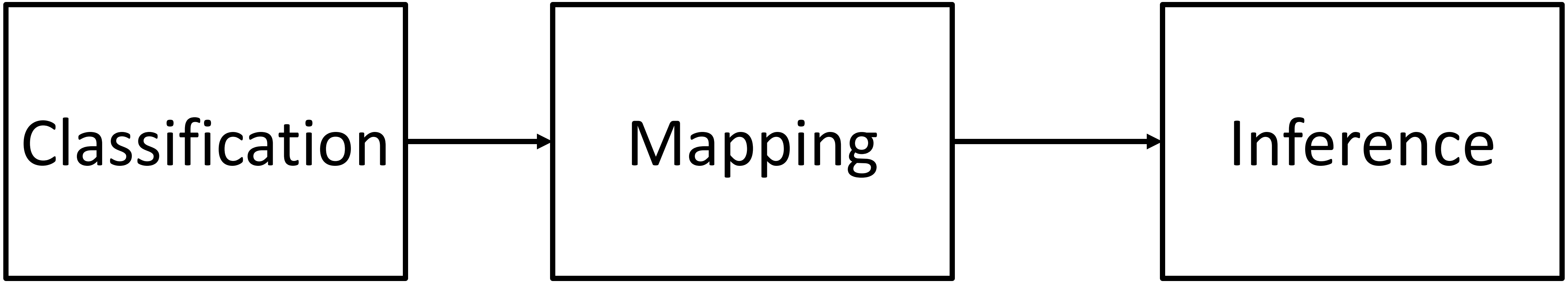
On the Inference from Satellite-based Classification Maps for Area Estimation

Sergii Skakun

University of Maryland, College Park, MD, USA

Land cover / land use mapping and area estimation



- Data
- Features
- Algorithms

- Mapping product

- Unbiased estimators with uncertainties

Problem statement: Area estimation

- Fundamental problem → **inference** from satellite-based maps

- Statistical framework: **sample-based** approach
  - Good practices [*Olofsson et al., 2014*] and
    CEOS Land Product Validation (LPV) protocols [*Tyukavina et al., 2025*]

  - **Stratified random sampling**
    - Maps used for strata

    - Map **inaccuracies will not impact the bias**: depends on the sampling design and estimator

    - Map **inaccuracies will impact the efficiency of stratification**

    - **Model-assisted** estimation to further improve the precision of the estimator
      [*Carfagna & Gallego, 2005*; *Gallego et al., 2010*; *McRoberts et al., 2024*]

Efficiency of stratification

- **Assumptions**
  - Sampling design: **stratified random sampling** with a pixel as sampling unit
  - Map's **classes** used as **strata**
  - Two maps: *A* and *B* used for stratification when estimating an area proportion $\hat{y}$

- **Efficiency** of stratification [*Gallego, 2007*]

$$\eta_{A,B} = \frac{var(\hat{y}_B) \times n_B}{var(\hat{y}_A) \times n_A}$$

n_B and n_A are sample size, and $var(\hat{y}_B)$ and $var(\hat{y}_A)$ are estimated variances of $\hat{y}$

- **Question**:
  To what extent do gains in map quality (accuracy) improve stratification efficiency?

Remote Sensing of Environment 326 (2025) 114805

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

The impact of map accuracy on area estimation with remotely sensed data within the stratified random sampling design

Sergii Skakun [*]

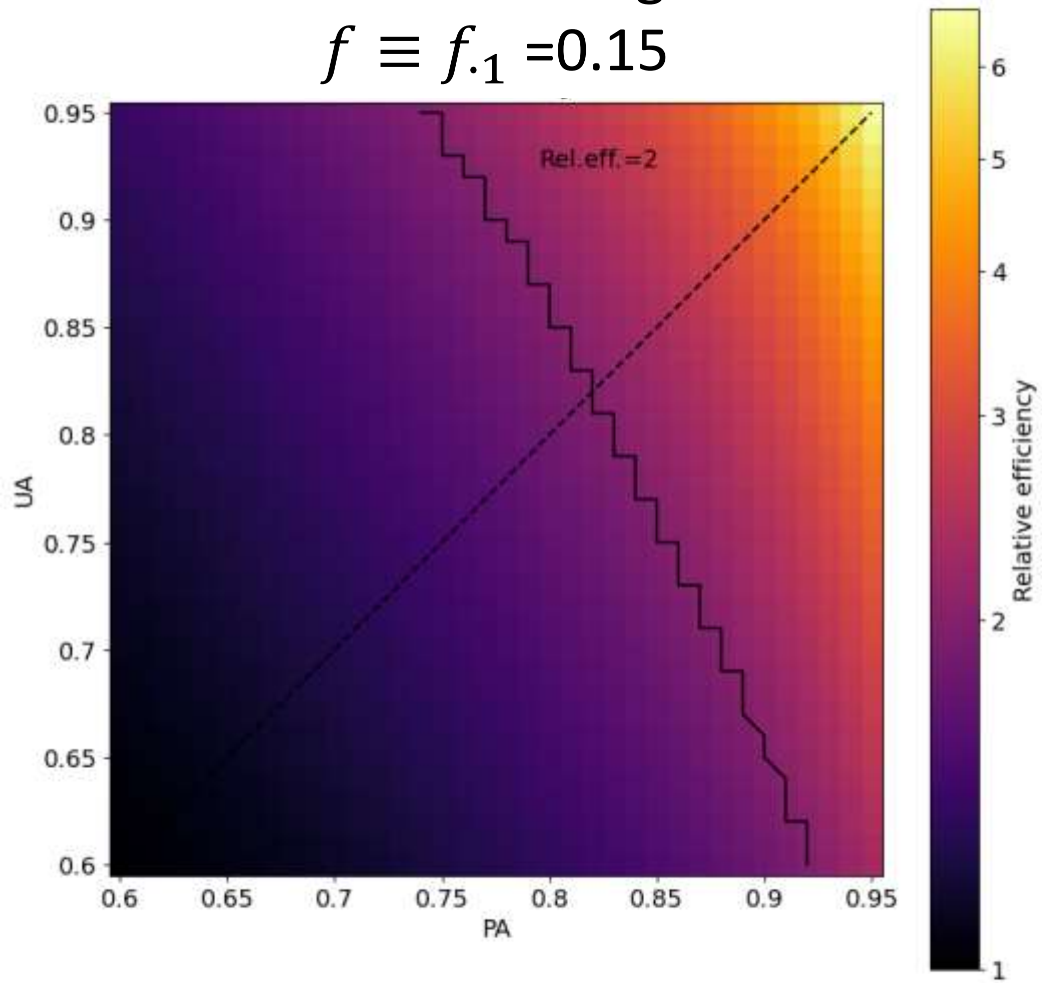
Department of Geographical Sciences, College of Information, University of Maryland, College Park, MD 20742, USA

