DIGITALES ARCHIV

ZBW – Leibniz-Informationszentrum Wirtschaft
ZBW – Leibniz Information Centre for Economics

Book

Technological innovation for agricultural statistics : key indicators for Asia and the Pacific 2018 : special supplement

Provided in Cooperation with:
Asian Development Bank (ADB), Manila

Reference: (2018). Technological innovation for agricultural statistics : key indicators for Asia and the Pacific 2018 : special supplement. Mandaluyong City, Metro Manila, Philippines : ADB. doi:10.22617/FLS189528-3.

This Version is available at:
http://hdl.handle.net/11159/2811

Kontakt/Contact
ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: *rights[at]zbw.eu*
https://www.zbw.eu/econis-archiv/

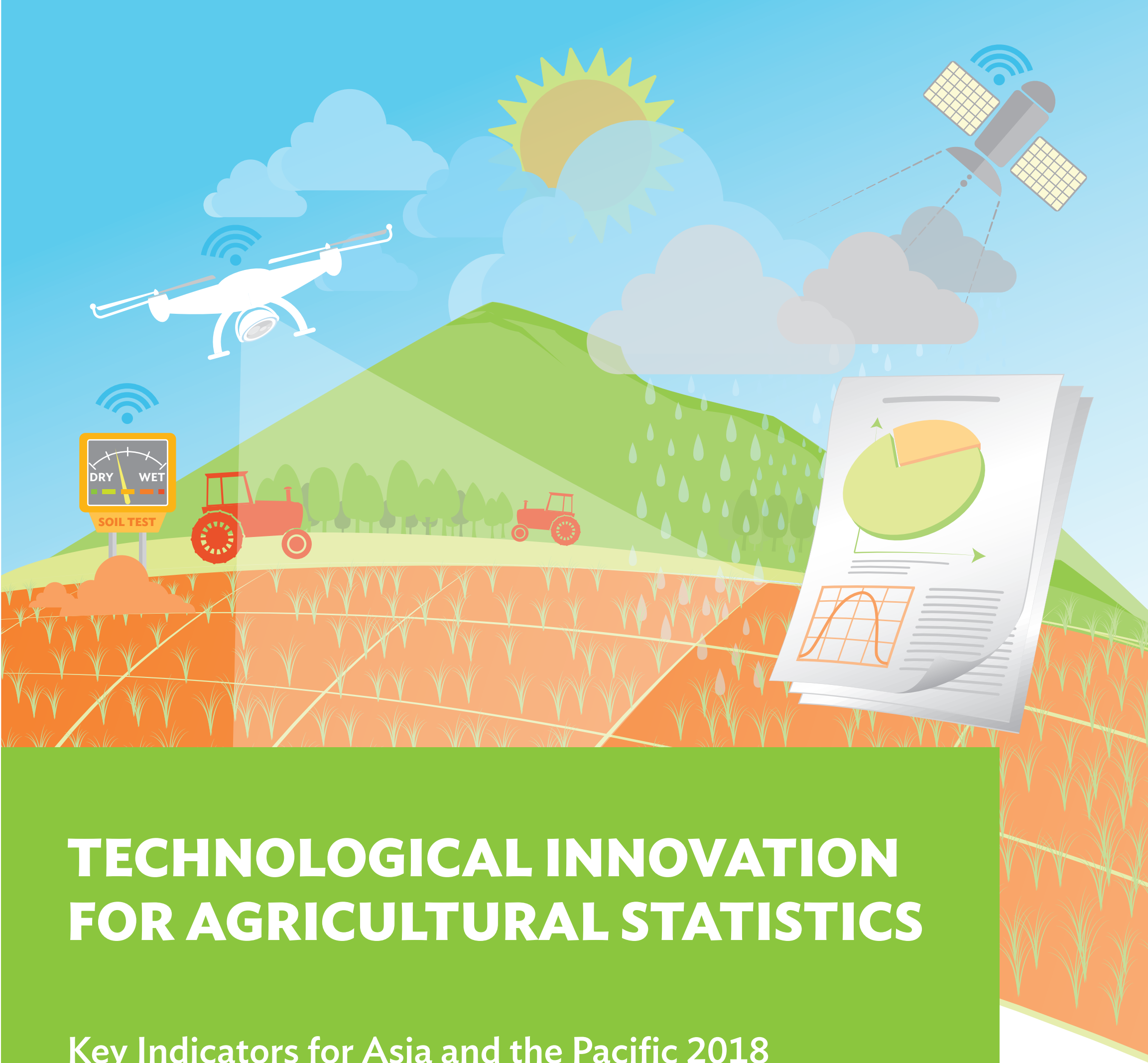
Standard-Nutzungsbedingungen:
Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

https://zbw.eu/econis-archiv/termsofuse

Terms of use:
This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der Leibniz-Gemeinschaft



TECHNOLOGICAL INNOVATION FOR AGRICULTURAL STATISTICS

Key Indicators for Asia and the Pacific 2018 Special Supplement

SEPTEMBER 2018

ASIAN DEVELOPMENT BANK

ADB

TECHNOLOGICAL INNOVATION FOR AGRICULTURAL STATISTICS

Key Indicators for Asia and the Pacific 2018
Special Supplement

ASIAN DEVELOPMENT BANK

ADB

Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO)

© 2018 Asian Development Bank
6 ADB Avenue, Mandaluyong City, 1550 Metro Manila, Philippines
Tel +63 2 632 4444; Fax +63 2 636 2444
www.adb.org

Some rights reserved. Published in 2018.

ISBN 978-92-9261-312-9 (print), 978-92-9261-313-6 (electronic)
Publication Stock No. FLS189528-3
DOI: http://dx.doi.org/10.22617/FLS189528-3

The views expressed in this publication are those of the authors and do not necessarily reflect the views and policies of the Asian Development Bank (ADB) or its Board of Governors or the governments they represent.

ADB does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use. The mention of specific companies or products of manufacturers does not imply that they are endorsed or recommended by ADB in preference to others of a similar nature that are not mentioned.

By making any designation of or reference to a particular territory or geographic area, or by using the term "country" in this document, ADB does not intend to make any judgments as to the legal or other status of any territory or area.

This work is available under the Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO) https://creativecommons.org/licenses/by/3.0/igo/. By using the content of this publication, you agree to be bound by the terms of this license. For attribution, translations, adaptations, and permissions, please read the provisions and terms of use at https://www.adb.org/terms-use #openaccess

This CC license does not apply to non-ADB copyright materials in this publication. If the material is attributed to another source, please contact the copyright owner or publisher of that source for permission to reproduce it. ADB cannot be held liable for any claims that arise as a result of your use of the material.

Please contact pubsmarketing@adb.org if you have questions or comments with respect to content, or if you wish to obtain copyright permission for your intended use that does not fall within these terms, or for permission to use the ADB logo.

Notes:
In this publication, "$" refers to US dollars.
Corrigenda to ADB publications may be found at http://www.adb.org/publications/corrigenda

Foreword

The Asia and Pacific region has experienced exceptional economic growth over the past two decades, accompanied by a dramatic reduction in extreme poverty. However, more moderate forms of poverty remain widespread, especially in rural areas and among households that rely on agriculture. Given the strong link between agricultural development and poverty alleviation, it is crucial that policies for the agriculture sector be evidence-based, timely, and reliable. However, most countries in the region continue to depend on traditional methods for collecting agricultural statistics that are susceptible to significant measurement errors due to the subjective nature of data collection. Remote sensing technology offers more objective methods to enhance availability and quality of agricultural statistics.

In 2010, the United Nations Statistical Commission initiated the Global Strategy to Improve Agricultural and Rural Statistics to enhance and ensure the sustainability of agricultural and rural statistics. The Asian Development Bank (ADB) joined the initiative as an implementing partner to develop technical assistance projects that pilot technological innovations to enhance the quality, quantity, and timeliness of agricultural statistics.

This report is a special supplement to the *Key Indicators for Asia and the Pacific 2018*. It presents a summary of the methodological research activities undertaken by ADB in collaboration with the national statistics offices and ministries of agriculture of three countries: the Lao People's Democratic Republic, Thailand, and Viet Nam. Specifically, the report explores the usefulness of remote sensing for land area measurement, yield estimation, and the development of a sampling frame. It also sheds light on other innovations such as drones, computer-assisted personal interviewing, and artificial intelligence, all of which are expected to revolutionize field data collection methods, improve the quantity and quality of agricultural data, and bolster evidence-based policymaking for agricultural development.

This report was produced by ADB's Development Economics and Indicators Division, under the overall guidance of Rana Hasan. The publication was prepared by Lakshman Nagraj Rao and Jude David Roque, with technical support from Anna Christine Durante, Pamela Lapitan, and

Dave Pipon. Rea Jean Tabaco and Lea Rotairo provided excellent research assistance. Kaushal Joshi reviewed the manuscript and provided valuable feedback. Copy editing was performed by Paul Dent, while the publication's cover design, layout, page design, and typesetting was carried out by Rhommell Rico. We are extremely grateful to the participating national statistics offices and ministries of agriculture for their active involvement in every stage of the project. Their contribution in terms of data collection, cleaning, and processing, alongside expert inputs, local knowledge, and practical advice was vital for the successful completion of the project.

We hope that this publication will be instrumental in promoting the role of technology in producing high-quality, timely, and cost-effective agricultural statistics to support policymaking for the agriculture sector.

Yasuyuki Sawada
Chief Economist and Director General
Economic Research and Regional Cooperation Department
Asian Development Bank

Contents

Figures

Boxes

Table

Abbreviations

ADB	Asian Development Bank
CAPI	computer-assisted personal interviewing
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
GCVI	Green Chlorophyll Vegetation Index
GIS	geographic information system
GPS	global positioning system
GSARS	Global Strategy to Improve Agricultural and Rural Statistics
ha	hectare
IRRI	International Rice Research Institute
m	meter
MAF	Ministry of Agriculture and Forestry
MoA	ministry of agriculture
MODIS	Moderate Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index
NSDS	National Strategy for Development of Statistics
NSO	national statistics office
PPP	purchasing power parity
RSC	Regional Steering Committee
SDG	Sustainable Development Goal
UNESCAP	United Nations Economic and Social Commission for Asia and the Pacific

Highlights

- **The importance of agricultural development in achieving poverty reduction is undisputed.** Accordingly, agriculture has been given special attention in the Sustainable Development Goals, with Target 2.3 aiming to double the agricultural productivity and incomes of small-scale food producers.

