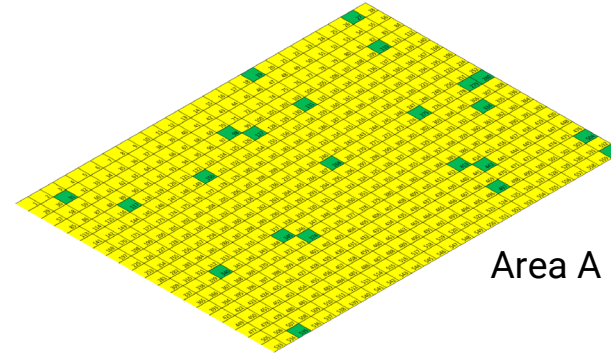
2. Sample size determination



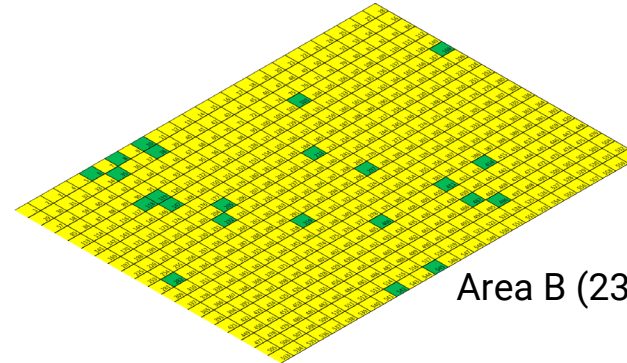
Gpower ver. 3.1



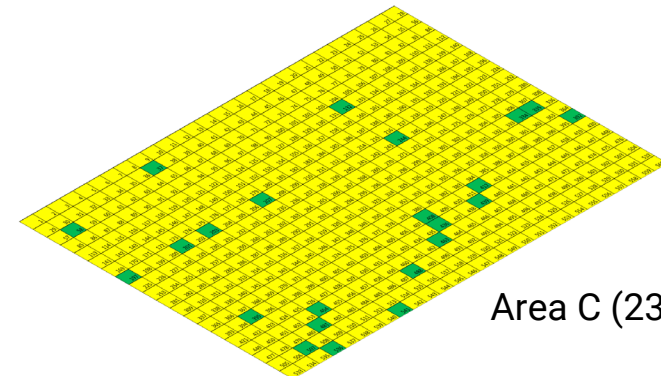
3. Randomization of sampling areas



Area A (23 samples)



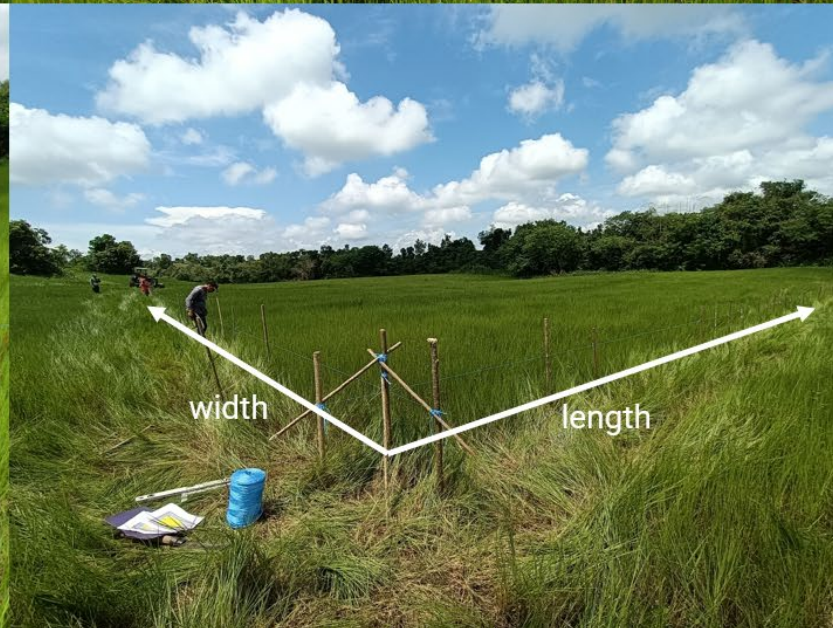
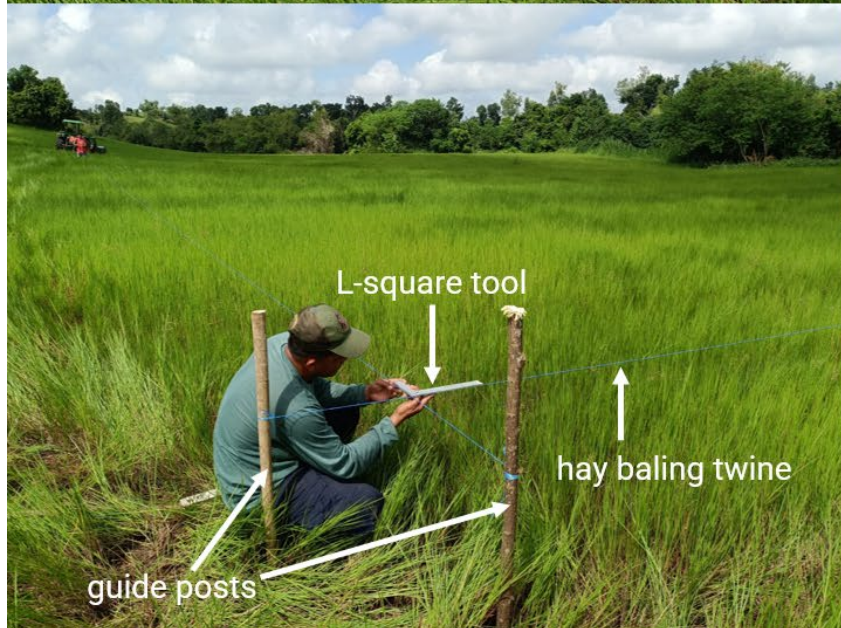
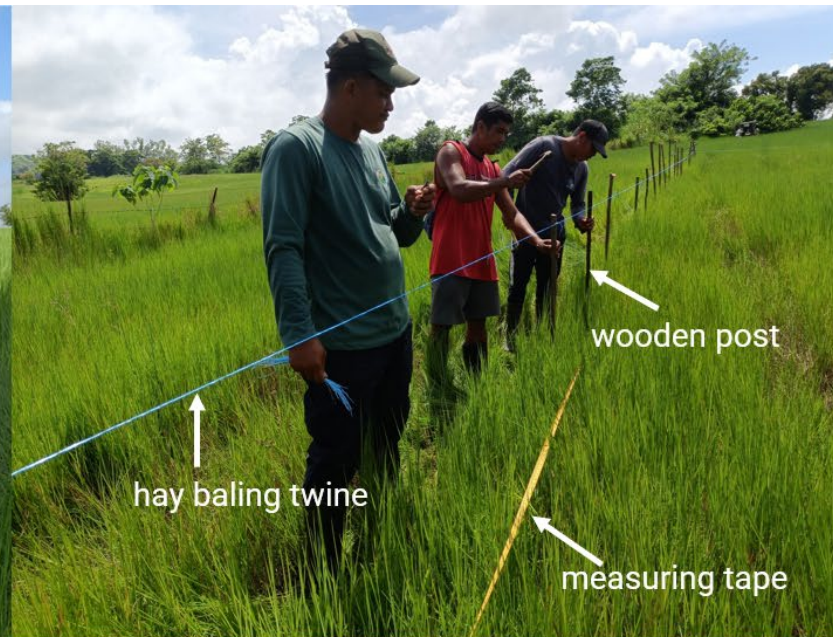
Area B (23 samples)