https://doi.org/10.1016/j.rse.2025.114805

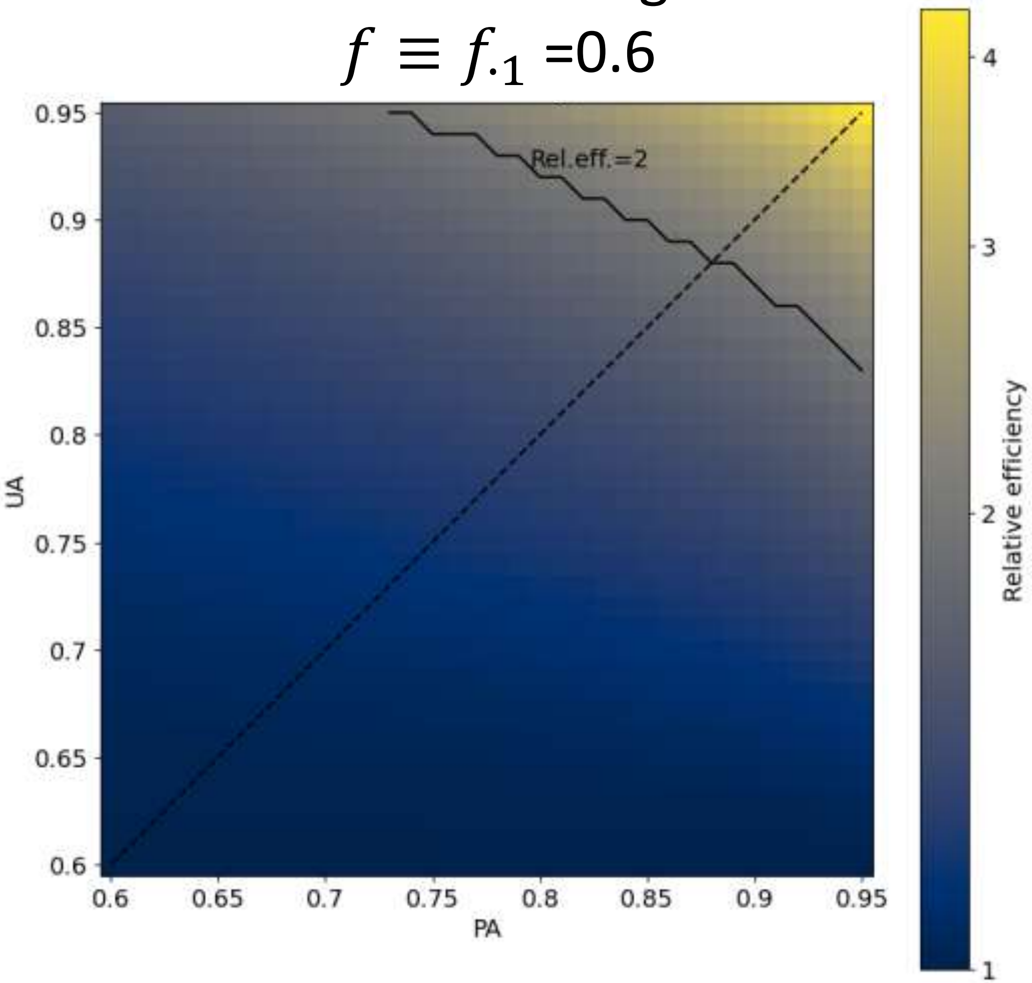
Google Colab with code is available

Relative efficiency of stratification

True area of the target class
$f \equiv f_{\cdot 1}$ =0.15

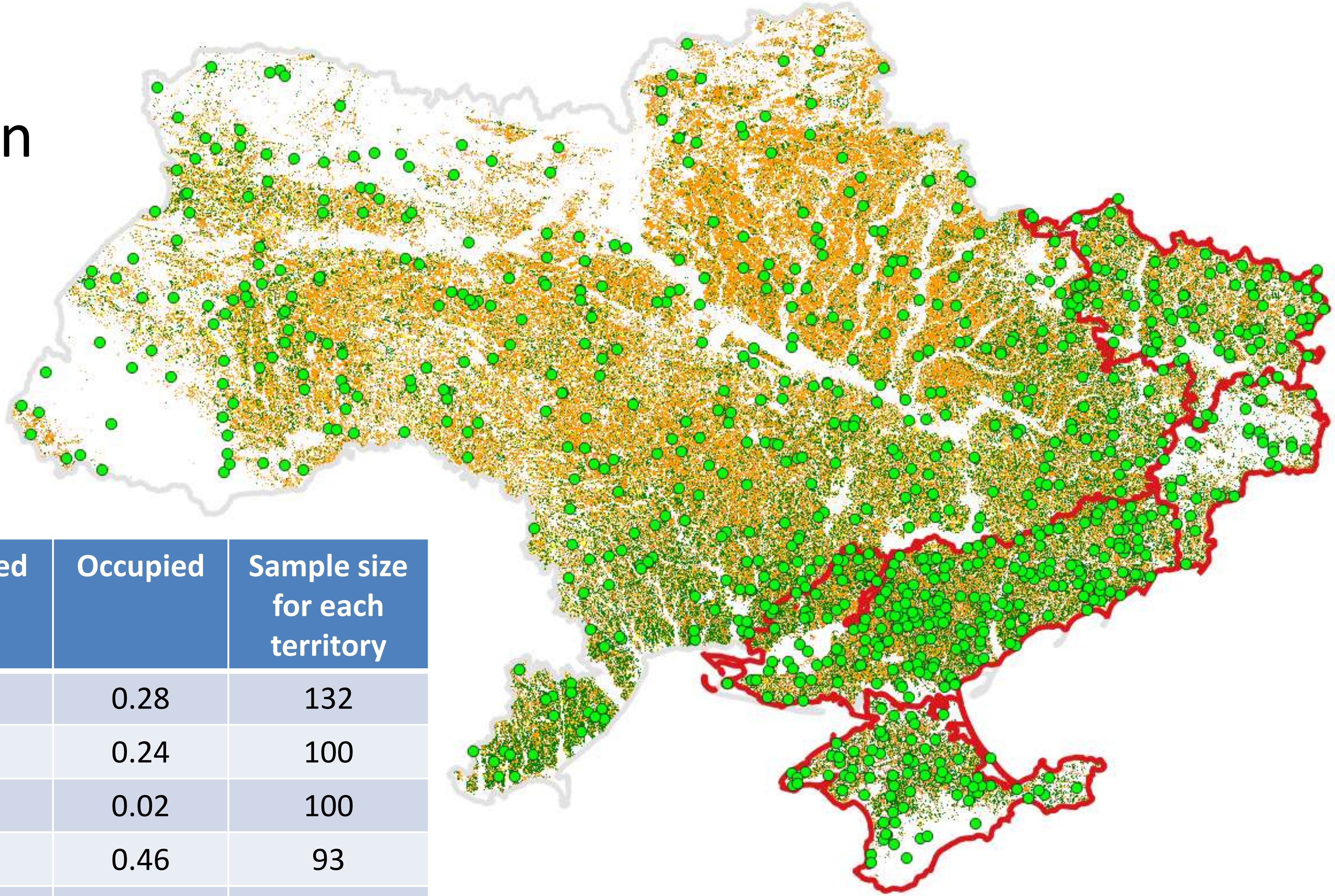


True area of the target class
$f \equiv f_{\cdot 1}$ =0.6



Relative efficiency $\eta_{A,B}$ calculated for each pair (v_1, u_1) (map *A*) in relation to the map at $(v_1, u_1) = (0.6, 0.6)$ (map *B*): $f = 0.15$ (left) and $f = 0.60$ (right). In each figure the solid line shows multiple solutions in the PA/UA space when a relative efficiency of 2 can be reached.