- **Despite this, limited efforts have been made to improve the accuracy and timeliness of agricultural statistics.** Most countries across Asia and the Pacific still rely on administrative reporting systems or sample surveys as data collection methods, which may be prone to significant measurement errors.

- **Recognizing this challenge, the Asian Development Bank (ADB) joined the Global Strategy to Improve Agricultural and Rural Statistics (GSARS) as an implementing partner.** The ADB supported this initiative by developing technical assistance projects that piloted innovative technologies to improve the availability and quality of agricultural statistics.

- **Specifically, the ADB piloted the use of remote sensing technology as an alternative to existing methods for generating key paddy rice statistics.** These methods were explored in collaboration with the national statistics offices and ministries of agriculture of three countries: the Lao People's Democratic Republic, Thailand, and Viet Nam. In each country, one province was selected for testing these new methods.

- **As a first step, a systematic comparison of existing objective and subjective methods to estimate plot area, rice production, and yield was conducted.** Farmer self-reports, the predominant way of collecting information on plot area and production in the countries, was compared with area measurements using global positioning system (GPS) technology and production estimates obtained from crop-cutting derived yields and GPS-based area estimates. Similarly, paddy rice yield estimated from farmer self-reports was compared to crop-cutting derived yield estimates.

- **Significant differences in paddy rice statistics were observed between objective and subjective data collection methods.** The differences were found to be nonlinear across the land size distribution and nonuniform in the direction of reporting bias. This confirmed that existing subjective methods implemented by the countries may not be sufficient to generate accurate paddy rice statistics.

- **ADB conducted three methodological studies to explore the viability of using satellite data as an alternative to traditional methods for estimating paddy rice statistics.** More specifically, these studies explored the use of remote sensing for land area measurement, rice yield estimation, and the development of a sampling frame. The objective was to compare the precision and costs associated between the two methods to assess the viability of remote sensing.

- **In the first study, plot boundaries were traced on high-resolution Google Earth images to estimate area, which was compared with GPS derived plot area estimates.** Results show that the plot areas derived from Google Earth images are statistically similar to plot area estimates from GPS, and are achievable at 38% lower per-unit costs. The lower costs are achieved because enumerators do not have to walk around the perimeter of the plot, but instead accomplish the task through tracing the plot boundaries using high-resolution Google Earth images and subsequent digitization using geographic information system (GIS) tools.

- **The second study employed a novel data fusion technique in Thai Binh province, Viet Nam, to generate a spatially disaggregated rice yield map.** Data fusion is a technique by which two satellite images with different spatial and temporal resolutions can be combined to produce a fused product with improved overall resolution. This approach significantly penetrated cloud cover and resulted in a strong relationship between satellite derived vegetation indices and crop-cutting derived yields.

- **The utility of land-use maps developed from satellite data while constructing a sampling frame was explored in the third study.** Estimates for paddy rice areas from the resulting sampling frame were compared to estimates from administrative data. Given that satellite-based estimation methodology is transparent, eliminates under-coverage stemming from an outdated population-based frame, and allowed for the calculation of confidence intervals, they are likely to be more reliable when compared to official estimates.

- **The results of the three methodological studies provide strong evidence in favor of satellite data as a viable alternative to existing methods to generate paddy rice statistics.** Multilaterals like the ADB can leverage across different networks to optimally design technical assistance projects for countries interested in exploring and mainstreaming remote sensing technology for agricultural statistics.

- **Other technological innovations such as drones, computer-assisted personal interviewing (CAPI), and artificial intelligence hold much promise for the future of agricultural statistics.** Drones have made it possible to gather information down to the crop level relatively quickly, which can further improve crop yield prediction. CAPI will boost administrative and survey data quality, availability and costs. Meanwhile, the more novel field of artificial intelligence can transform agricultural statistics through machine learning algorithms that can provide real-time data to facilitate more accurate forecasting. The ultimate vision is for these different technologies to work in tandem and foster better quality data in support of evidence-based policy making for agriculture.

Introduction

About one-third of the labor force in developing Asia relies on agriculture as its main source of livelihood, excluding high income countries where less than 5% are employed by the sector (ADB 2018). Research has shown that agricultural development is critical for reducing poverty in developing countries and that gains in agricultural productivity are central to macroeconomic theories of structural transformation (ADB 2013, Cervantes-Godoy and Dewbre 2010, Christiaensen et al. 2011, Klasen and Reimers 2016). The link between poverty reduction and agriculture has also been given a special emphasis in the Sustainable Development Goals (SDGs), with Target 2.3 specifically aiming to "double the agricultural productivity and incomes of small-scale food producers, in particular women, indigenous peoples, family farmers, pastoralists and fishers, including through secure and equal access to land, other productive resources and inputs, knowledge, financial services, markets and opportunities for value addition and non-farm employment."

Although small-scale food producers play an important role in shaping the agricultural sector with significant positive spillovers to other economic sectors, limited efforts have been made to collect accurate data on this group to facilitate evidence-based planning, management, and monitoring (Carletto et al. 2015). The problem is partly systemic, wherein most government- and donor-supported agricultural projects targeting large-scale agricultural investments do not devote adequate resources to ensure that the underlying data used to measure outcomes are timely, of high-quality, and reflect the reality on the ground. This is exacerbated by other factors such as lack of coordination across a variety of data producers, weak methodological processes, limited human and capital infrastructure, inadequate capacities to collect and analyze data from a policy perspective, and poor-quality metadata and dissemination tools.[1] Such bottlenecks in agricultural data collection and analysis can hamper policymaking and consequently affect the lives of millions of small-scale farmers.

To improve the quality and quantity of agricultural and rural statistics across the world, the Global Strategy to Improve Agricultural and Rural Statistics (GSARS) was endorsed by the United Nations Statistical Commission

[1] FAO. 2010. GSARS. Washington, DC.

in February 2010.[2] One of the key focus areas of the global strategy is the use of technology to improve the timeliness and efficiency of collecting, processing, safeguarding, and disseminating agricultural data alongside lowering associated costs. This also feeds into closing the gap between Tier I and Tier II or Tier III indicators within Goal 2 of the SDGs.[3] While national statistics offices (NSOs) in the region have made tremendous progress in compiling data relevant to the SDGs, they have not been at the forefront of harnessing the power of information and communication technology. Agricultural statistics in most developing countries continue to be compiled through relatively outdated and often costly means. Modern technology is underutilized, and big data methods are virtually untouched.

To pilot innovative agricultural data collection methods in line with the global strategy, such as remote sensing technology and computer-assisted personal interviewing (CAPI), ADB implemented a technical assistance project between 2013 and 2017 through strategic partnerships with NSOs and ministries of agriculture (MoAs) in provinces of four of its developing member countries; Savannakhet in the Lao People's Democratic Republic (Lao PDR), Nueva Ecija in the Philippines, Ang Thong in Thailand, and Thai Binh in Viet Nam.[4] The project was sponsored by the Japan Fund for Poverty Reduction and implemented in partnership with the Japan Aerospace Exploration Agency.

A significant component of the project dealt with understanding the nature of measurement errors that often arise while generating crop area, yield, and production using farmer self-reports or administrative data. By piloting new and cost-effective technologies to estimate these statistics for paddy rice, and comparing them to estimates derived from current best practices and existing in-country methods, the project was able to provide proof of concept for future scale-up of these novel methods.

---

2 The GSARS is based on three pillars: (i) the establishment of a minimum set of core data that countries will provide to meet the current and emerging demands; (ii) the integration of agriculture into their National Statistical Systems in order to meet the needs of data policymakers and other data users and to ensure data comparability across countries and over time; and (iii) helping countries to enhance the sustainability of the National Agricultural Statistical System through governance and statistical capacity building.

3 Tier 1 indicators of the SDGs are conceptually clear, have internationally established methodologies and standards, and are being regularly produced. Tier 2 indicators are conceptually clear, have internationally established methodologies and standards, but data are not regularly produced by countries. Tier 3 indicators are those with no internationally established methodology or standards.

4 ADB. Innovative Data Collection methods for Agricultural and Rural Statistics. https://www.adb.org/projects/46399-001/main.

This section summarizes the major findings and valuable lessons from the pilot activities conducted by ADB, in collaboration with national government agencies, in the Lao PDR, Thailand, and Viet Nam.[5]

Existing Methods for Collecting Agricultural and Rural Statistics in Asia and the Pacific

In most countries across Asia and the Pacific, NSOs are responsible for the compilation of agricultural and rural statistics at the national level. The arrangements may vary across countries, with some countries such as the Philippines designating the task only to the NSO, while others such as Viet Nam dividing the responsibility between the NSO and the MoA. In most cases, data are collected through administrative reporting systems, household surveys, and censuses.

Administrative reporting systems

Administrative reporting systems involve collecting data at the lowest administrative levels (e.g., villages or municipalities), followed by a process of data aggregation and transmission up the numerous administrative levels, until national estimates are produced. The information at the lowest level is typically collected by agricultural personnel or heads of villages or municipalities, who observe harvests and interview key people in the localities, such as farmers and traders.

For example, in the Lao PDR, agricultural statistics are compiled from administrative reports collected by the technical departments of the Ministry of Agriculture and Forestry (MAF). Data at the lowest administrative level are gathered from a collective unit, which consists of 10 to 12 households in a village. These data are subsequently compiled by personnel at MAF district offices, then forwarded to MAF provincial offices. Different technical sections at MAF provincial offices review the data based on specific commodity types, then send summary reports to the relevant technical department of MAF headquarters in Vientiane. The Statistics Division at MAF compiles and disseminates data collected from each of the MAF technical departments in the form of publications (ADB 2016).