Area C (23 samples)

Methodology: Experimental layout

15



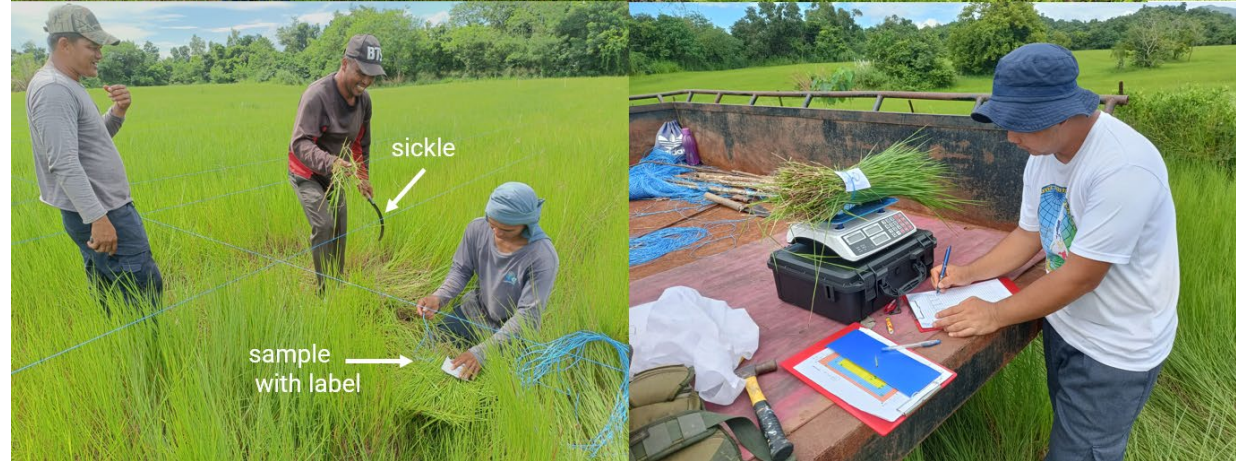
Methodology: Drone data collection



Methodology: Field data collection

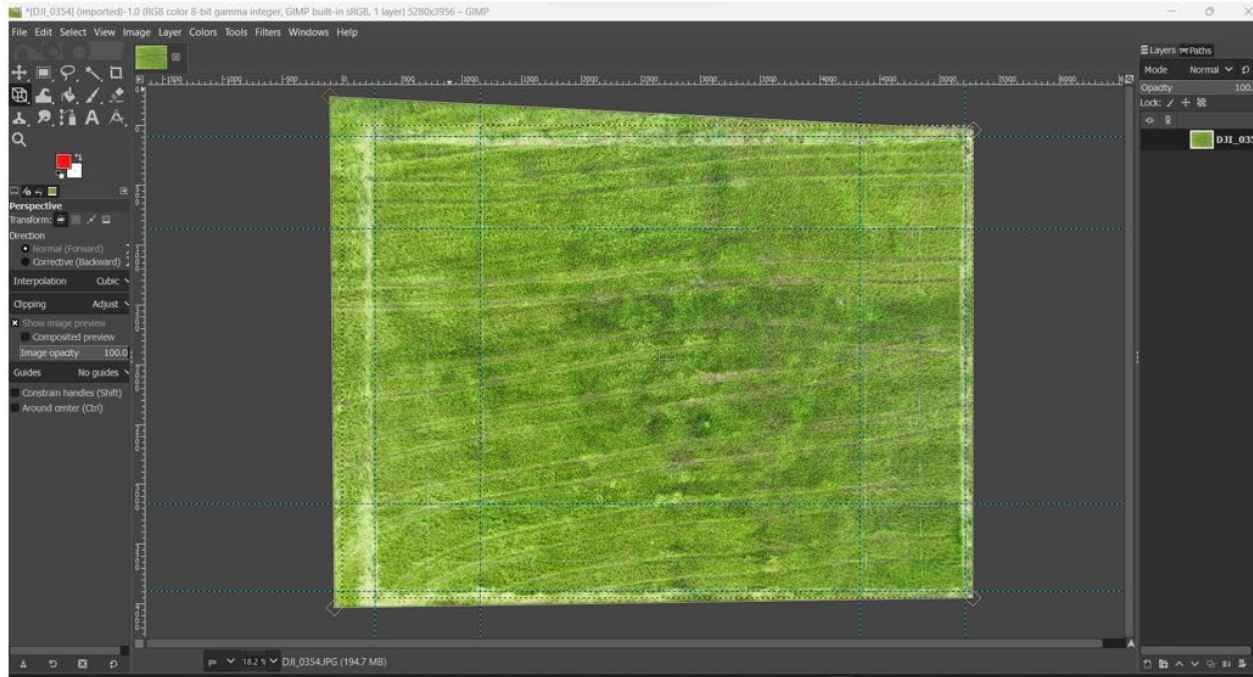


Forage sampling

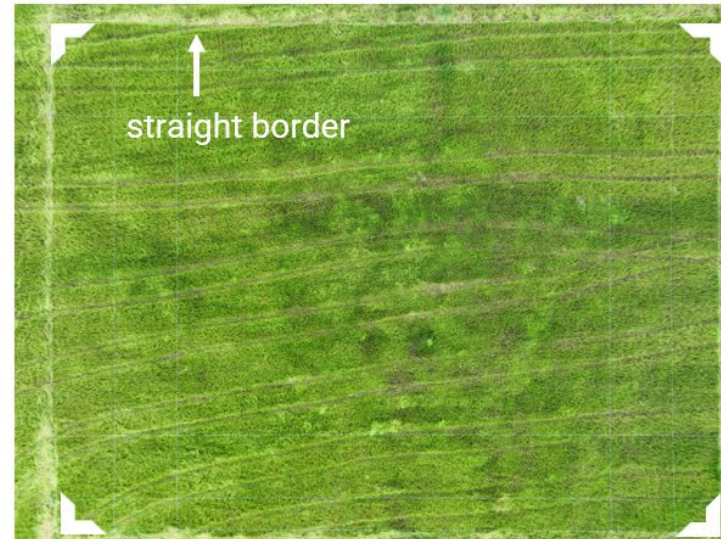
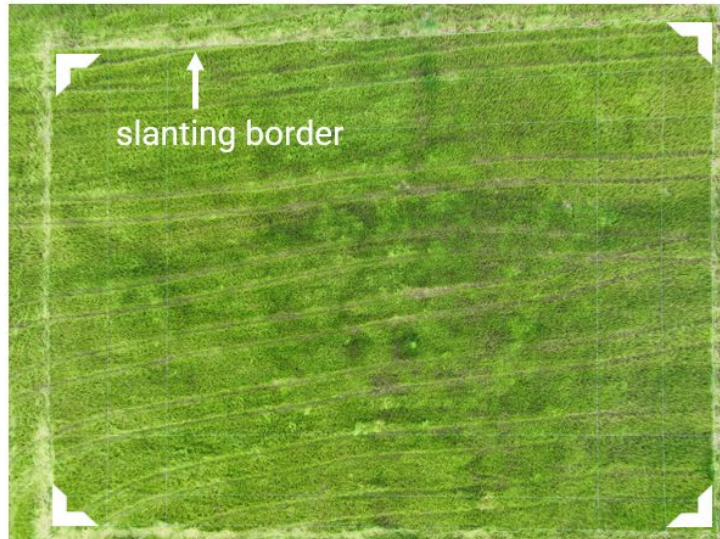