Winter cereal area estimation in 2022

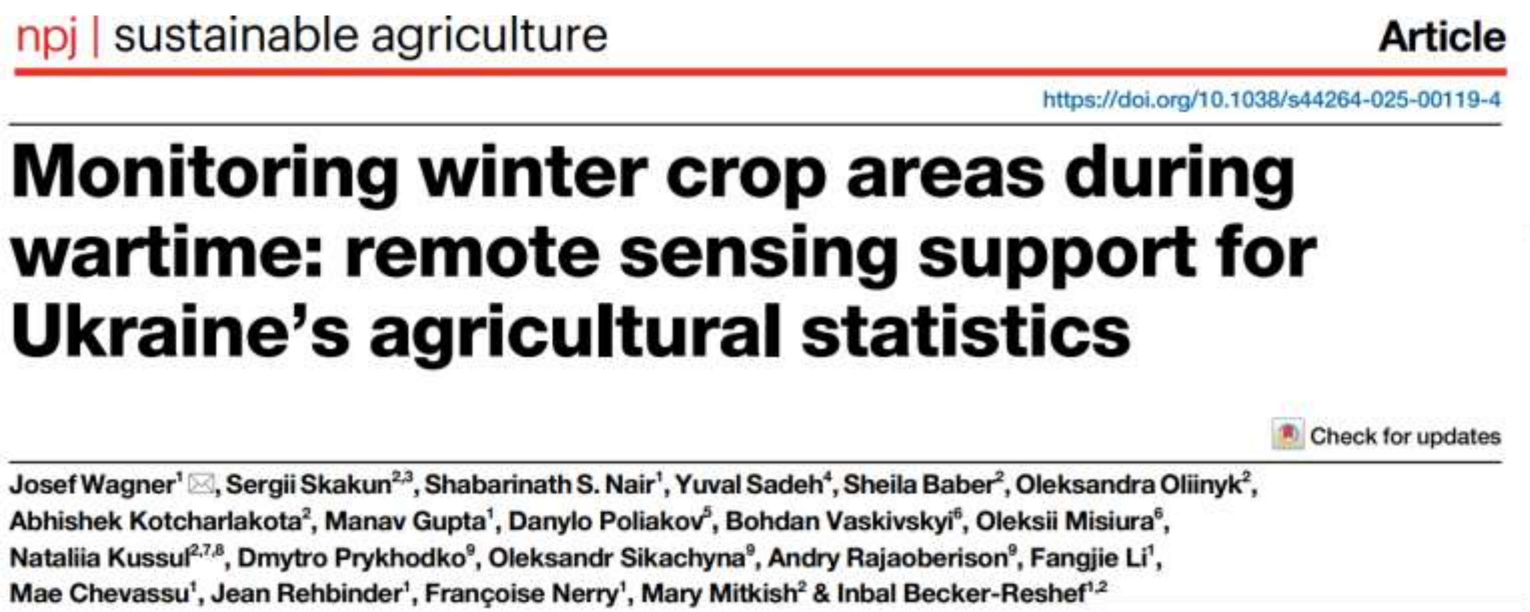


Stratum	Unoccupied	Occupied	Sample size for each territory
Summer crops	0.31	0.28	132
Winter cereal	0.13	0.24	100
Rapeseed	0.02	0.02	100
Non-cropland	0.53	0.46	93
Total	1	1	425

Final areas: temporary occupied territories (as of July 2022)

	Summer	Winter cereals	Winter rapeseed	Non-cropland
Estimated area [ha]	4,185,540	2,400,960	173,567	5,676,877
95% CI of area [ha]	±499,631	±367,998	±13,217	±497,106

CI = confidence interval at 95%


npj | sustainable agriculture
Article
https://doi.org/10.1038/s44264-025-00119-4

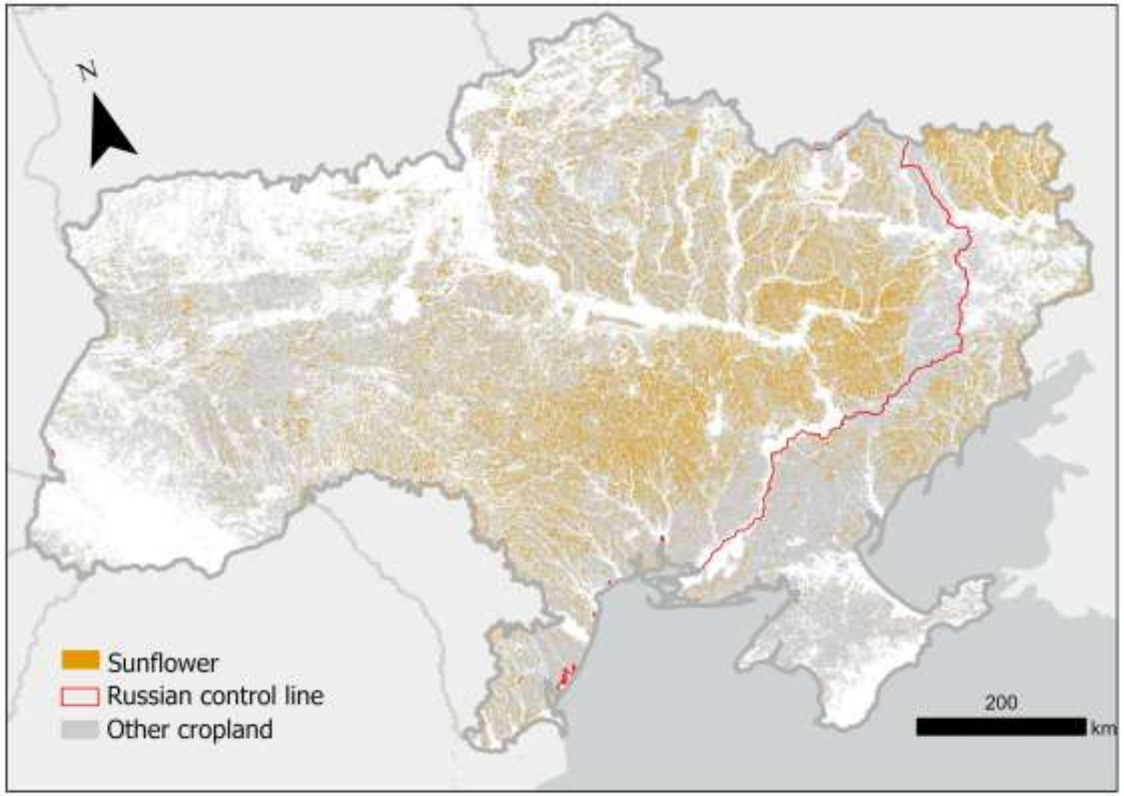
Monitoring winter crop areas during wartime: remote sensing support for Ukraine's agricultural statistics

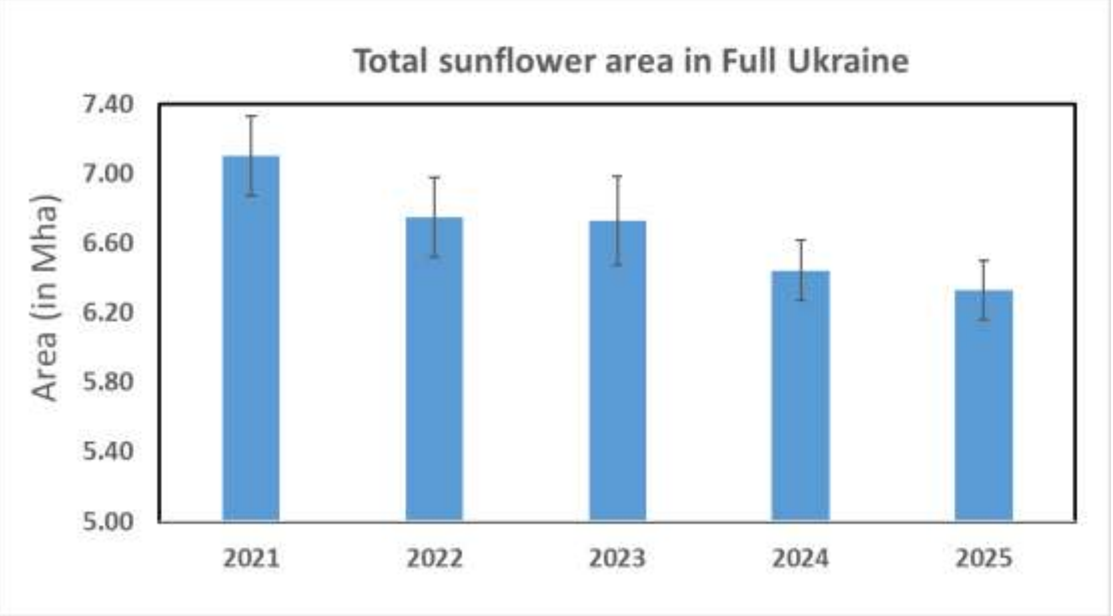
Check for updates

Josef Wagner[1], Sergii Skakun[2,3], Shabarinath S. Nair[1], Yuval Sadeh[4], Sheila Baber[2], Oleksandra Oliinyk[2], Abhishek Kotcharlakota[2], Manav Gupta[1], Danylo Poliakov[5], Bohdan Vaskivskyi[6], Oleksii Misiura[6], Nataliia Kussul[2,7,8], Dmytro Prykhodko[9], Oleksandr Sikachyna[9], Andry Rajaoberison[9], Fangjie Li[1], Mae Chevassu[1], Jean Rehbinder[1], Françoise Nerry[1], Mary Mitkish[2] & Inbal Becker-Reshef[1,2]

Sunflower area in Ukraine (2021 to 2025)