5 Results from the Philippines are not presented in this report as not all field activities could be completed due to the occurrence of Typhoon Koppu.

In India, data on irrigated areas are derived from administrative reports of state government departments, while other indicators are compiled by the Directorate of Economics and Statistics within the Ministry of Agriculture, based on village land records. Land-use statistics are also derived from village land records maintained by village accountants (known as Patwari's), wherein information on land-use classifications are recorded and aggregated at successive administrative levels. The compilation of land-use statistics based on administrative records in India is done in addition to a nationwide land mapping conducted by the National Remote Sensing Agency.

The process of collecting data using an administrative reporting system is usually less expensive than conducting large-scale surveys or censuses. However, due to the subjective manner in which administrative data are collected and lacking in external validation procedures, information may be prone to significant measurement errors. Nevertheless, administrative data continues to be the main source of information for policymaking, so the GSARS has recommended the development of robust validation systems, coupled with the training of staff at all levels in government agencies, to improve the quality of such data.[6]

Sample surveys

Many countries across Asia and the Pacific regularly conduct sample surveys to collect agricultural data. These surveys are either conducted on a standalone basis to gather information on specific commodity types (e.g., livestock surveys or crop production surveys) or are integrated into larger household surveys, with modules devoted to specific agricultural indicators (e.g., household consumption, agricultural labor force, and food demand or security). Surveys may be conducted on a monthly, quarterly, biannual, or annual basis, depending on the production cycle of the crop, the resources available to the agency tasked with collecting such data, and the field methods implemented.

---

[6] GSARS. 2017. Improving the Methodology for Using Administrative Data in an Agricultural Statistics System: Final Report. *Technical Report No. 24*. Global Strategy Technical Report: Rome.

For example, monthly surveys conducted by the Philippines Statistics Authority include the Rice and Corn Stocks Survey, the Commercial and Municipal Fishery Survey, and the Farm Prices Survey. Meanwhile, quarterly surveys include production surveys on rice and corn, other crops, livestock, fisheries, and food consumption. The Philippines also conducts agricultural labor surveys biannually. Costs and returns surveys for specific commodities have also been conducted, albeit on an ad-hoc basis. In contrast, Bhutan and the Lao PDR have integrated agricultural modules into general household surveys, with questions on crop, livestock, and fishery production as well as landholding.[7]

Since these surveys follow well-defined sampling strategies, they are likely to produce better quality estimates than administrative data. However, the possibility of under- or over-estimating results may persist. For instance, a farmer may be asked about the specific area cultivated for rice by the farming household but could instead respond with the household's total landholding (including areas used for purposes other than rice cultivation), thereby overestimating results. On the other hand, underestimation of rice-cultivated areas may result if respondents leave out areas given to share-croppers, but which are still used for rice cultivation. Moreover, sample surveys are usually conducted at the end of a harvesting cycle, and there may be a lag of more than a year by the time the results are published. This is a critical shortcoming, particularly if decisions on imports are urgently needed to address food security issues related to an agroclimatic shock.

Census of agriculture

National agricultural censuses are conducted with the objective of collecting data on various agriculture-related topics and cover the entire population of a country. An agricultural census typically serves as the basis for developing a sampling frame to implement agricultural surveys. Based on the recommendations of the FAO, as indicated in the World Programme for the Census of Agriculture 2020, countries should consider including the following in the supplementary census modules: demographic and social characteristics, landholding, irrigation and water management, crops, livestock, agricultural

---

7 Refers to the Bhutan Living Standards Survey and the Lao Expenditure and Consumption Survey, which are large-scale and multi-topic household surveys.

practices, agricultural services, farm labor, household food security, aquaculture, forestry, and management of the holding. [8]

Due to the intensive resources required to plan and implement them, agricultural censuses are usually conducted every 10 years in most countries. However, some countries conduct agricultural censuses every 5 years. India is one such country, conducting its first agricultural census in 1970–1971, with the latest round of data collection conducted in 2015–2016. Likewise, Viet Nam's Rural, Agricultural and Fishery Census has been conducted every 5 years since 2001.

Data Collection Activities in Project Areas

As mentioned earlier, three provinces, namely Savannakhet in the Lao PDR, Ang Thong in Thailand, and Thai Binh in Viet Nam, were selected after extensive consultations with the counterpart government agencies using the criteria of the existence of substantial extent of contiguous paddy rice area. The field activities conducted as part of this project focus on the rainy season of 2015 in each country. Crop-cutting surveys were implemented during the harvesting period associated with the rainy season of 2015, while the farmer recall survey was implemented two to three months after the harvesting was completed.

Detailed questionnaires, manuals, and associated training materials were developed and adapted to the local context of each country. The surveys were administered on paper by field enumerators and subsequently verified by field supervisors. The completed questionnaires were returned to the headquarters of each government agency where they were double data entered and cleaned. In addition to collecting area, yield, and production-related data, ancillary information on the household, plot characteristics, crop variety, etc. were also collected.

This study employs a three-stage stratified sampling methodology for which the sampling frame used was developed using land-use maps. While the details on the selection criteria of the plots is presented in more detail in the

[8] FAO. 2017. World Programme for the Census of Agriculture 2020 Volume 1: Programme, Concepts and Definitions. Rome.

third methodological study, a total of 256, 135, and 253 plots were sampled in Ang Thong, Savannakhet, and Thai Binh respectively. Once adjusted using sampling weights, the information from the plots can be scaled up to provide provincial level estimates for key paddy rice statistics.

Measurement Error in Land Area, Yield, and Production Estimates

Agricultural statisticians often measure crop productivity in terms of harvested yield, which is mathematically defined as total production divided by the total area planted.[9] In a few countries, such as India and Bangladesh, yields are directly measured through crop-cutting surveys, whereby a sample of subplots are identified and harvested by field enumerators, with sampling weight adjustments made to obtain unbiased estimates. However, many countries still rely on administrative data collection systems or farmer recall surveys for land area and production measures, since these two variables are easily incorporated into a questionnaire and require little additional time or money to collect. Both methods assume that farmers are willing and capable of providing reasonably accurate estimates of land area and production.

Land area

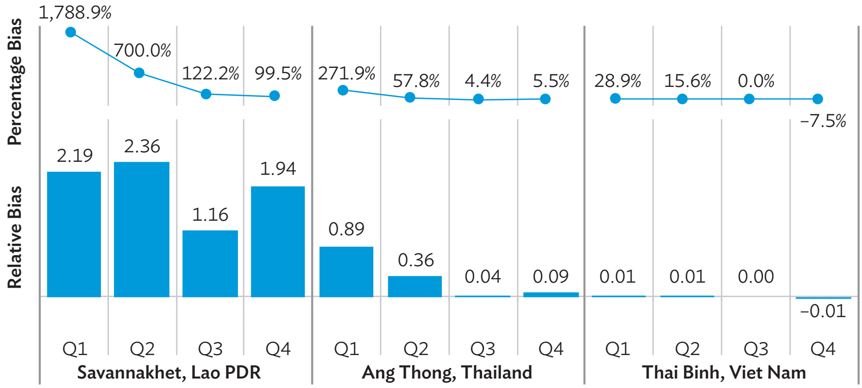
As part of the ADB-supported pilot project, data on plot sizes were collected using self-reported farmer estimates and global positioning system (GPS) devices in the three provinces. To avoid any biases, self-reported estimates for plot size were collected prior to conducting the GPS mapping of the plots. Since GPS is considered to be the new gold standard for land area measurement[10], we assumed it to be the objective value and compared it with plot sizes estimated by the farmers. Figure 1 provides a snapshot of the bias of land area measurement found in the three pilot provinces.

---

9 Two other measures of yield often cited in the agronomical literature are biological yield and economic yield. Biological yield refers to the maximum attainable yield in the absence of any preharvest losses. Economic yield is the value obtained after eliminating any losses incurred in postharvest operations such as drying, shelling, etc.

10 Previously, the compass-and-rope method was the gold standard by the FAO. However, through numerous validation studies conducted across the world, GPS-based area measurement has gained traction as the new gold standard (Carletto et al. 2016).

Figure 1: Land Area Measurement Bias, Self-Reported versus Global Positioning System
(ha)



ha = hectares, Lao PDR = Lao People's Democratic Republic, Q = quartile based on area measured by global positioning system (GPS).
Note: Relative Bias refers to Self-Reported Area – GPS Area in hectares. Percentage bias is defined as Relative Bias*100/GPS Area.
Source: Asian Development Bank estimates using field survey data from the Lao PDR, Thailand, and Viet Nam.

The overall differences between self-reported and GPS-measured plots for the full sample are minimal for Thai Binh. Given the socialist structure of the country, and its well-documented land recording system, this result is not entirely surprising. However, in Savannakhet and Ang Thong, self-reported plot sizes diverge significantly from GPS-based measures. The differences are nonlinear across the land size distribution and nonuniform in the direction of reporting bias. For example, in Savannakhet, the difference is upward biased by 1,788.9% and 700.0% in the first two quartiles, respectively. Likewise in Ang Thong, farmers in the first two quartiles of the land distribution overreported plot sizes by 271.9% and 57.8%, respectively. These descriptive statistics are also presented as densities of the land distribution in Figure 2, disaggregated by measurement method and country.

The results presented in Figure 1 are consistent with recent academic literature on the topic, mainly focused on pilot studies conducted across Africa. Goldstein and Udry (1999), using data from the eastern region of Ghana, found a correlation of 0.15 between GPS and self-reported land size, attributing this to field measurements in the country being historically based on length and not area. Statistically significant differences were also found by Carletto, Savastano, and Zezza (2013) in Uganda, wherein the magnitude varied by plot size. The differences were larger at the tail ends of the distribution of plot sizes, but were smaller for medium-sized plots. Meanwhile, Carletto, Gourlay, and Winters (2015) observed systematic overreporting of plot sizes for smaller plots and underreporting for larger plots. Similarly, Dillon et al. (2017), found that GPS measurements of land area in Nigeria were close to compass-and-rope estimates and more reliable than farmer estimates.