Sub-sampling and preparation for lab analysis

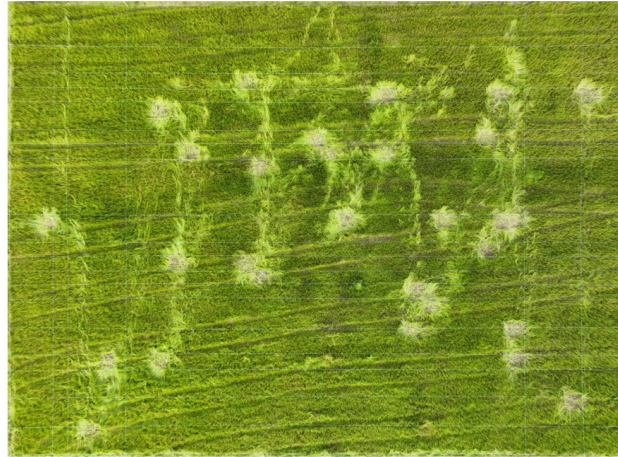
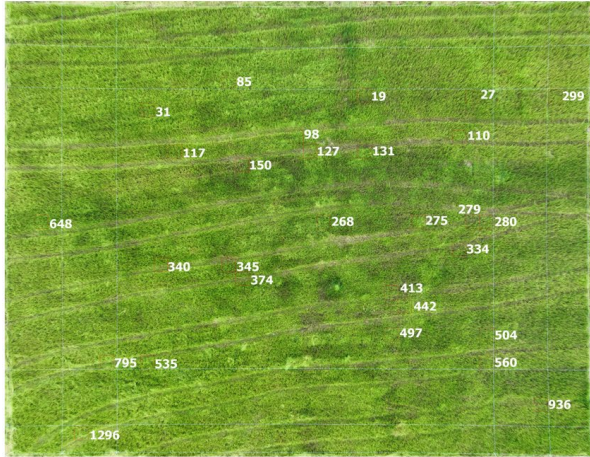
Methodology: Image processing



Perspective correction and cropping

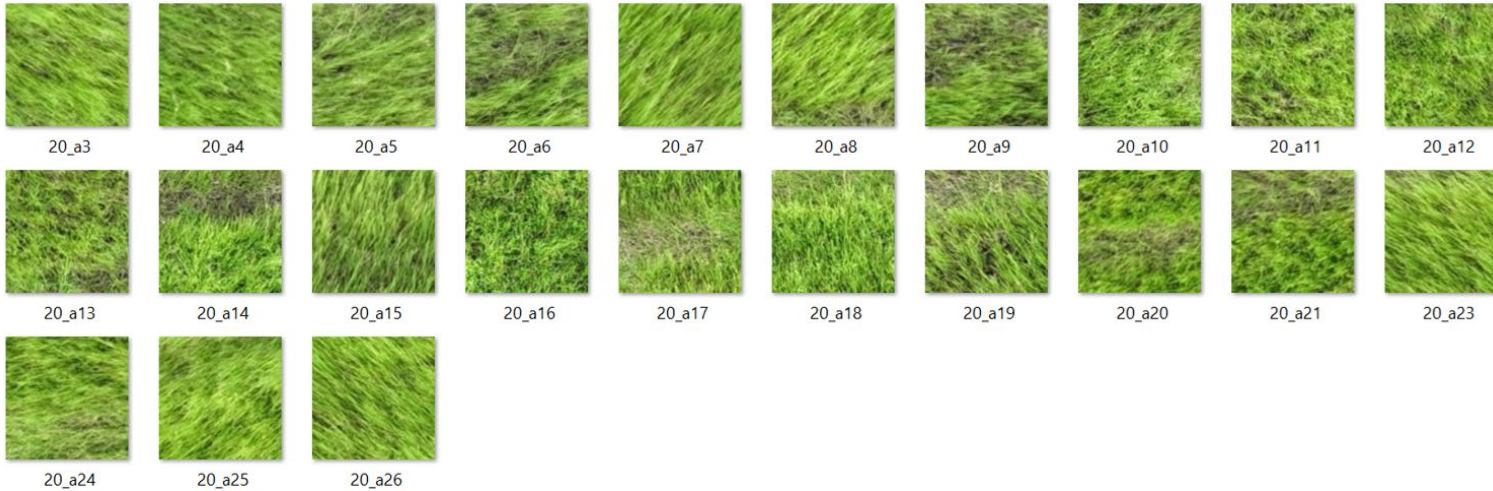


Methodology: Image processing

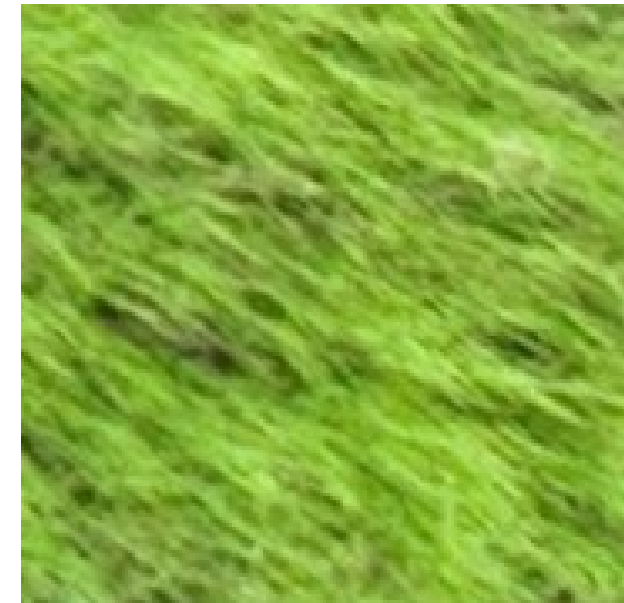


cropping guide (before and after harvest)

Cropping individual images



cropped images



cropped image dimension (180 x 180 pixel)