- Total sunflower area in Ukraine has been continuously decreasing since 2021
- There is a negligible decrease in sunflower area from 6.44 Mha in 2024 to 6.33 Mha in 2025
- Sunflower area decreased in both free and occupied regions (major decrease in Kherson and Mykolaiv regions)





Science of Remote Sensing
Volume 10, December 2024, 100139

Estimation of sunflower planted areas in Ukraine during full-scale Russian invasion: Insights from Sentinel-1 SAR data

Abdul Qadir [a], Sergii Skakun [a, b], Inbal Becker-Reshef [a], Nataliia Kussul [a, c, d], Andrii Shelestov [c, d]

Conclusions

- When **maps** are used for **area estimation** using a **sample-based** approach within the **stratified random sampling design** with error-free reference data:
  - **Map quality** impacts stratification **efficiency**
    - A more accurate map
      → **smaller sample size** to reach the desired precision of the estimate
      → **improved precision** when the sample size is fixed
    - Analytical expressions for K=2
  - A criterion (**relative efficiency**) ← **implications** of **accuracy** increase
    - Depends on **class-specific PA** and **UA**, and **target class area**
    - The impact is not linear and contributions of **PA**, **UA**, and f are not equal
    - OA, F-score are not adequate metrics

Back up slides

Error matrix

- Population matrix (in terms of **area proportions**)

		Reference				
		Class 1	Class 2	…	Class K	Total
Map	Class 1	f_{11}	f_{12}	…	f_{1K}	$f_{1\cdot}$
	Class 2	f_{21}	f_{22}	…	f_{2K}	$f_{2\cdot}$
	…	…	…	…	…	…
	Class k	f_{K1}	f_{K2}	…	f_{KK}	$f_{K\cdot}$
	Total	$f_{\cdot 1}$	$f_{\cdot 2}$	…	$f_{\cdot K}$	1

$\boldsymbol{f_{k\cdot}}$ is the **mapped** area proportion

$\boldsymbol{f_{\cdot k}}$ is the **reference** area proportion

Producer's accuracy (PA): $v_k = {f_{kk}}/{f_{\cdot k}}$

User's accuracy (UA): $u_k = {f_{kk}}/{f_{k\cdot}}$

- A **bias of the pixel counting estimator**

Through PA/UA $\quad b_k = f_{k\cdot} - f_{\cdot k} = f_{\cdot k}\underbrace{\left({v_k}/{u_k} - 1\right)}_{\text{Relative bias}}$

Relative efficiency

- Elements of the confusion matrix estimated from the **sample**

- Sample size [*Olofsson et al., 2014*; *Cochran, 1977, Eq. 5.25*]

$$n = \frac{\left(\sum_{k=1}^{K} w_k s_k\right)^2}{V_{target} + {}^{1}/_{N} \sum_{k=1}^{K} w_k s_k^2} \approx \frac{\left(\sum_{k=1}^{K} w_k s_k\right)^2}{V_{target}}$$

where V_{target} is the target variance, $w_k \equiv f_{k\cdot}$ is the stratum weight, N is population size, and s_k is the standard deviation within stratum k

for overall accuracy $s_k(\hat{f}_{kk}) = \sqrt{u_k(1 - u_k)}$

for area estimation $s_i(\hat{f}_{\cdot k}) = \sqrt{\frac{n_{ik}}{n_{i\cdot}}\left(1 - \frac{n_{ik}}{n_{i\cdot}}\right)}$

Relative efficiency: binary case K=2

- Relative efficiency for *K*=2 is the function of target area $\boldsymbol{f}$, $\boldsymbol{v_1}$=***PA***, $\boldsymbol{u_1}$=**UA**:
  - Need **3 variables** to fully describe 2x2 confusion matrix

$$\eta_{A,B} = \frac{var(\hat{y}_B)\times n_B}{var(\hat{y}_A)\times n_A} =$$

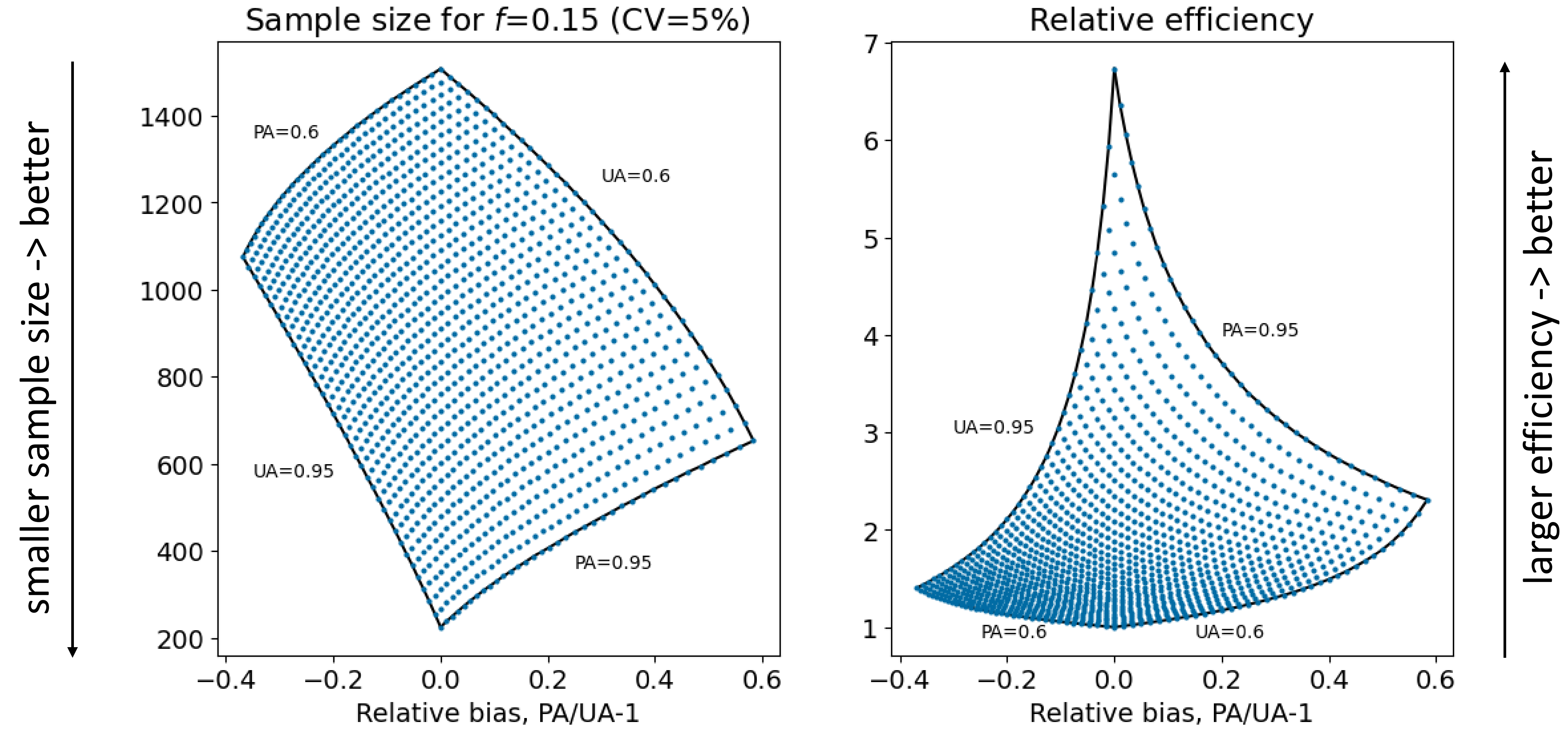
$$= \frac{\left(w_1^B\sqrt{u_1^B(1-u_1^B)}+w_2^B\sqrt{u_2^B(1-u_2^B)}\right)^2}{\left(w_1^A\sqrt{u_1^A(1-u_1^A)}+w_2^A\sqrt{u_2^A(1-u_2^A)}\right)^2} = \frac{\left(f v_1^B\sqrt{\frac{1}{u_1^B}-1}+(1-f)v_2^B\sqrt{\frac{1}{u_2^B}-1}\right)^2}{\left(f v_1^A\sqrt{\frac{1}{u_1^A}-1}+(1-f)v_2^A\sqrt{\frac{1}{u_2^A}-1}\right)^2}$$