The postulated reasons for mismeasurement in plot size using farmer self-reports include:

(i) use of nonstandard local area measurement units which may exhibit spatial variation.

(ii) intentional underreporting or overreporting by farmers for strategic reasons, such as minimizing property taxes or gaining access to government programs.

(iii) natural tendency for farmers to report farm size by rounding off numbers and providing approximations.

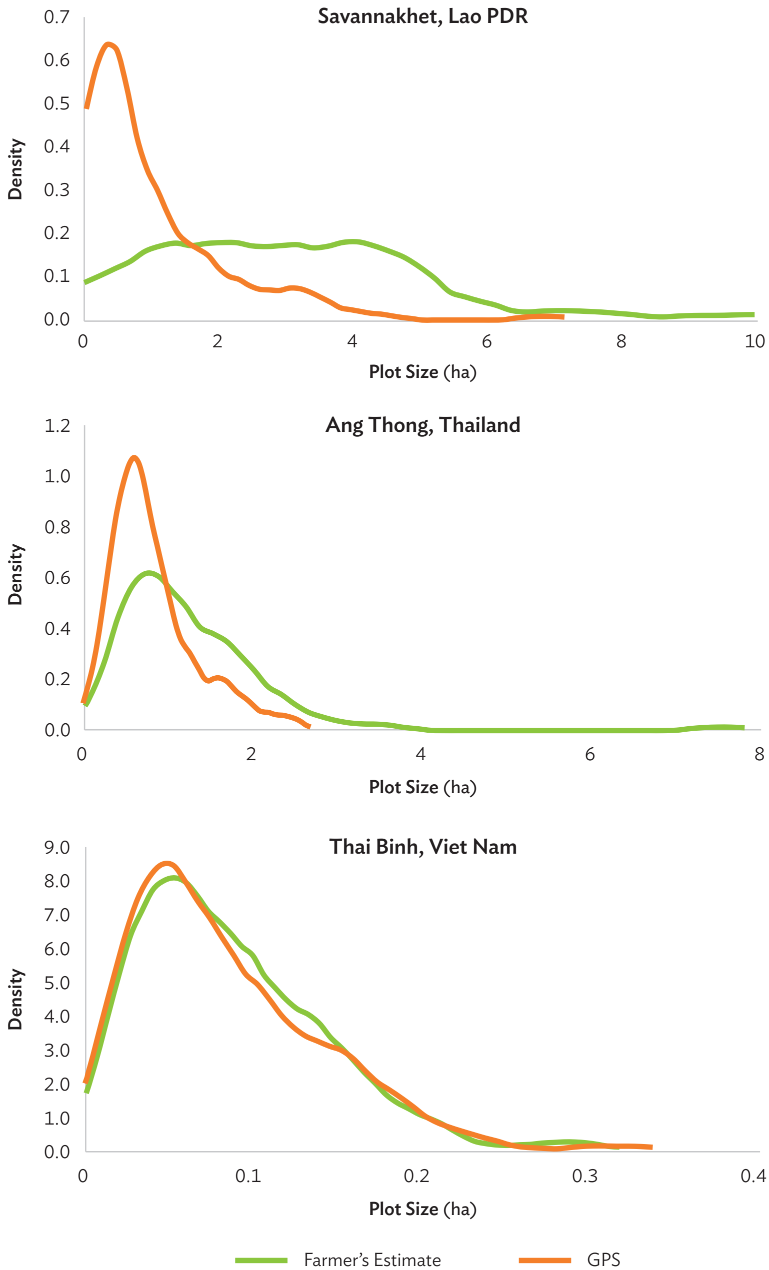
(iv) the terrain on which the plot exists (e.g., slope, shape, etc.).

That said, GPS devices may be subject to an overall position error ranging from 0.5 meters to 4 meters due to satellite position and signal propagation (Hofmann-Wellenhof, Lichtenegger, and Wasle 2008). Since these errors are most likely random, the resulting estimates will remain unbiased.

Production

Extensive academic research in agronomy has shown that the edges of a plot may be more productive than the interior of a plot (Little and Hills 1978, Barchia and Cooper 1996, Ward et al. 2016, Holman and Bednarz 2001). Several reasons have been hypothesized for this phenomena, including increased sunlight exposure, differences in pests or pollination, greater nutrient flow due to reduced competition, and greater water availability (Bevis et al. 2017). For this reason, Fermont and Benson (2011) suggest that production estimates derived from crop-cutting may be biased if the selection of crop-cutting subplots is not random. To avoid any systematic biases in calculating true production, a randomized selection of subplots for crop-cutting was implemented in the three provinces. The derived estimates were statistically weighted at the plot

Figure 2: Distribution of Differences in Plot Area
(ha)

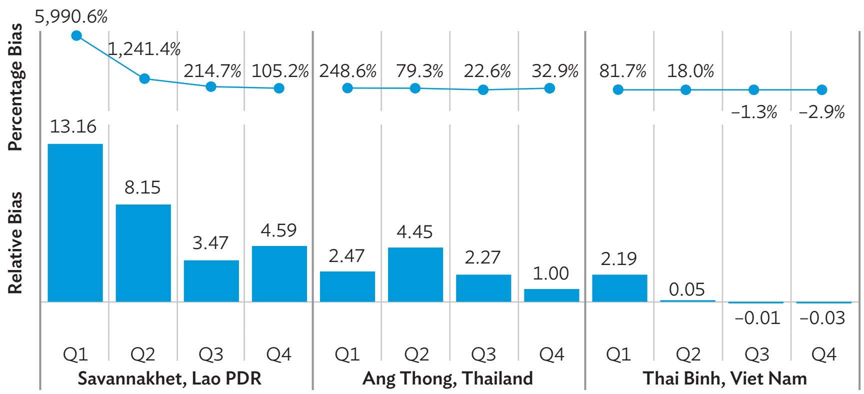


GPS = global positioning system, ha = hectare, Lao PDR = Lao People's Democratic Republic.
Source: Adapted from A. Dillon and L.N. Rao. 2018. Land Measurement Bias: Comparisons from Global Positioning System, Self-Reports, and Satellite Data. Asian Development Bank, *Economics Working Paper Series*. No. 540.

level, using well-defined sampling techniques, and multiplied to GPS derived area estimates to obtain more objectively measured production estimates at the plot level which in subsequent discussions we refer to as ‘objectively measured’ production estimates as compared with production estimates based on farmer’s recall or ‘self-reported’ production.[11]

Figure 3 depicts the descriptive differences between average ‘self-reported’ production and ‘objectively measured’ production.[12] The differences in production estimates are observed to be more drastic than area estimates, with the largest difference seen in Savannakhet, especially in the lower quartiles. Though these differences are nonlinear across the land size distribution, they are not uniform in the direction of the reporting bias. In Ang Thong, average self-reported production diverges significantly from objectively measured production, especially for the two lowest quartiles. The differences are the lowest in Thai Binh, but the direction of bias switches between Quartile 2 and Quartile 3.

Figure 3: Production Measurement Bias, Self-Reported versus Objectively Measured (tons)



Lao PDR = Lao People’s Democratic Republic, Q = quartile based on area measured by GPS.

Note: Relative Bias refers to Self-Reported Production − Objectively Measured Production in tons. Percentage bias is defined as Relative Bias*100/Objectively Measured Production. Tons refers to production in metric tons.

Source: Asian Development Bank estimates using field survey data from the Lao PDR, Thailand, and Viet Nam.

[11] A systematic comparison of yields derived from randomized subplot harvesting and full plot harvesting by Gourlay et al. (2017) has shown no significant differences, strengthening our case for true production estimation using crop-cutting.

[12] Objectively measured production is calculated using yields from crop-cutting and area from GPS.

The results of this study are consistent with the findings from the academic literature. Gourlay et al. (2017) show that farmer-reported production estimates exhibit an upward bias in the lower half of the plot area distribution relative to actual estimates. Similarly, Desiere and Jolliffe (2018) note that the degree of error between farmer estimates and crop-cutting measures is systematic in nature, with production more significantly overestimated by farmers on smaller plots in Ethiopia.

While measurement error in land area has been widely studied across different contexts, measurement error in production has only started to receive attention recently (Gourlay et al. 2017, Desiere and Jolliffe 2018). Gourlay et al. (2017) provide a conceptual basis for potential errors:

(i) recall bias, wherein farmers may find it difficult to aggregate production from a previously completed agricultural season;

(ii) rounding-off error, wherein farmers may tend to report production values as whole numbers or as a function of the number of bags that were used to store the crop, when in reality these were not whole numbers (e.g., half a bag);

(iii) intentional bias, which can work both upward and downward (e.g., if a farmer expects to receive an incentive, such as a tax break, or a disincentive, such as higher taxes, for reporting specific values);

(iv) the use of nonstandard units to measure production varying across space and time; and

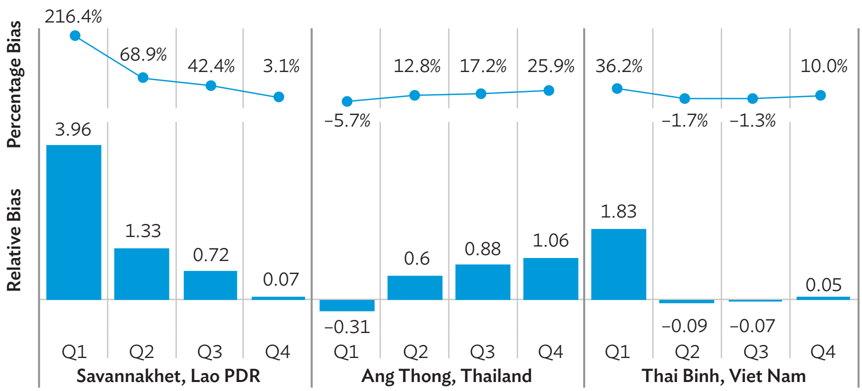
(v) crop condition and state (e.g., shelled versus unshelled, dried, threshed, etc.).