Methodology: Image processing

Vegetation Index	Formula*	Usage	Reference
Excess Green Index (<i>ExG</i>)	$2 \times g - r - b$	Forage yield	Viljanen et al. (2018) referred from Woebbecke et al. (1995); Borra-Serrano et al. (2019)
Excess Red Index (<i>ExR</i>)	$1.4 \times r - b$	Forage yield and quality	Viljanen et al. (2018) referred from Meyer et al. (1999); Oliveira et al. (2022)
Excess Green-Red Index (<i>ExGR</i>)	$ExG - ExR$	Forage yield	Borra-Serrano et al. (2019) referred from Camargo Neto (2004)
Green Leaf Index (<i>GLI</i>)	$\frac{(2 \times G - R - B)}{(2 \times G + R + B)}$	Forage yield	Pranga et al. (2021) referred from Gobron et al. (2000)
Normalized Green Red Difference Index (<i>NGRDI</i>)	$\frac{G - R}{G + R}$	Forage yield	Lussem et al. (2018) referred from Tucker (1979)
Modified Green Red Vegetation Index (<i>MGRVI</i>)	$\frac{(G^2 - R^2)}{(G^2 + R^2)}$	Forage yield and quality	2018 Viljanen referred from Bendig et al. (2014); Michez 2020 referred from Bendig et al. (2015).
Red Green Blue Vegetation Index (<i>RGBVI</i>)	$\frac{(G^2) - (B \times R)}{(G^2) + (B \times R)}$	Forage yield and quality	Viljanen et al. (2018) referred from Bendig et al. (2015); Oliveira et al. (2022) referred from Bendig et al. (2015)

Normalized Green Blue Index (<i>NGBI</i>)	$\frac{G - B}{G + B}$	Forage yield and quality	Michez et al. (2020) referred from Metternicht (2003) Grijalva-Gómez et al. (2022)
---	-----------------------	--------------------------	--

* Each band is represented as $R = red$, $G = green$, $B = blue$, and derived $r = \frac{R}{(R+G+B)}$,

$$g = \frac{G}{(R+G+B)}, b = \frac{B}{(R+G+B)}.$$

Extraction of vegetation indices

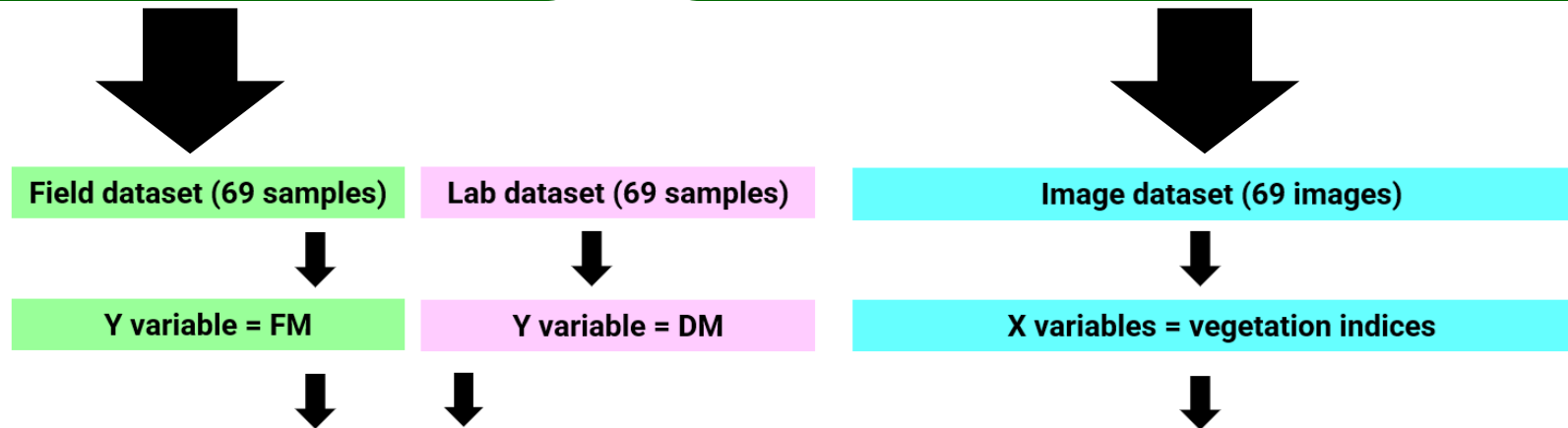
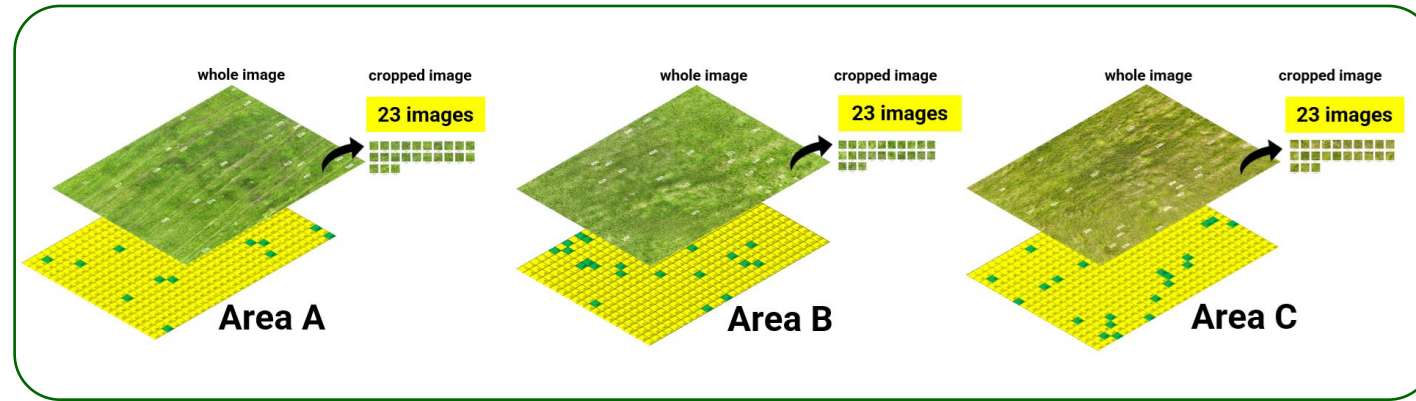
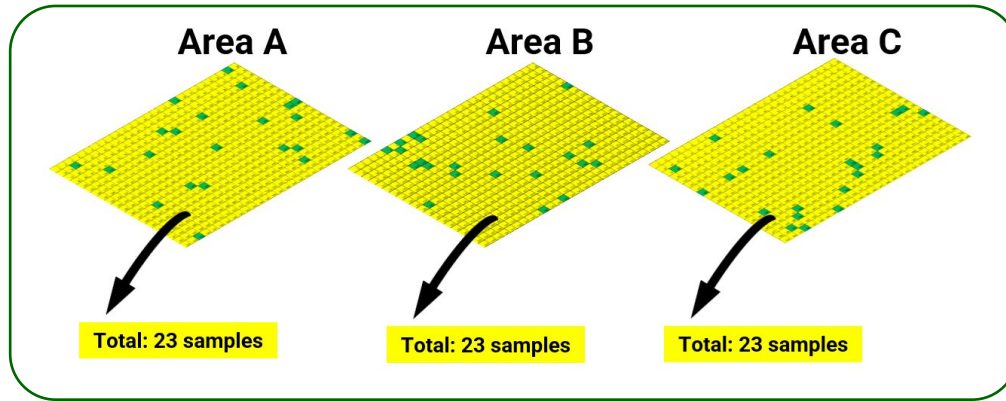