Relative efficiency of stratification

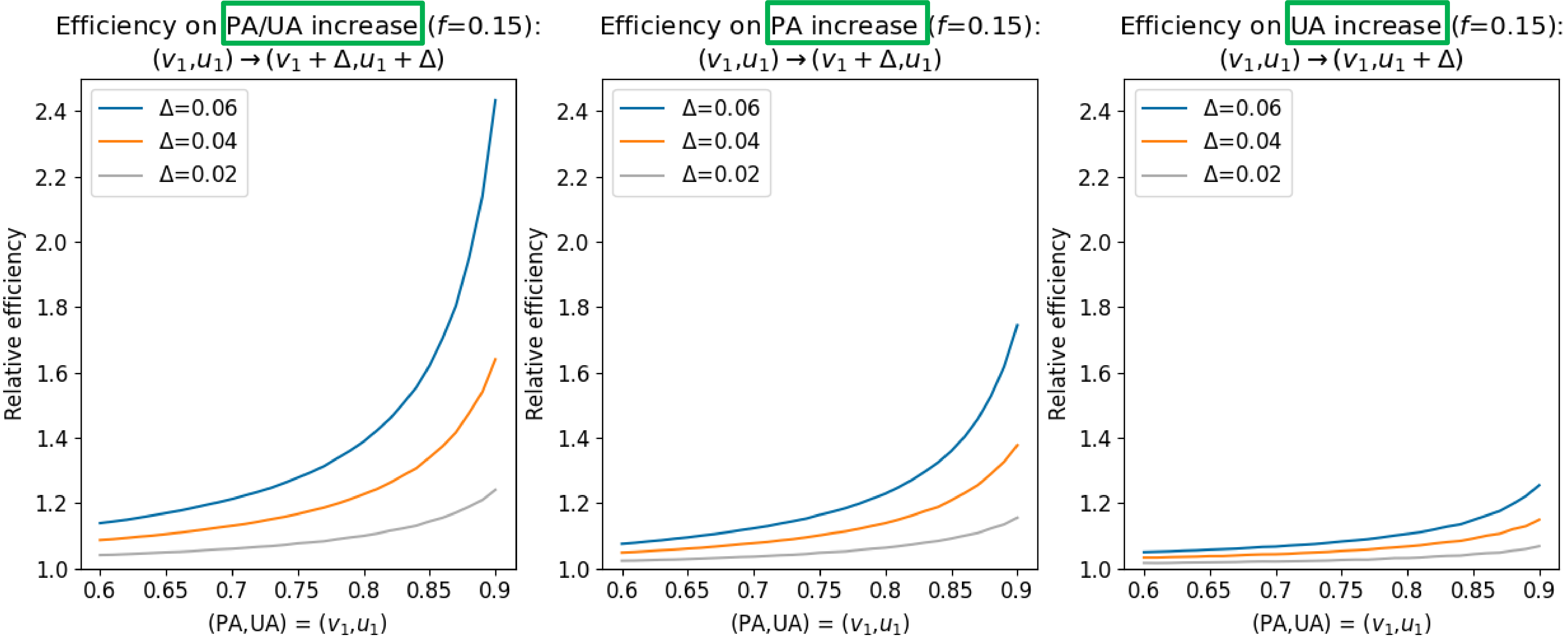
- Two classes K=2, true area of the target class $f \equiv f_{.1}$ =0.15 and target CV=5%.

	A PA=0.9 UA=0.7	B PA=0.7 UA=0.9	C PA=0.8 UA=0.8	D PA=0.6 UA=0.6
	$f = 0.15$			
Overall	0.927	0.943	0.940	0.880
F-score=$\frac{2v_1u_1}{v_1+u_1}$	0.788	0.788	0.800	0.600
Sample size n	693	934	836	1508
Rel. eff. η	**2.18**	1.61	1.80	1.00
Rel. eff. (to SRS)	**3.27**	2.43	2.71	1.50

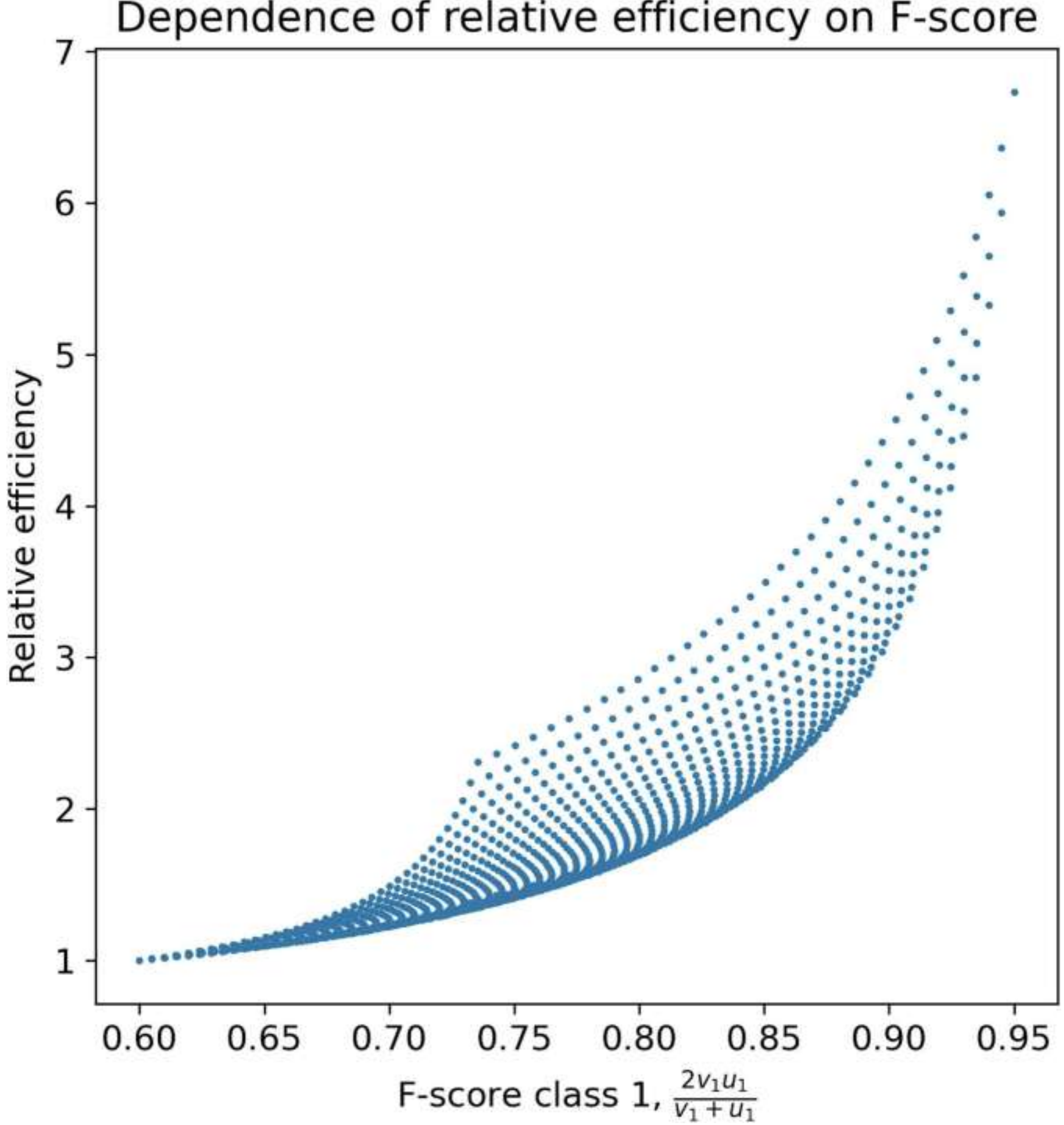
Simple random sampling (SRS), sample size n=2,267



Dependence of the total sample size and relative efficiency (in relation to $(v_1, u_1) = (0.6, 0.6)$) on the relative bias.
Case $f = 0.15$.