Yield

Self-reported area and production information can be combined to generate self-reported yields, which can then be compared with yield estimates based on crop-cutting. The results of such a comparison are presented in Figure 4. The crop-cutting derived yields are based on weighted estimates of produce found in a randomly selected 2.5 meter (m) by 2.5 m subplot of each sample plot. These were statistically adjusted to represent values in metric tons per hectare.

The overall trend was an upward bias for yield, albeit with variations by quartiles of plot size across countries. In the Savannakhet, farmers tended to overestimate yields systematically, with the degree of overestimation decreasing across plot area quartiles. In Ang Thong, the degree of overestimation increased across plot area quartiles, while in Thai Binh, an upward bias was observed in the smallest and largest quartiles. The overall difference between self-reported and crop-cutting derived estimates was about 10% in both Ang Thong and Thai Binh.

Figure 4: Distribution for Differences in Yield
(tons/ha)



ha = hectare, Lao PDR = Lao People's Democratic Republic, Q = quartile based on area measured by GPS.
Note: Relative Bias refers to Self-Reported Yield – Crop-Cutting Yield in tons per hectare. Percentage bias is defined as Relative Bias*100/Crop-Cutting Yield. Tons refers to production in metric tons.
Source: Asian Development Bank estimates using field survey data from the Lao PDR, Thailand, and Viet Nam.

Abay et al. (2018) provide a theoretical framework for measurement error in yield and its econometric implications, while estimating a production function and testing it on data from Ethiopia. They show that the direction of bias in production and area can be different and the combination of the two biases could lead to biases in yield estimation. Additional research using field data across different contexts is needed to empirically test this hypothesis.

In summary, this report found significant deviations for paddy rice area, production, and yield. Such stark differences are not trivial from a policy perspective. Improvements in such data and estimates may provide some answers to the long-standing debate on the relationship between plot size and productivity, which is discussed in Box 1.

Box 1: The Century Old Mystery of the Relationship Between Plot Size and Productivity for Small-Scale Farmers

For over 100 years, there has been debate amongst academics about who is more productive, the farmer with a larger plot or the farmer with a smaller plot. From an empirical perspective, the academic literature since the 1920s has found results in favor of farmers with smaller plot sizes, starting with Chayanov in Russia, followed by Sen in India, and a host of other studies across Asia, Africa, and Latin America (Dillon and Rao 2018).

This finding is at odds with economic theory, which predicts that, if such a phenomenon were observed in a freely functioning market, larger agricultural plots would be subdivided into smaller plots to achieve gains in efficiency until equilibrium is achieved. It is also at odds with the prevalence of large tracts of land found in developed economies. However, the argument for smaller plot sizes has often been used as the rationale for land distribution in many countries.

Several explanations have been postulated for the existence of the inverse relationship, such as differences in labor endowments between larger and smaller landowners (Eswaran and Kotwal 1985); missing markets for land, labor, and credit (Assunção and Ghatak 2003, Barrett 1996, Carter and Wiebe 1990); and omitted variable bias (Barrett, Bellemare, and Hou 2010; Bhalla and Roy 1988; Chen, Huffman, and Rozelle 2011). None of these theories has been empirically corroborated across different contexts.

A fourth explanation, which has recently gained traction, is that this inverse relationship is a statistical anomaly stemming from measurement error in land area and production reported by farmers (Gourlay et al. 2017, Abay et al. 2018, Desiere and Jolliffe 2018). However, there is limited empirical evidence on the matter across different settings.

Using data from ADB-supported pilot project, the authors of this report were able to statistically test how correcting either for land area measurement bias or both land area and production measurement biases affected the relationship between plot size and productivity. They found that using the accurate global positioning system (GPS) measure for the independent variable, i.e., land area, without correcting for biases in yield resulted in a strong and statistically significant negative relationship, as often found in the literature. However, correcting for yield values using objective crop-cutting derived measures, along with land areas measured by GPS, made the relationship not statistically significant and very close to zero, indicating no impact of plot size on productivity.

	Savannakhet, Lao People's Democratic Republic		**Ang Thong, Thailand**		**Thai Binh, Viet Nam**	
	Yield from FARMER RECALL	Yield from CROP CUTTING	Yield from FARMER RECALL	Yield from CROP CUTTING	Yield from FARMER RECALL	Yield from CROP CUTTING
Land Area from GPS	−0.943***	−0.005	−0.822***	0.026	−0.671***	−0.029

*** indicates statistical significance at $p<0.01$

These results provide preliminary evidence that smaller plots might not be more efficient than larger plots, at least in these three countries. This does not mean that development policies supporting land redistribution from larger plots to smaller plots should be discouraged. Instead, they must rely on justifications such as social and economic equity, rather than empirical findings on the plot size-productivity relationship.

Source: Asian Development Bank estimates using field survey data from the Lao People's Democratic Republic, Thailand, and Viet Nam.

Technology for Agricultural Statistics: A Potential Game-Changer

The findings presented in the previous section provide evidence of significant differences between estimates obtained from objective measurement techniques and those obtained from subjective measurement techniques in the three provinces. Since the objective measurement methods discussed above are more labor-intensive, more time-consuming, and costlier than subjective methods, alternative methods that provide reliable estimates with less effort and at lower costs need to be explored. Remote sensing and satellite-based technologies present a window of opportunity and may be game-changers for existing field measurement methods.

While some countries across Asia and the Pacific, such as the People's Republic of China and India, are already at the forefront of applying such technologies for agricultural statistics, others are still in the process of exploring their merits and limitations. Strategic partnerships with international organizations such as ADB and South-South cooperation among countries can build capacities within NSOs and will lead to long-term dividends in terms of better quality and more timely data.

Apart from their applications in crop area and production estimation, remote sensing and satellite-based technologies can also be tapped to monitor plot conditions and crop growth. Agrometeorological information, crop damage assessments during times of calamity, and forecasting are other interesting applications of remote sensing. For example, after Typhoon Haiyan devastated the northeastern part of Leyte Province in the Philippines in 2013, satellite images were released by the International Rice Research Institute (IRRI) in collaboration with sarmap and the Department of Agriculture-Philippine Rice Research Institute. These satellite images were used to construct pre- and post-typhoon rice area maps, which revealed that nearly 1,800 ha of standing rice crops in the province had been damaged by flooding (IRRI 2013).

Carletto et al. (2016), on the use of remote sensing for household surveys state that, "[l]ittle research is available on the implementation of using remote sensing imagery for area measurement in household surveys. As technology advances and image resolution improves along with affordability, the use of this method becomes more feasible, and is likely to hold promise particularly for the measurement of large plots".

As part of this project, ADB spearheaded three research initiatives to explore:

(i) the use of remote sensing for land area measurement,
(ii) estimating rice yields from space, and
(iii) the use of remote sensing for developing a sampling frame.

The results of these three research initiatives are presented below.

Remote sensing for land area measurement

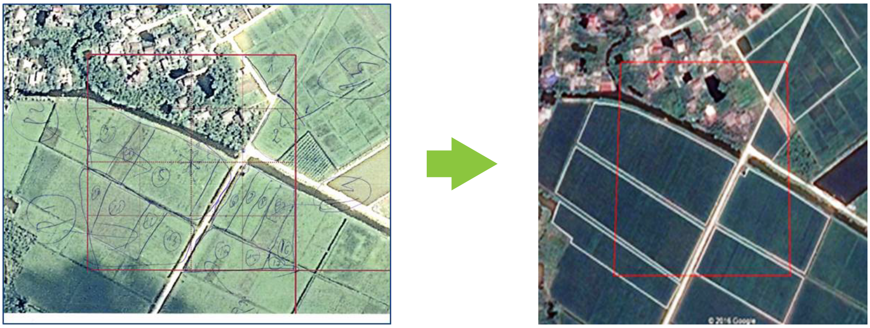
While GPS-based plot area measurements are more accurate than self-reports by famers, there are time, cost, and effort implications. For an enumerator, GPS-based plot measurement involves traversing the plot boundary on foot, which may not be feasible for larger plots or those with standing crops. Given that enumerator wages tend to be one of the largest cost items in conducting surveys, any opportunity for savings while maintaining data quality would be preferable.

The ADB-supported project piloted a third method of plot area measurement, which involved the use of high-resolution Google Earth maps. Under this methodology, plot boundaries were traced onto printed Google Earth maps in consultation with village heads and farmers.[13] These hand-drawn plot boundaries were subsequently digitized using a geographic information system (GIS) software. From a field implementation perspective, the advantage of using the Google Earth method was that enumerators did not have to walk around the plot boundary, as is the case with the GPS method. Figure 5 provides a pictorial representation of this method.

The plot area estimates derived from Google Earth were subsequently compared with GPS-based and farmer-reported plot area estimates (Figure 6). The average difference between GPS-based plot area estimates and those derived from Google Earth was small and not statistically significant,

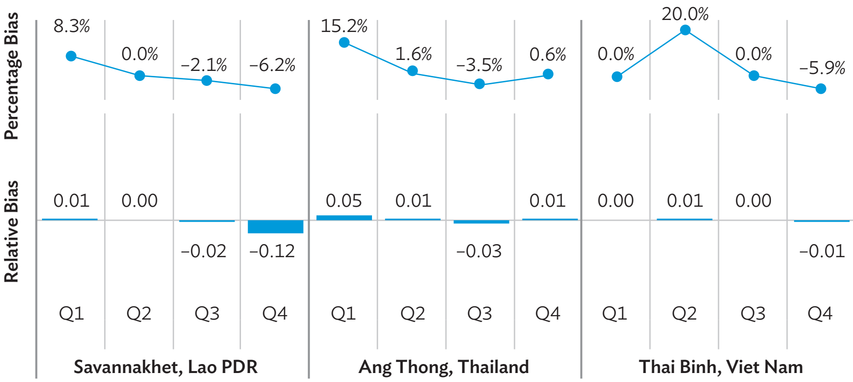
[13] Tracing the plot boundaries onto a handheld device itself would have been ideal, but at the time, the team was not aware of any freely available software options that had such a feature. That said, handheld devices were still useful in identifying point locations at all times, enabling enumerators to match farmer-identified corners of the plot with their corresponding location on the handheld device, which could be easily translated to a printed Google Earth map.