ver. 3.1



ver. 1.3.9

Methodology: Merging the datasets (field, lab, image)



sample	FM	DM	ExG	ExR	ExGR	GLI	NGRDI	MGRVI	RGBVI	NGBI
A03	1.010	0.549	0.4428	-0.1530	0.5958	0.2978	0.5196	0.8007	0.6066	0.1340
A04	1.132	0.558	0.4630	-0.1444	0.6074	0.3099	0.5224	0.8037	0.6189	0.1508
A05	0.923	0.466	0.3603	-0.1020	0.4624	0.2468	0.4209	0.6999	0.5053	0.1121
A06	0.947	0.507	0.3437	-0.0890	0.4327	0.2357	0.3983	0.6678	0.4801	0.1084
A07	1.164	0.667	0.4793	-0.1696	0.6489	0.3196	0.5574	0.8319	0.6451	0.1463
A08	0.887	0.500	0.4282	-0.1574	0.5856	0.2879	0.5159	0.7847	0.5920	0.1219
A09	1.012	0.504	0.3489	-0.1018	0.4507	0.2381	0.4132	0.6773	0.4869	0.1035
A10	1.219	0.607	0.4367	-0.1316	0.5683	0.2928	0.4944	0.7664	0.5847	0.1412

Methodology: Statistical modeling

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

where:

Y = dependent variable

β_0 = intercept term

β_i = slope of the regression factors X_i , $i = 1, 2, \dots, n$

X_i = independent variable, $i = 1, 2, \dots, n$

ε = error term

Faraway (2005)

Forage yield prediction models:

Model 1: Fresh Matter (FM), kg = f (linear VI, quadratic VI, cubic VI) + ε

Model 2: Dry Matter (DM), kg = f (linear VI, quadratic VI, cubic VI) + ε

Correlation and scatterplot

Correlation between X and Y variables



SAS REG Procedure ADJRSQ for model selection

$$R^2 = SSR / SST$$

where:

SSR (Sum of Squares Regression)

SSE (Sum of Squares Error)

SST (Sum of Squares Total) = $SSR + SSE$

$$\text{Adjusted } R^2 = 1 - \left[(1 - R^2) \times \frac{n-1}{n-k-1} \right]$$

where:

R^2 = the R^2 of the model

n = the number of samples

k = the number of predictor variables



Measured vs. predicted

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

where:

$\hat{y}_i$ = predicted values by the single altitude model

y_i = measured values in the field and in the laboratory

n = number of samples

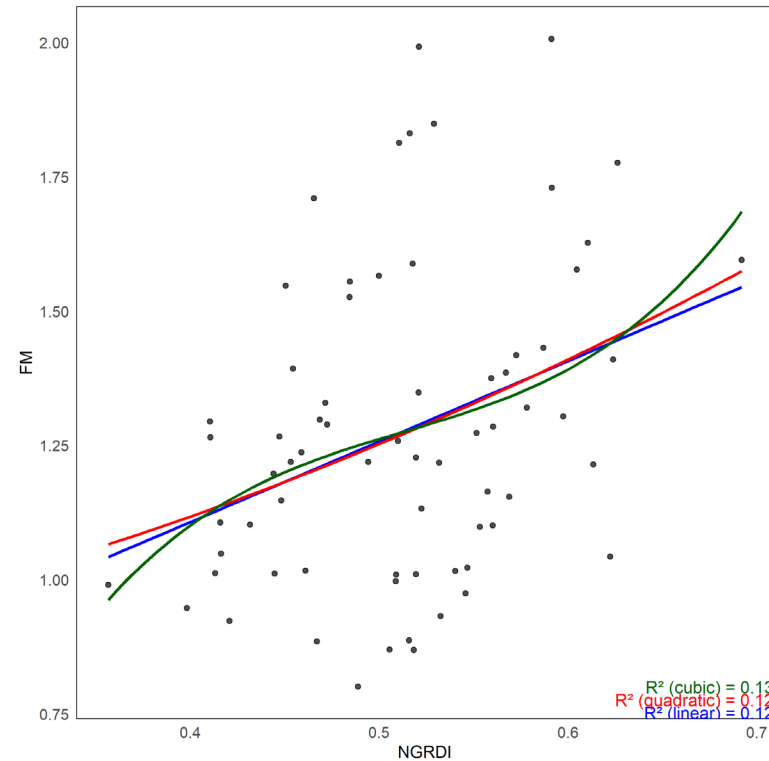
$$NMRSE (\%) = \frac{RMSE}{\text{mean of measured values}}$$



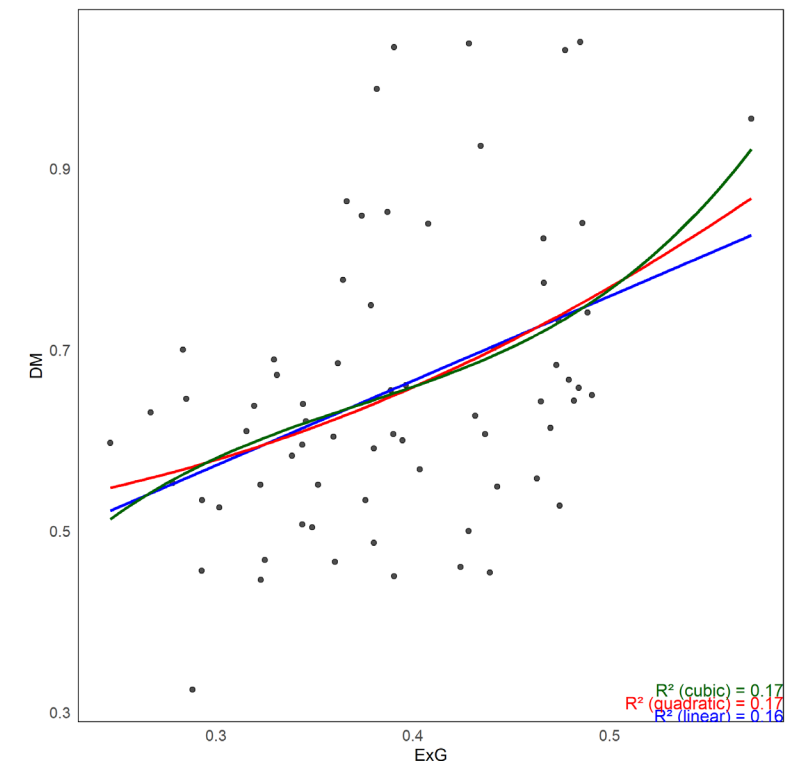
Results and discussion: Correlation analysis and scatterplot