Dependence of improvements of relative efficiency when PA and UA increase by $\Delta = +0.02$, $+0.04$ and $+0.06$. Efficiency increase when both PA and UA increase by Δ is shown in left figure; PA increase only is center figure; and UA increase only is right figure.



Dependence of relative efficiency (in relation to =(0.6,0.6)) on the F-score of class 1

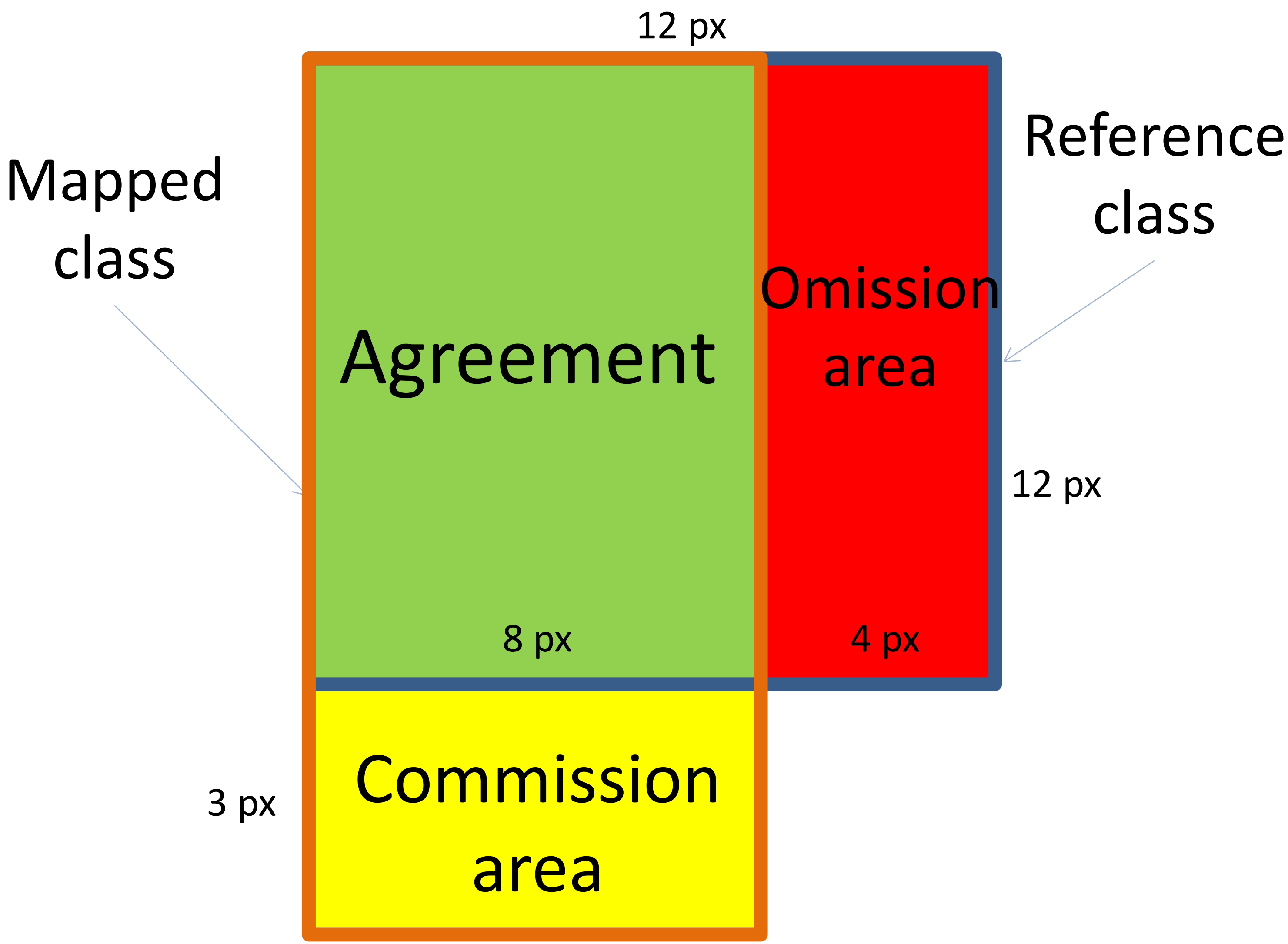
Problem statement: Area estimation

- Areas can be estimated w/o satellite-based maps!

- Maps
  - **Wall-to-wall** coverage → known satellite-based information on **population** (subject to spatial resolution)
  - Omission and commission **errors** → information on the population might be **inaccurate**
  - "Naïve" use: a ***pixel counting estimator*** is biased [*Gallego, 2004*; *Gallego et al., 2010*; *McRoberts, 2011*]

Land cover / land use mapping and area estimation

- Pixel counting is a biased estimator



Reference area:
12 x 12 = 144 px

Mapped area:
8 x 15 = 120 px (bias -17%)

PA = 8 x 12 / (12 x 12) = 66.7%

UA = 8 x 12 / (8 x 15) = 80%

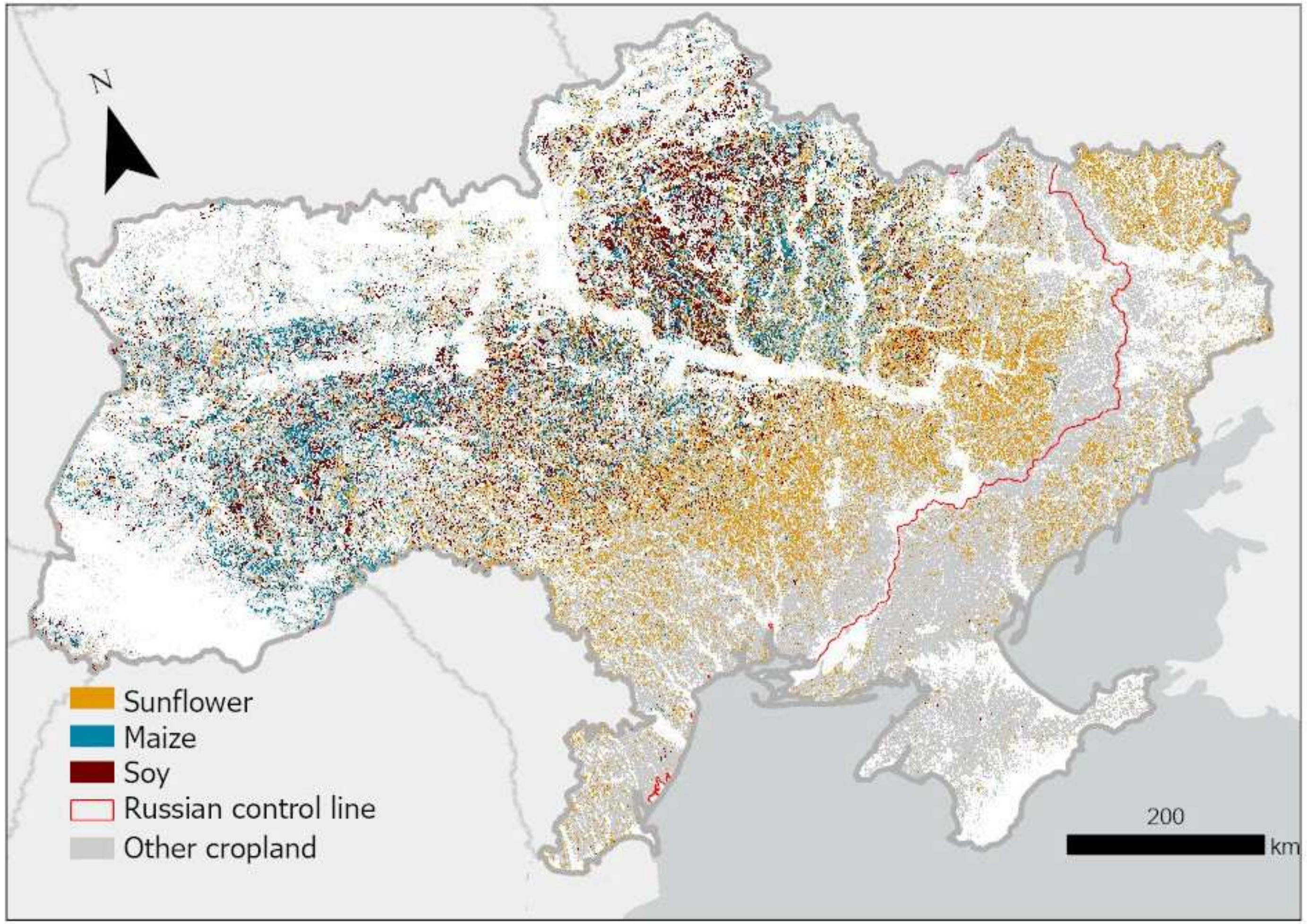
Rel. bias = PA/UA - 1

Summer crops area estimation in Ukraine

Sunflower 2025 (in Ha)	Estimated	SE	CI	CV, %
Full Ukraine	6,330,106	168,637	330,529	2.7
Free territories	5,459,592	135,484	265,549	2.5
Occupied territories	870,514	100,412	196,808	11.5