Figure 5: Mapping Plot Boundaries Using Google Earth Images



Source: Asian Development Bank depiction using Google Earth. http://www.earth.google.com (accessed 13 October 2016).

Figure 6: Distribution for Differences in Area
(ha)



GPS = global positioning system, Lao PDR = Lao People's Democratic Republic, Q = quartile based on area measured by GPS.

Note: Relative Bias refers to Google Estimates – GPS Estimates in hectares. Percentage bias is defined as Relative Bias*100/GPS Estimates.

Source: Asian Development Bank estimates using field survey data from the Lao PDR, Thailand, and Viet Nam.

except in the case of Thai Binh, where the average plot size is significantly smaller than those in the other two countries. The differences at the plot level can be attributed to various reasons including the improper tracking of plot perimeters by enumerators, misdirection and misinformation from farmers regarding plot information, machine measurement errors related to the calibration of GPS devices, or the quality of satellite coverage.[14] This shows that the estimates derived from Google Earth are statistically equivalent to GPS-measured plots.

However, improvements in data quality must always be compared with implementation costs when deciding whether to implement any new survey method. A back-of-the-envelope calculation shows that the average cost per plot of using the Google Earth method was nearly 38% lower than the GPS-based method. These results have had important implications for field methods, and propelled the usage of CAPI platforms, such as Survey Solutions, in developing plot area measurements using digital tracing methods (Box 2).

Estimating rice yields from space

As described earlier, rice yields have been traditionally estimated using either administrative records or sample surveys, which can be time-consuming. This has led to lags in the results reaching policymakers for effective planning in the agriculture sector. Remote sensing may be a viable alternative for estimating rice yields efficiently, although global research efforts are still in the early stages (Guan et al. 2018). For the purposes of this discussion, the focus will be on optical images, which can be thought of as a regular photograph taken from space but at varying spatial and temporal resolutions (Box 3), akin to how the human eye would see the world.[15]

One of the challenges with using optical satellite data is that there may be few clear images in areas with significant cloud coverage, thereby limiting the number of usable images for constructing indices that provide a proxy for crop yield. In addition, very high-resolution satellite data (sub-5 meter) are prohibitively expensive for large areas and require significant computing resources to analyze.

[14] More details provided in https://www.adb.org/sites/default/files/publication/409421/ewp-540-land-measurement-bias.pdf.

[15] Cloud-free satellite images, also known as radar data, can penetrate through clouds because they rely on microwave sensors. However, the longer wavelength of microwaves may lead to the signals penetrating the surface of the plant and reflecting the bands corresponding the ground underneath. Also, radar data are not as frequently available and are significantly more expensive than optical data.

Box 2: Land Area Measurement Using Survey Solutions

Acquiring high-quality and timely data on agricultural land area at the plot level is daunting. Most surveys collect such data through farmer-recall methods, which may be prone to significant measurement error for several reasons. Currently, the best-quality land measurement data can be obtained either from the compass-and-rope method or by using global positioning system (GPS) devices. Both methods require traversing the perimeter of a plot, which is time-consuming, costly, and operationally challenging.

The use of satellite images for digital tracing has the potential to improve data quality, time, and cost when it comes to land area measurement. The World Bank has taken the concept of digital tracing to the next level in its free computer-assisted personal interviewing (CAPI) platform called Survey Solutions, by introducing a new "Geography" feature. This feature allows embedding offline base maps for both GPS recording and area tracing.

The feature is available in four options: single point, multiple points, path, and area. Unlike other CAPI platforms that only permit recording a GPS coordinate on the go and require connecting to a satellite on the field, Survey Solutions provides enumerators with the flexibility to pin the location (i.e., record the coordinates of the observation) on a pre-downloaded satellite image of the study area in an offline mode. The same can be accomplished for multiple points. For paths, Survey Solutions allows the recording of coordinates at the beginning, end, and vertex points of a line and computes distance on the fly. Finally, the area function allows enumerators to draw and save a polygon on the base map (satellite image) using a finger or a digital pen. It also computes the area of the polygon when needed. These features are accessible offline as long as the base images of the study area are pre-downloaded.

The new feature holds promise in generating plot area boundaries in places where it is hard to link with a satellite to obtain location coordinates. It is also potentially useful in conducting agricultural censuses, wherein enumerators may identify plot areas owned by households remotely using the feature's tracing method without having to traverse the plot manually. The identified plots may then be linked to the survey questionnaire, facilitating the creation of a spatial database of all agricultural parcels in a given area. This can be a starting point for the use of digital frames, and will reduce survey duration, costs, and enumerator effort on the field.

Geography Feature of Survey Solutions

Source: Asian Development Bank estimates using the Survey Solutions platform.

Box 3: Explaining Spatial and Temporal Resolution from a Remote Sensing Perspective

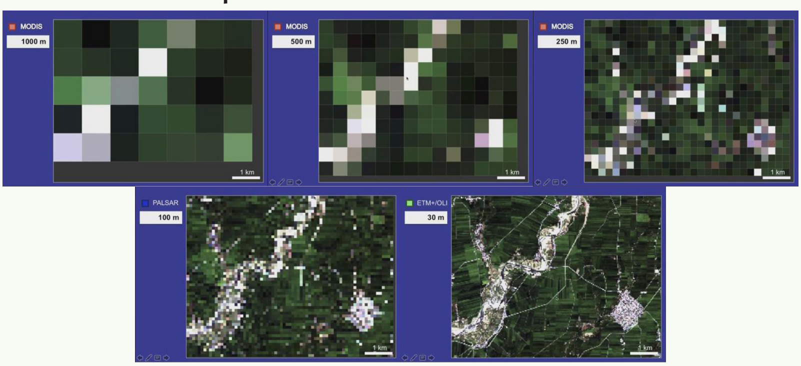
Spatial Resolution

Spatial resolution is defined as the size of the smallest image element (also known as a pixel) in a scene. Clearer and sharper image details can be seen in high-resolution satellite images compared to low-resolution images. A series of images provided in the figure below help illustrate the differences in spatial resolutions of the different sensors and how spatial resolution affects the sharpness of the image. At lower spatial resolutions—higher pixel sizes of 1000 meter (m), 500 m, and 250 m—the images are blurry and the area covered by the image is difficult to identify. However, at higher spatial resolutions (particularly in the smaller pixel sizes of 30 m and 10 m), vegetation, roads, and other objects on the image are clearly identifiable.

Temporal Resolution

Temporal resolution refers to how often a given sensor can capture an image over a given area. Different sensors have different return times—the period it takes for a satellite to return to the same position over the Earth or to collect an image over the same area. The table below presents the return times of some selected sensors. Here, spatial and temporal resolution seem to conflict with each other for freely available satellite data. For example, Moderate Resolution Imaging Spectroradiometer (MODIS) data are more frequently available but have a lower spatial resolution.

Spatial Resolution for Different Sensors



ETM = Enhanced Thematic Mapper, m = meter, MODIS = Moderate Resolution Imaging Spectroradiometer, OLI = Operational Land Imager

Source: Adapted from Asian Development Bank's online training course on the use of remote sensing for paddy rice estimation (http://www.adbx.online/courses/course-v1:ADBx+RS202+2018_01/about).

Characteristics of Selected Sensors

Satellite	Source	Spatial Resolution	Temporal Resolution	Cost	Sensor Type
MODIS	NASA	1km/500m/250m	1–2 days	Free	Optical
ALOS-2	JAXA	100m	14 days	Costly	Radar
Landsat	USGS/NASA	30m	16 days	Free	Optical

ALOS = Advanced Land Observational Satellite, JAXA = Japan Aerospace Exploration Agency, km = kilometer, m = meter, MODIS = Moderate Resolution Imaging Spectroradiometer, NASA = National Aeronautics and Space Administration, USGS = United States Geological Survey

Source: Adapted from Asian Development Bank's online training course on the use of remote sensing for paddy rice estimation (http://www.adbx.online/courses/course-v1:ADBx+RS202+2018_01/about).

To address the issue of cloud cover, improve spatial and temporal resolution of freely existing optical images, and promote the use of remote sensing technology for estimating paddy rice yield in developing countries, ADB piloted a novel data fusion technique in Thai Binh province, Viet Nam. Data fusion is a method that combines two freely available sources of satellite data, one with better spatial resolution and another with better temporal resolution. This approach significantly penetrated cloud coverage, improved the overall resolution of fused satellite image, and allowed the researchers to prepare a spatially explicit map of rice yield for Thai Binh.

The pilot study used Landsat and Moderate Resolution Imaging Spectroradiometer (MODIS) data. Landsat images have a 30 m spatial resolution, but a temporal resolution of 16 days. MODIS on the other hand, passes through the study area almost every day (high temporal resolution) but is available only at 250 m spatial resolution. Merging these two sources of data through a process referred to as data fusion[16] enhances the spatial resolution of the final data to 30 m while improving the temporal availability to 1–2 days. This is observable in Figure 7 where the fused data is seen to have a greater number of clear observations than the original Landsat data.