	FM	DM	ExG	ExR	ExGR	GLI	NGRDI	MGRVI	RGBVI	NGBI
FM	1	0.93	0.29	-0.2	0.37	0.28	0.34	0.29	0.33	0.12
DM	0.93	1	0.4	-0.03	0.36	0.4	0.27	0.24	0.36	0.3
ExG	0.29	0.4	1	0.1	0.77	1	0.52	0.56	0.86	0.85
ExR	-0.2	-0.03	0.1	1	-0.56	0.1	-0.79	-0.75	-0.41	0.61
ExGR	0.37	0.36	0.77	-0.56	1	0.77	0.95	0.95	0.98	0.31
GLI	0.28	0.4	1	0.1	0.77	1	0.53	0.56	0.86	0.85
NGRDI	0.34	0.27	0.52	-0.79	0.95	0.53	1	0.99	0.88	0
MGRVI	0.29	0.24	0.56	-0.75	0.95	0.56	0.99	1	0.89	0.04
RGBVI	0.33	0.36	0.86	-0.41	0.98	0.86	0.88	0.89	1	0.47
NGBI	0.12	0.3	0.85	0.61	0.31	0.85	0	0.04	0.47	1

Y vs. X @ 20 m



FM vs. NGRDI



DM vs. ExG

Data has linear and non-linear relationship

Results and discussion: Model selection

The performance of different model selection methods for each response variable

Response variable	Model selection method*	No. of X variables	R ²	Adjusted R ²
Fresh Matter (FM)	Adjusted R²	14	0.7659	0.7052
	Backward	10	0.7514	0.6870
	Stepwise	3	0.3501	0.3201
Dry Matter (DM)	Adjusted R²	13	0.7260	0.6612
	Backward	8	0.7161	0.6287
	Stepwise	4	0.3459	0.3050

Note: Values in bold are the highest for the best model selection method.

Best model selection method: **Adjusted R²**

Results and discussion: Optimum regression model

Regression equation: Fresh Matter (FM) @ 20 m altitude

$$\begin{aligned} \text{Fresh Matter} = & \beta_0 + \beta_1 \text{ExG} + \beta_2 \text{ExR} + \beta_3 \text{GLI} + \beta_4 \text{NGRDI} + \beta_5 \text{RGBVI} + \beta_6 \text{NGBI} \\ & + \beta_7 \text{ExG}^2 + \beta_8 \text{ExR}^2 + \beta_9 \text{ExGR}^2 + \beta_{10} \text{GLI}^2 + \beta_{11} \text{NGRDI}^2 + \beta_{12} \text{NGBI}^2 \\ & + \beta_{13} \text{GLI}^3 + \beta_{14} \text{NGRDI}^3 + \varepsilon \end{aligned}$$

$R^2 = 0.77$
Adjusted $R^2 = 0.71$

Other studies: RGB
 $R^2 = 0.77$ (Batistoti et al., 2019)

Other studies: multispectral
 $R^2 = 0.48$ (Na et al., 2017)

The variables and statistics for the fresh matter regression equation

	Variable	DF	Slope name	Estimate	Std. error	t value	Pr > t
linear: 6	Intercept	1	β_0	150.3698	38.0116	3.96	0.0002
	ExG	1	β_1	1034.7310	220.7954	4.69	<.0001
	ExR	1	β_2	-1026.0900	281.8926	-3.64	0.0006
	GLI	1	β_3	-809.8120	371.3810	-2.18	0.0336
	NGRDI	1	β_4	-1399.0200	204.0757	-6.86	<.0001
	RGBVI	1	β_5	178.3928	115.7565	1.54	0.1291
quadratic: 6	NGBI	1	β_6	619.7890	381.2711	1.63	0.1099
	ExG ²	1	β_7	-1414.3800	233.6360	-6.05	<.0001
	ExR ²	1	β_8	-2323.1700	286.3350	-8.11	<.0001
	ExGR ²	1	β_9	634.0347	175.1353	3.62	0.0006
	GLI ²	1	β_{10}	-3028.6200	452.4019	-6.69	<.0001
	NGRDI ²	1	β_{11}	1268.9160	250.9211	5.06	<.0001
cubic: 2	NGBI ²	1	β_{12}	2715.2520	455.8973	5.96	<.0001
	GLI ³	1	β_{13}	2012.3970	532.5837	3.78	0.0004
	NGRDI ³	1	β_{14}	-405.7690	79.3804	-5.11	<.0001

Results and discussion: Optimum regression model

Regression equation: Dry Matter (DM) @ 20 m altitude

$$\begin{aligned} \text{Dry Matter} = & \beta_0 + \beta_1 \text{ExG} + \beta_2 \text{ExR} + \beta_3 \text{ExGR} + \beta_4 \text{NGRDI} + \beta_5 \text{NGBI} + \beta_6 \text{ExG}^2 \\ & + \beta_7 \text{ExR}^2 + \beta_8 \text{ExGR}^2 + \beta_9 \text{GLI}^2 + \beta_{10} \text{NGBI}^2 + \beta_{11} \text{ExG}^3 + \beta_{12} \text{ExGR}^3 \\ & + \beta_{13} \text{RGBVI}^3 + \varepsilon \end{aligned}$$

$R^2 = 0.73$

Adjusted $R^2 = 0.66$

Other studies: RGB

$R^2 = 0.81$ (Borra-Serrano et al., 2019)

$R^2 = 0.62$ (Lussem et al., 2018)

$R^2 = 0.62$ (Grijalva-Gómez et al., 2022)

$R^2 = 0.60$ (Rueda-Ayala et al., 2019)

$R^2 = 0.59 \sim 0.81$ (Grüner et al., 2019)

Other studies:

$R^2 = 0.92$ (RGB + hyperspectral) (Franceschini et al., 2019)

$R^2 = 0.70$ (RGB + multispectral) (Michez et al., 2019)

The variables and statistics for the dry matter regression equation

Variable	DF	Slope name	Estimate	Std. error	t value	Pr > t
Intercept	1	β_0	25.7797	8.1909	3.15	0.0027
ExG	1	β_1	904.8600	306.4938	2.95	0.0046
ExR	1	β_2	-844.8230	299.2221	-2.82	0.0066
ExGR	1	β_3	-717.9900	297.5880	-2.41	0.0192
NGRDI	1	β_4	-196.5650	29.1911	-6.73	<.0001
NGBI	1	β_5	225.6276	88.0893	2.56	0.0132
ExG ²	1	β_6	-920.8820	185.9445	-4.95	<.0001
ExR ²	1	β_7	-1111.4200	156.2413	-7.11	<.0001
ExGR ²	1	β_8	725.1813	97.8646	7.41	<.0001
GLI ²	1	β_9	-1664.5100	231.7639	-7.18	<.0001
NGBI ²	1	β_{10}	1600.1780	233.2480	6.86	<.0001
ExG ³	1	β_{11}	249.8425	54.0022	4.63	<.0001
ExGR ³	1	β_{12}	-122.0840	20.9793	-5.82	<.0001
RGBVI ³	1	β_{13}	74.8398	16.3255	4.58	<.0001

linear: 5

quadratic: 5

cubic: 3

Results and discussion: Models and vegetation indices (VI)