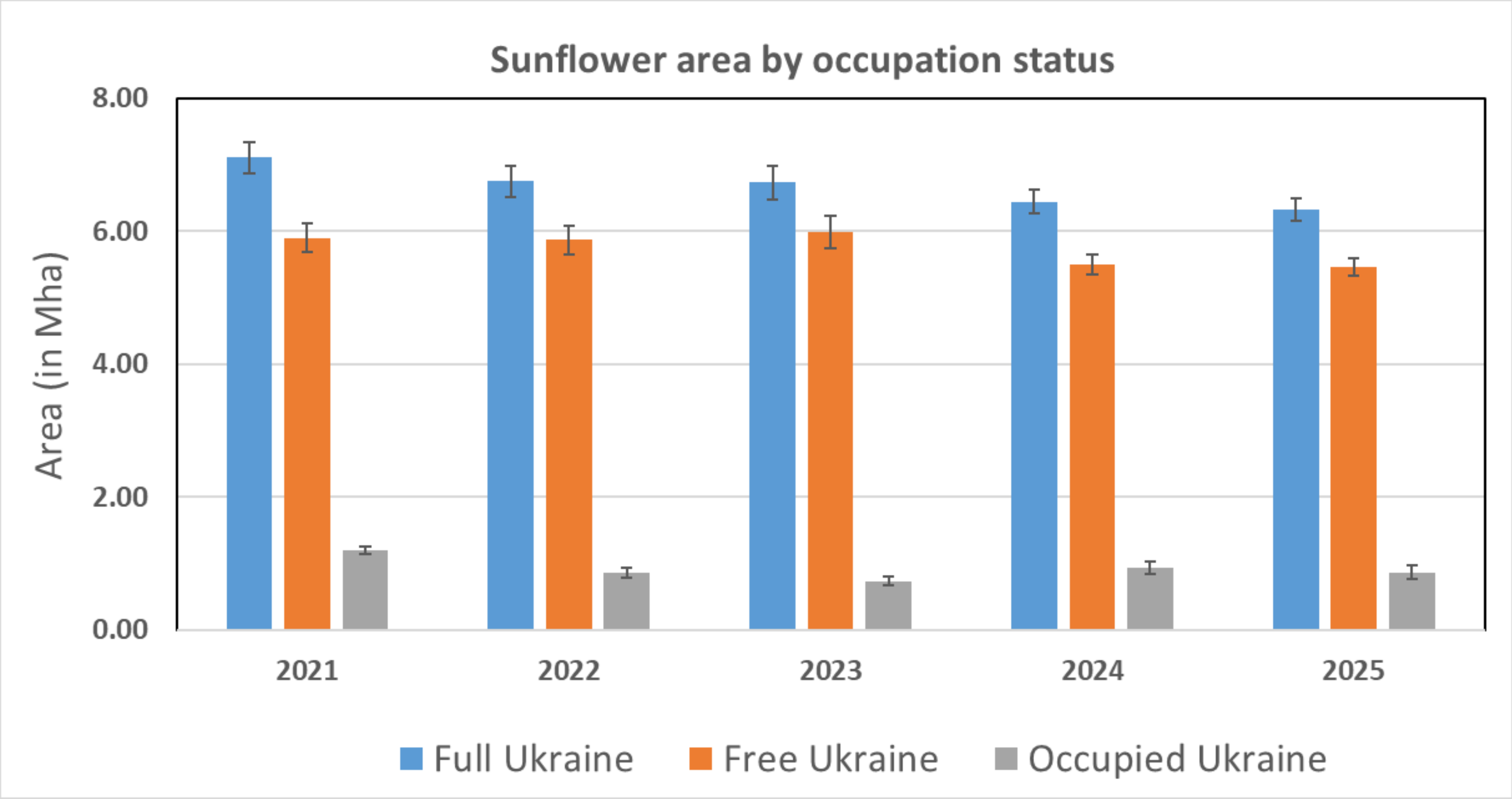
Maize 2025 (in Ha)	Estimated	SE	CI	CV (%)
Full Ukraine	4,314,107	210,942	413,446	4.89
Free territories	4,244,909	200,981	393,922	4.73
Occupied territories	69,198	64,057	125,552	92.57

Soy 2025 (in Ha)	Estimated	SE	CI	CV (%)
Full Ukraine	3,157,966	240,891	472,145	7.63
Free territories	3,152,080	240,886	472,137	7.64
Occupied territories	5,886	1,472	2,884	25.00



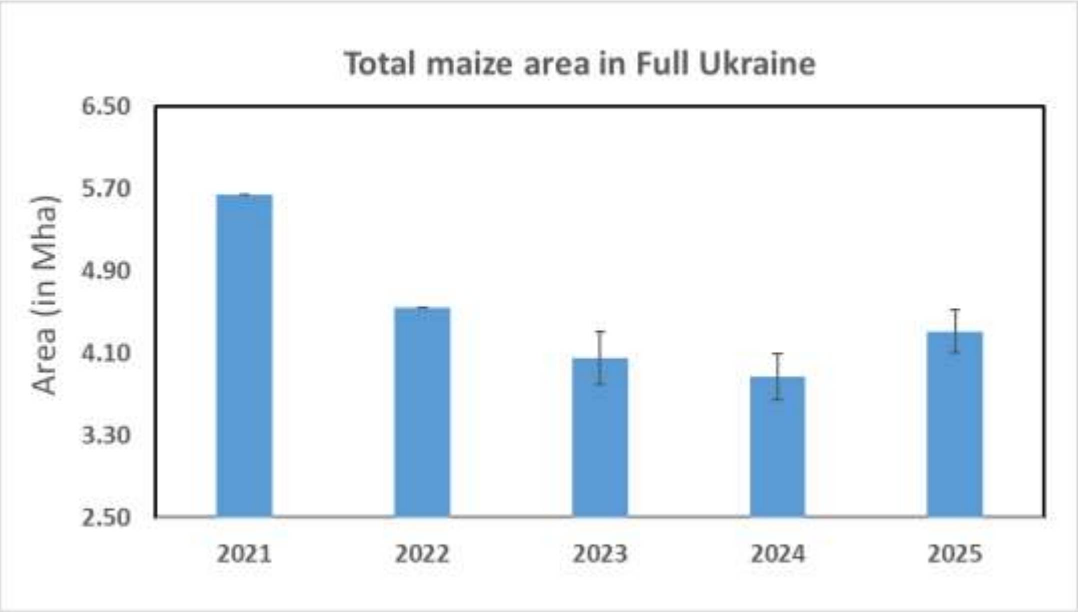
Sunflower area by occupation (2021 to 2025)