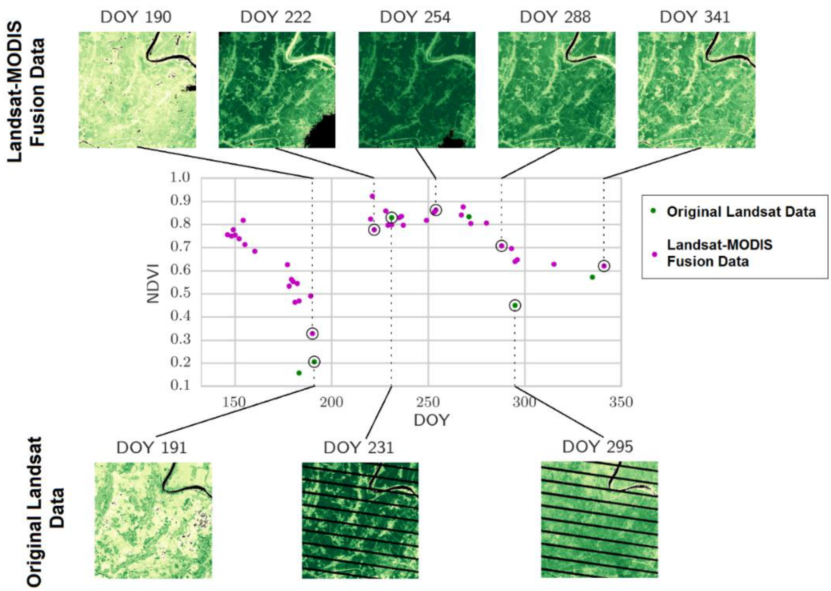
Land cover classification

To identify paddy rice areas from other types of land cover, ADB researchers classified the fused satellite data covering Thai Binh into six categories using the International Geosphere-Biosphere Programme classification scheme (Friedl et al. 2002) and a random forest classifier algorithm (Breiman 2001). The principle behind this kind of classification can be explained as a three-step process:

---

[16] A mature Landsat–MODIS fusion algorithm, the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) was employed to create a fused data product. The STARFM model blends Landsat and MODIS data to generate synthetic daily surface reflectance products at Landsat spatial resolution based on a deterministic weighting function computed by spectral similarity, temporal difference, and spatial distance (Gao et al. 2006). The algorithm requires Landsat and MODIS pair images for the same date with clear day quality.

Figure 7: Normalized Difference Vegetation Index Time Series



DOY = date of year, MODIS = Moderate Resolution Imaging Spectroradiometer.

Note: The series shows a 30 meter (m) by 30 m pixel that combines the original Landsat data (in green points) and the Landsat–MODIS fused data (in purple points). The top and bottom rows show the image data (3,000 m by 3,000 m) that correspond to different time stamps, and the corresponding DOY and Normalized Difference Vegetation Index values at the central of the image. The second rice-growing cycle starts around DOY 200.

Source: Adapted from K. Guan et al. 2018. Measuring Rice Yield from Space: The Case of Thai Binh Province, Viet Nam. Asian Development Bank *Economics Working Paper Series*. No. 541.

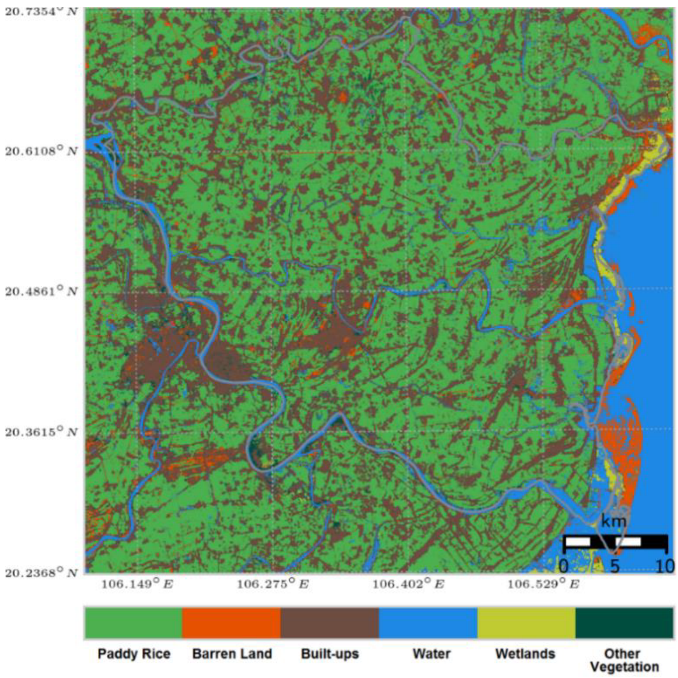
Step 1: Sample pixels from the fused satellite images were selected as evenly as possible across the spatial extent of Thai Binh. The land cover of these pixels were validated through ground observations and visual interpretation of high-resolution images such as those derived from Google Earth.

Step 2: Once a sufficient number of training pixels were obtained for each of the six categories, the threshold values for each of the land cover categories were established and applied to the whole study area using a machine learning algorithm. This generated a first prototype for a land-cover classification map.

Step 3: A random number of pixels were again selected across the six categories (different from the training pixels) and again verified through ground observations or high-resolution satellite data to construct various classification accuracy statistics. This study exhibited close to 91% accuracy for paddy rice areas.

The classification map created for Thai Binh province is shown in Figure 8.

Figure 8: Classified Land Cover Map for Thai Binh Province, Viet Nam



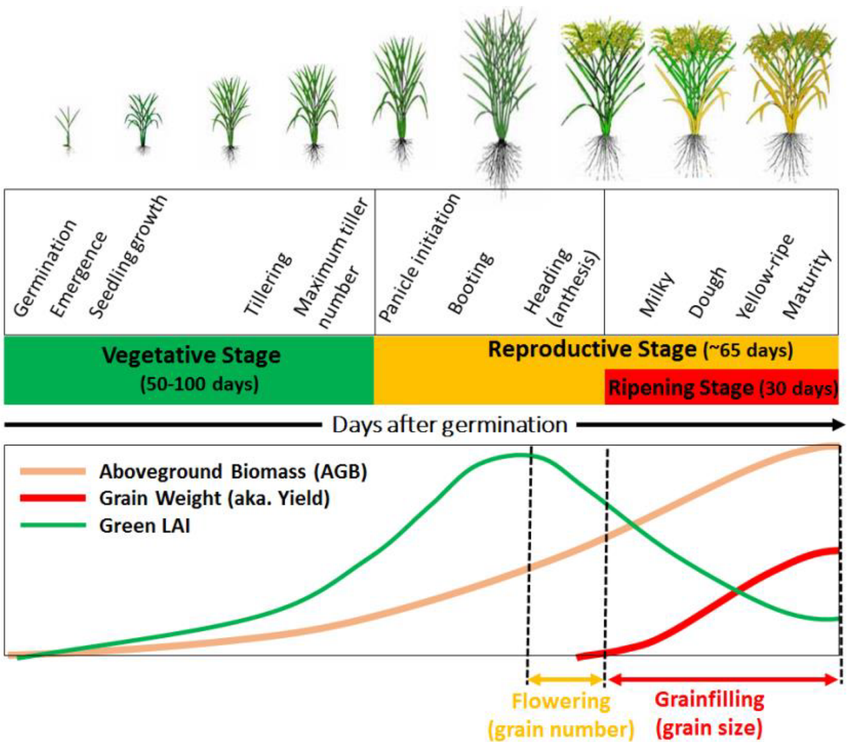
km = kilometer
Source: Adapted from K. Guan et al. 2018. Measuring Rice Yield from Space: The Case of Thai Binh Province, Viet Nam. Asian Development Bank *Economics Working Paper Series*. No. 541.

Yield estimation

Rice yields can be estimated using the land cover classification map. However, before getting into the statistical methodology of predicting rice yields and preparing spatially explicit rice maps, it is important to understand how the growth cycle of paddy rice relates to remote sensing technology.

The growth cycle of paddy rice is depicted in Figure 9, based on the classification of the IRRI (IRRI 2013). The 150-day cycle is broken into two broad stages, vegetative and reproductive. Both stages are further broken down into several substages and, while each substage is important, the flowering stage, which starts 1 day after the heading stage and takes about 7 days, is of particular relevance to this research. Given that each flower results in only one grain of rice, the flowering period largely determines the potential grain yield. Once the plant reaches its ripening stage, the number of grains is fixed, but the size of the grain can increase. Therefore, the final yield is a simple product of the number of grains and the average grain size per unit area.

Figure 9: Growth Cycle of Paddy Rice: A Conceptual Framework to Model Crop Yield

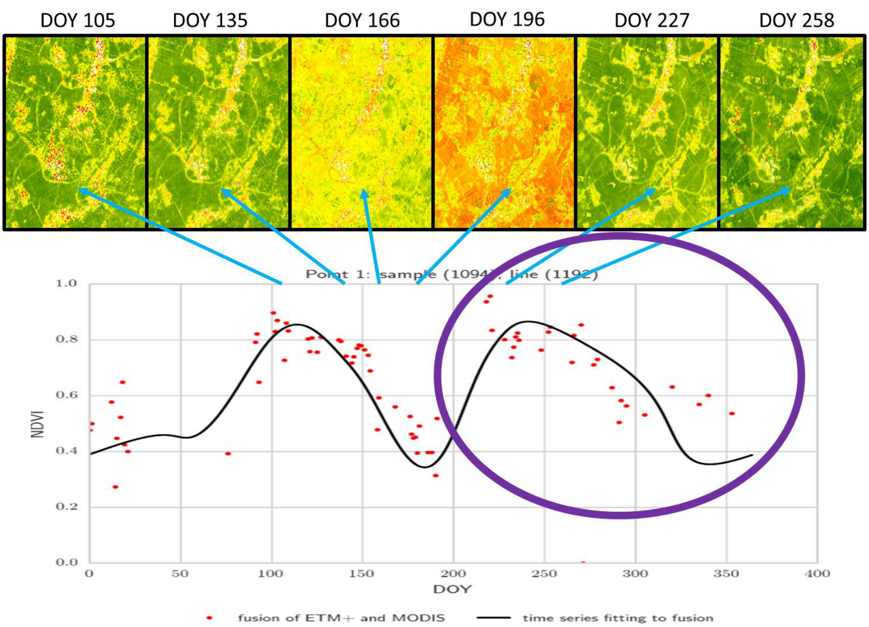


LAI = Leaf Area Index

Source: Adapted from: http://www.knowledgebank.irri.org/images/stories/crop-calendar-growth-dsr.jpg.