The combination of vegetation indices used in the model for each response variable

Response variable*	Term**	Vegetation indices (VI)							
		ExG	ExR	ExGR	GLI	NGRDI	MGRVI	RGBVI	NGBI
FM	X	✓	✓		✓	✓		✓	✓
	X^2	✓	✓	✓	✓	✓			✓
	X^3				✓	✓			
DM	X	✓	✓	✓		✓			✓
	X^2	✓	✓	✓	✓				✓
	X^3	✓		✓				✓	

*FM – Fresh Matter, DM – Dry Matter,

** X = linear, X^2 = quadratic, X^3 = cubic

Combination of linear, quadratic, and cubic terms captured the data pattern between Y and X variables

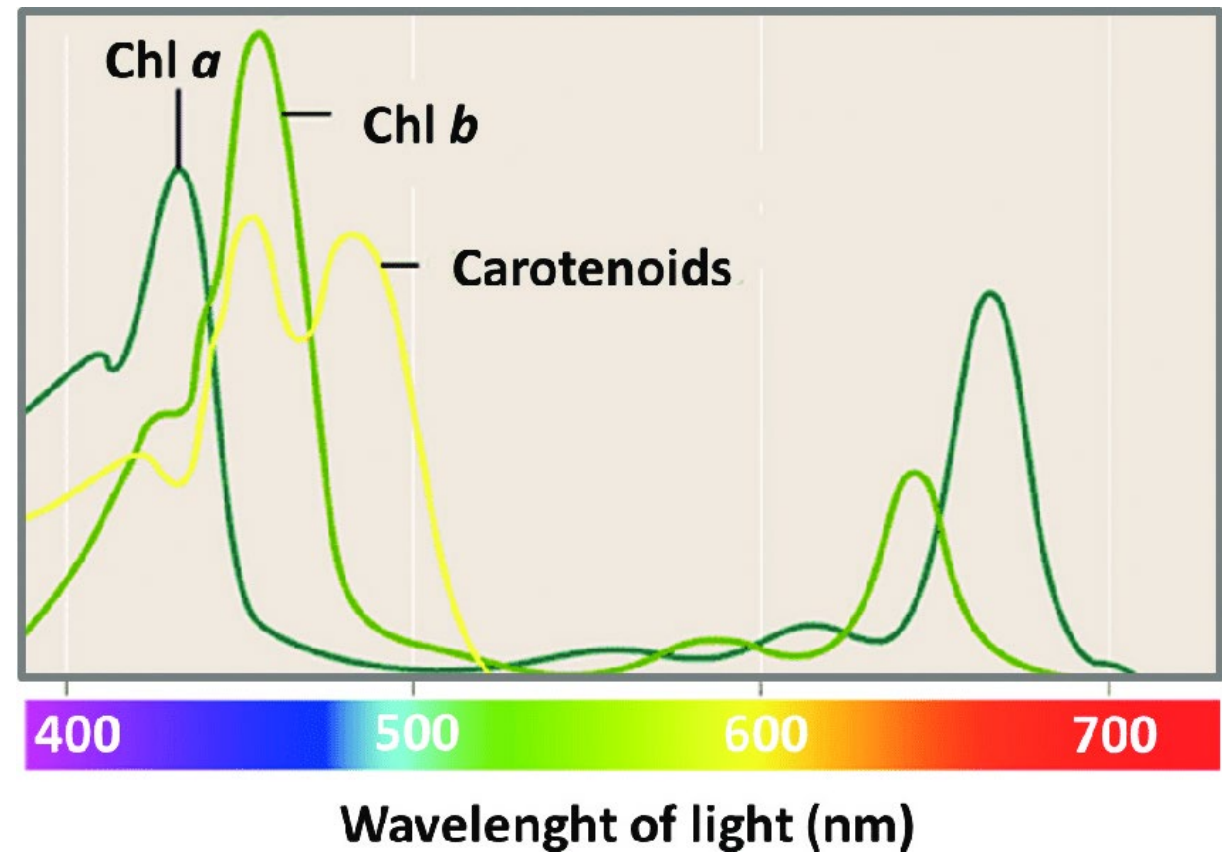
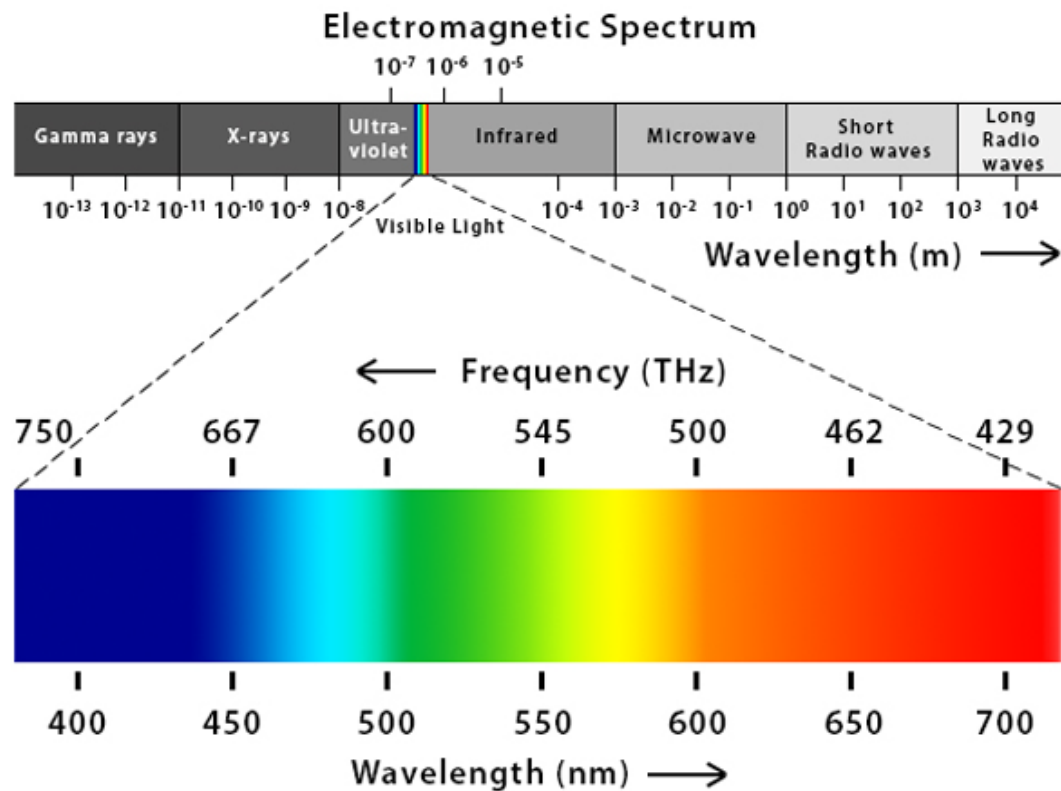
FM: GLI and NGRDI → most used VI