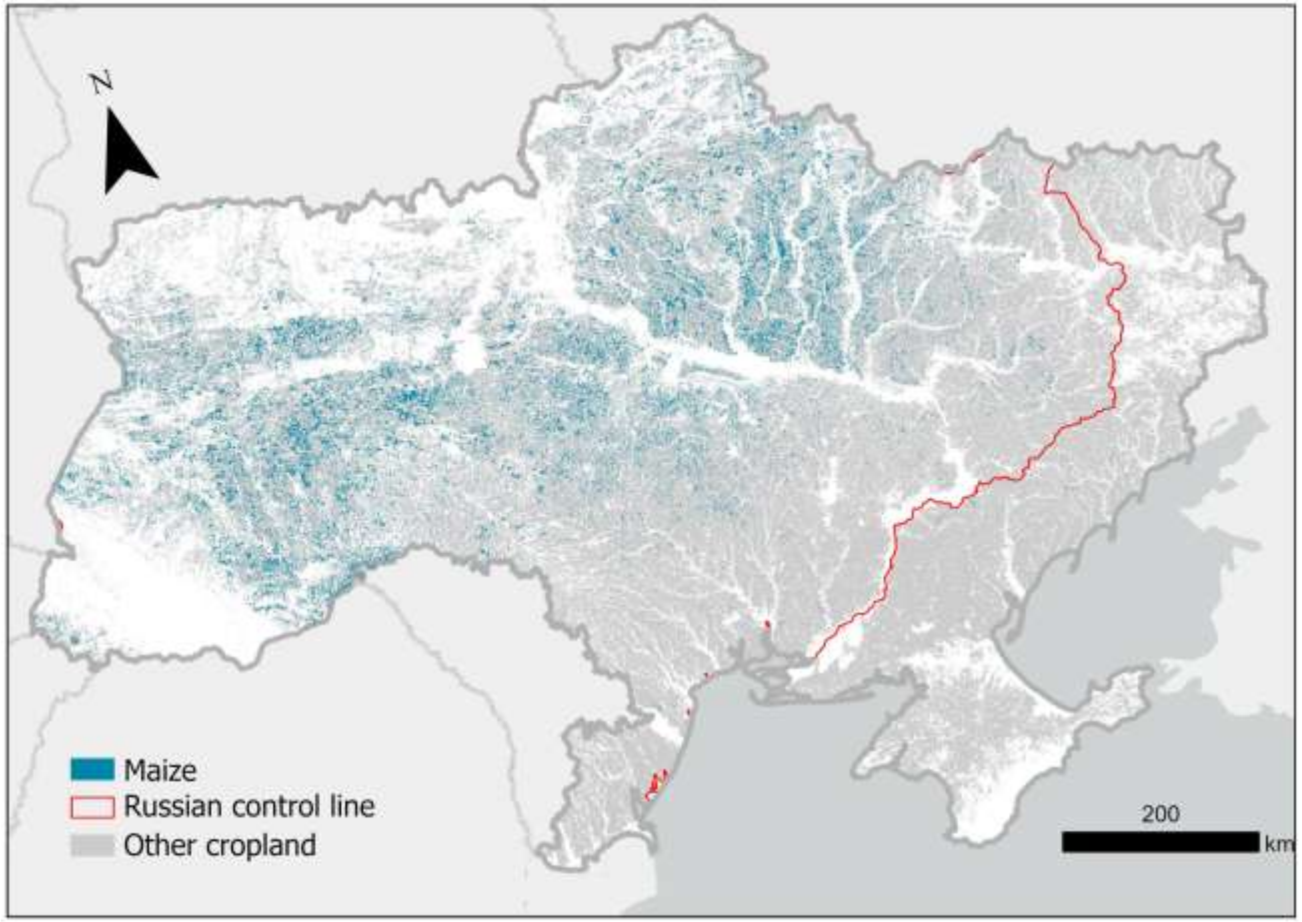
- Total sunflower area in the occupied region has been nearly constant in the last four years.
- Decreasing trend in the free Ukraine region from 2021 onwards.



Maize area in Ukraine (2021 to 2025)

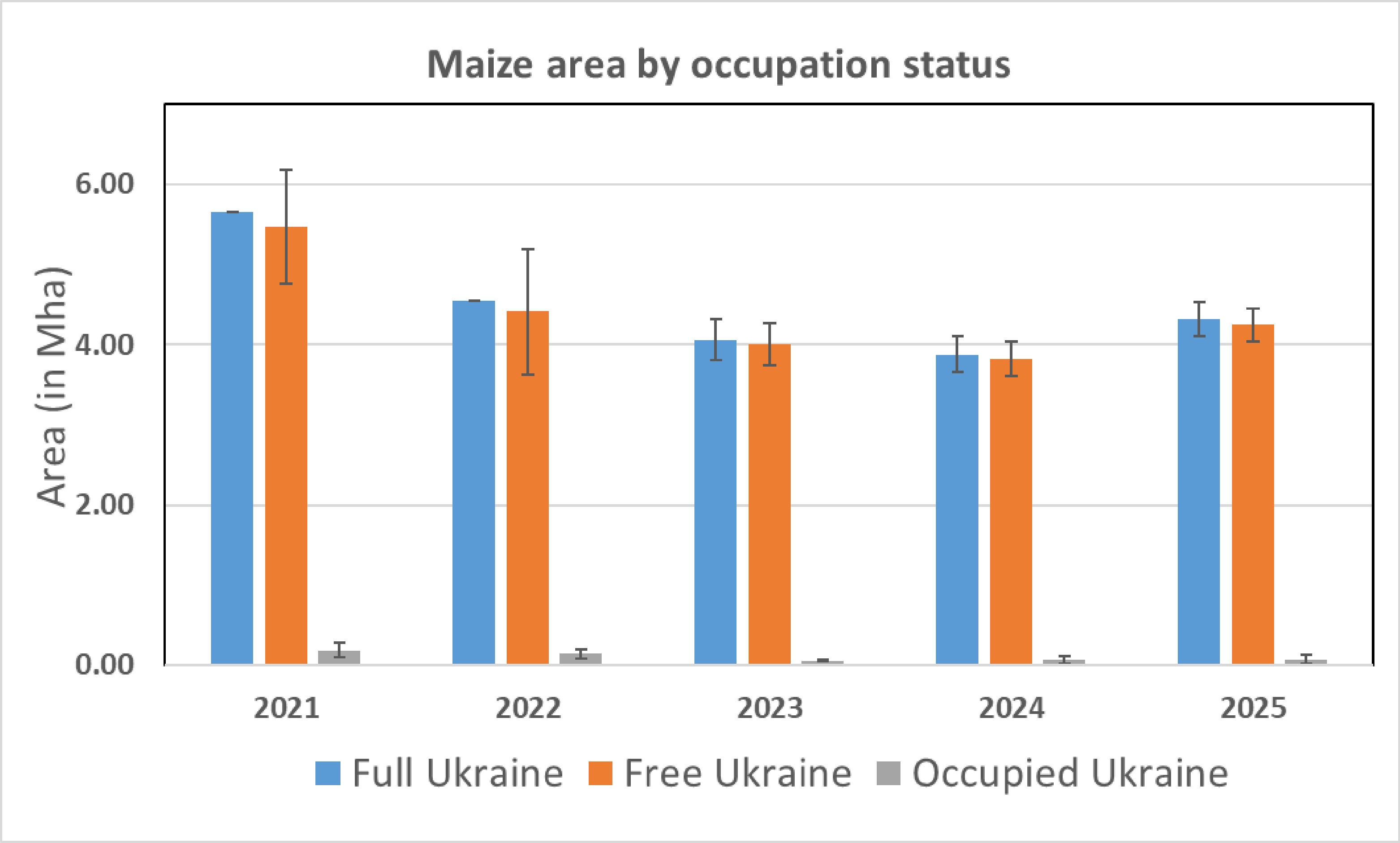
- Maize area increased in 2025 compared to 2024.
- There is approx. 11% increase in maize area compared to last year.
- Highest maize area in last 3 years.





Maize area by occupation (2021 to 2025)

- Maize area is nearly negligible in Occupied Ukraine over the last three years.
- In last three years average maize area in occupied region decreased to 0.6 Mha from 0.15 Mha.



Soy area in Ukraine (2025)

- Soy area estimated as 3.16 Mha. USDA is projecting 3Mha this year compared to last year's 2.9 Mha.
- Soy area harvested this year will be the highest in Ukraine's history.