DM: ExG and ExGR → most used VI

All VI emphasizes the greenness of vegetation.
Similar VI used by other authors for forage yield prediction

Results and discussion: Models and vegetation indices (VI)

Visible Light Spectrum



FM and DM: related to the chlorophyll absorption in the visible region
(Curran, 1989; Lugassi et al., 2015)

RGB-based indices are effective predictors
(Lussem et al., 2018; Grüner et al., 2020; Karila et al., 2022))

Results and discussion: Measured vs. Predicted

Comparison between the fresh and dry matter measured vs predicted using the optimum regression equation at 20 m altitude

Parameter	FM measured	FM predicted	DM measured	DM predicted	
Total Mean	1.276	1.276	0.657	0.657	 High similarity
Area A_Mean	1.129	1.175	0.597	0.627	
Area B_Mean	1.405	1.362	0.750	0.721	
Area C_Mean	1.294	1.290	0.624	0.623	
Total Correlation	0.8748		0.8511		 High correlation
Area A_Correlation	0.8038		0.8442		
Area B_Correlation	0.7937		0.7477		
Area C_Correlation	0.9513		0.9104		
Total RMSE (NMRSE)*	0.1392 (10.9%) 		0.0840 (12.8%) 		 Low RMSE & NMRSE
Area A_RMSE (NMRSE)	0.1478 (13.1%)		0.0774 (13.0%)		
Area B_RMSE (NMRSE)	0.1670 (11.9%)		0.1057 (14.1%)		
Area C_RMSE (NMRSE)	0.0917 (7.1%)		0.0630 (10.1%)		

*RMSE = Root Mean Square Error, NMRSE = Normalized Root Mean Square Error

Other studies:
 RGB + Hyperspectral camera
 FM NMRSE: **19%**
 DM NMRSE: **21%**
 (Karila et al., 2022)

Conclusion and perspective



Accuracy: RGB-based VI can accurately predict FM & DM



Practicality: Simple methodology (no photogrammetry, open-source software, cheap tools) is fine but sample size and randomization is priority



Computationally efficient: image processing algorithms and statistical model was sufficient



Scalability: farm to nationwide level if combined with satellite imagery and GIS



More samples
(whole production season)



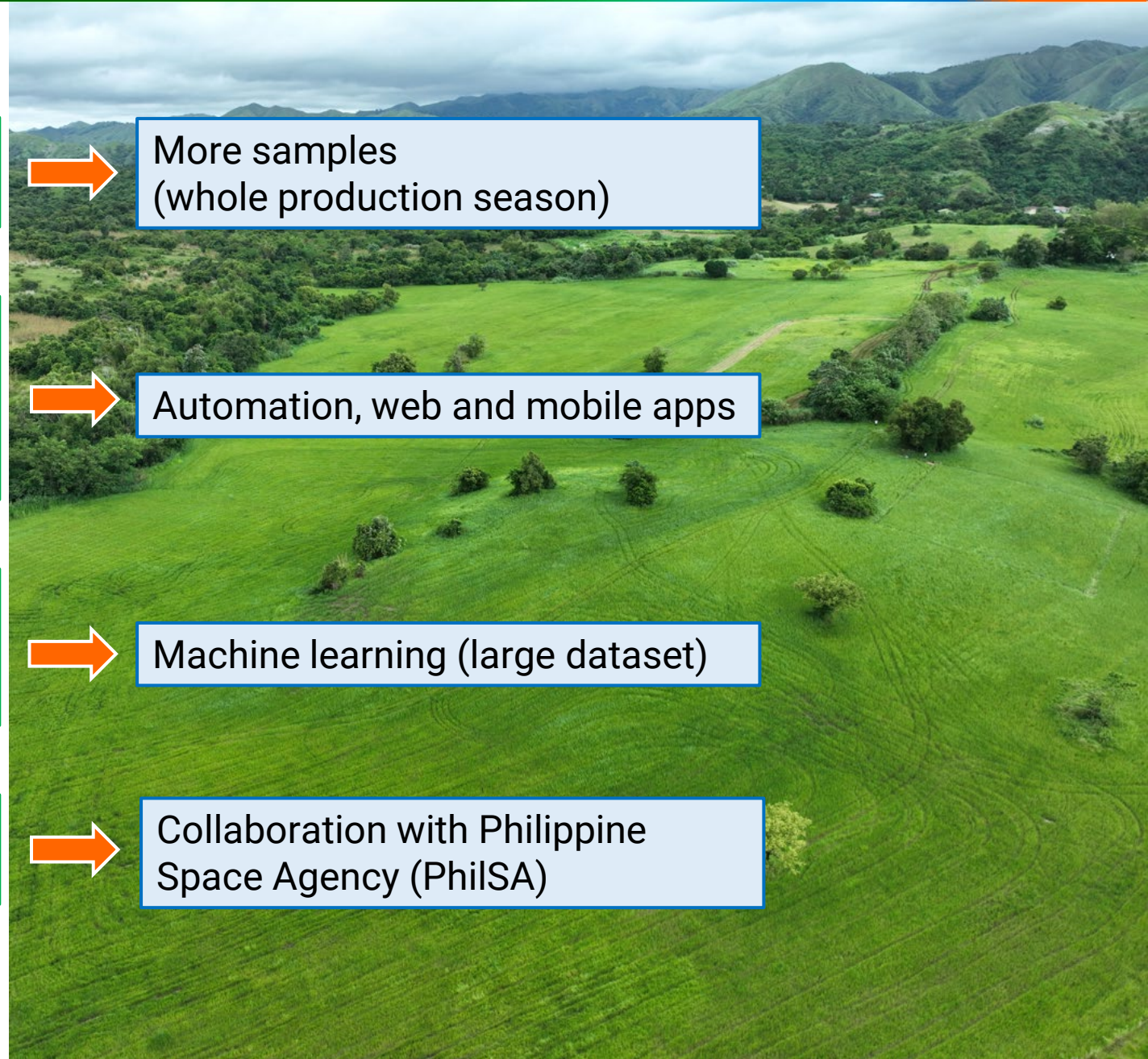
Automation, web and mobile apps



Machine learning (large dataset)



Collaboration with Philippine
Space Agency (PhilSA)



Thanks to my team in the Philippines...

44



感謝您的聆聽
Thank you!
Salamat po!



EAAP
European Federation
of Animal Science



**National Taiwan
University**



臺大國際學院
International College
National Taiwan University



Republic of the Philippines
Department of Agriculture
BUREAU OF ANIMAL INDUSTRY



Republic of the Philippines
DEPARTMENT OF SCIENCE AND TECHNOLOGY
Science Education Institute