As the resolution of the fused satellite product (30 m) does not permit observing rice grains directly, vegetation indices, which are simple graphical indicators derived from satellite data that serve as a proxy for crop growth, were used. One such vegetation index, the Normalized Difference Vegetation Index (NDVI), exhibits an interesting pattern for rice, reaching peak values around harvest time. This can be seen in Figure 10, whereby the peak values of the NDVI are associated with the two major harvesting seasons for paddy rice in Thai Binh—April to May and September to October. Focusing on the second growing season of rice, the study estimated that the average peak value attained for each sample plot on which crop-cutting was implemented, which is observed in Figure 10.

Figure 10: Plotting Time Series of Normalized Difference Vegetation Index Values

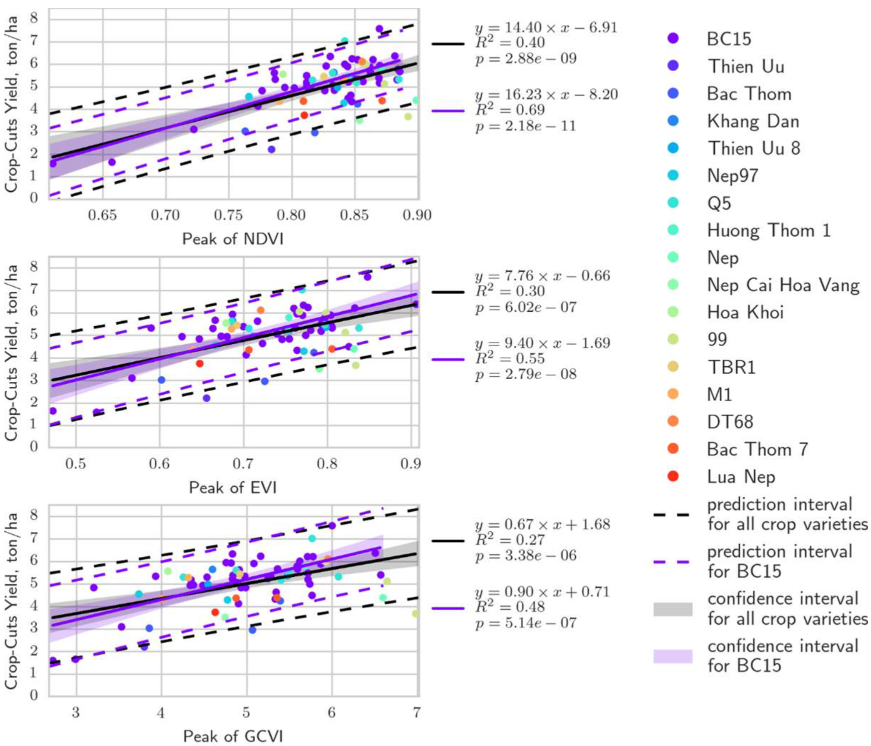


DOY = date of year, ETM = Enhanced Thematic Mapper, MODIS = Moderate Resolution Imaging Spectroradiometer, NDVI = Normalized Difference Vegetation Index.
Source: Asian Development Bank estimates.

Ground data on rice yields obtained through the implementation of crop-cutting at randomized locations across the province are also available. By spatially linking the peak NDVI values for pixels in the same location as the random crop-cutting sites, estimating a quantitative relationship between the

NDVI and crop-cutting yields is possible. The same concept is applied to other vegetation indices such as the Enhanced Vegetation Index (EVI) and Green Chlorophyll Vegetation Index (GCVI) and the results are presented in Figure 11. The NDVI was found to the have strongest relationship with crop-cutting yields.

Figure 11: Linear Regression Model between the Peak of Vegetation Indices and Crop Yield



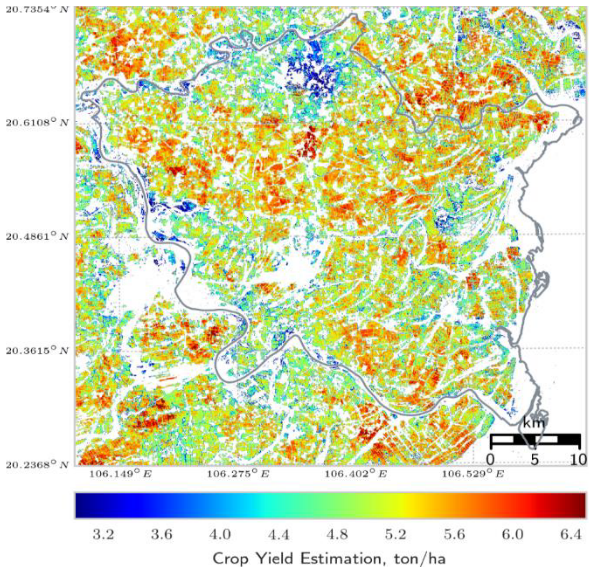
EVI = Enhanced Vegetation Index, GCVI = Green Chlorophyll Vegetation Index, ha = hectare, NDVI = Normalized Difference Vegetation Index.

Note: The vegetation indexes are NDVI, EVI, and GCVI. Crop yield for all the crop varieties are represented by the black line and BC15 by the purple line. Colors of the dots refer to different crop varieties.

Source: Adapted from K. Guan et al. 2018. Measuring Rice Yield from Space: The Case of Thai Binh Province, Viet Nam. Asian Development Bank *Economics Working Paper Series*. No. 541.

This estimated relationship between the NDVI and crop-cutting yield can now be used to produce a spatially explicit map of crop yield for Thai Binh (Figure 12). The mean value obtained from this exercise of 5.0 t/ha was found to be very close to the crop-cutting yield of 4.94 t/ha.

Figure 12: Spatially Explicit Yield Map Based on Normalized Difference Vegetation Index



ha = hectare, km = kilometer
Source: Adapted from K. Guan et al. 2018. Measuring Rice Yield from Space: The Case of Thai Binh Province, Viet Nam. Asian Development Bank *Economics Working Paper Series*. No. 541.

The results hold promise for adoption by NSOs and MoAs for multiple reasons. Firstly, all the data used in this study are freely available. Next, the innovative data fusion technique resulted in improving the spatial and temporal resolution and minimizing the cloud cover issue. While the methods implemented in this study are not straightforward, strategic capacity building programs, as was done in the pilot project, can help staff from NSOs and MoAs to learn and apply the techniques.

A few things need to be borne in mind should other countries plan on implementing similar activities. Firstly, this study only made use of crop-cutting data for one season for one province. Data for different provinces over time will be needed to fine-tune the methodology. Secondly, at the time of study, other satellite images with higher spatial and temporal resolution were not available. Recently, the European Space Agency launched the Sentinel 1 and Sentinel 2 satellites which have the capacity to generate data at significantly higher spatial (10 m) and temporal (every 5 days) resolutions. Testing the algorithms above with higher-resolution data is likely to yield stronger results. Finally, as satellite-based algorithms of yield estimation mature through artificial intelligence techniques, periodic and strategic field data will be necessary for validation purposes. That said, the infrastructure is already available for future scale-up in Viet Nam, and significant benefits are foreseen for other countries looking to adopt such methods.

Remote sensing for developing a sampling frame

The starting point for most agricultural surveys is a census-based sampling frame, which is basically a list of all agricultural households within the total population of a country. A sampling frame is used to draw a sample that is ultimately surveyed to obtain key agricultural estimates. However, in some countries, a complete frame may not be available if the reference is a census with low coverage, or the existing lists of sampling units change rapidly, rendering the list frame out of date (Griffin 2014). Field-listing activities may not reflect the reality on the ground if households systematically overreport or underreport agricultural holdings. Semi-nomadic households engaged in agriculture, or fully nomadic households without fixed dwellings, may not be properly represented under this approach, leading to substantial biases in agricultural statistics (Himelein et al. 2014).

To circumvent the problems with census-based sampling frames, the ADB project piloted the use of remotely sensed data and GIS techniques to construct an alternative frame based on land-use. This method is likely to produce more reliable estimates than a census-based sampling frame because it relies on complete coverage of land area instead of outdated population data (Cotter and Nealon 1987). While this does not eliminate the need to collect information on land area through subsequent enumeration, the full population is theoretically covered given that the basis for selection is land area and not households.

Several steps are involved in constructing a frame using remote sensing technology and GIS techniques. First, the geographic scope of the three provinces was identified and divided into nonoverlapping 200 m x 200 m square meshes (Figure 13). These meshes were stratified into four categories—high probability, medium probability, low probability, and very low probability —based on the expected likelihood of finding paddy rice area in each square mesh.[17]

Figure 13: Sample Mesh
(200m x 200m)

m = meter
Source: Asian Development Bank depiction using Google Earth. http://www.earth.google.com (accessed 13 October 2016).

Since it was not possible to visit all areas, a sample of 120 meshes was selected for each province, such that the number of selected meshes was greater in the stratum where the expected likelihood of finding rice-growing plots was highest, and lower in areas with low or very low likelihood of finding rice-growing plots.[18] The distribution of the total number of meshes in the frame by stratum and sample replacement meshes selected for Savannakhet, Ang Thong, and Thai Binh is shown in Table 1.

---

17 Two sources of rice maps were utilized to implement the stratification process: (i) rice extent maps using 2015 MODIS data produced by the IRRI and (ii) land-use maps from 2009 produced by the European Space Agency under its GLOBCOVER initiative. These sources allowed for identification of land most recently used for growing rice and the compilation of information on areas where rice cultivation has been the standard land-use for several years.

18 The total number of meshes was based on the expected number of rice plots to be found and interviewed in each stratum using data from pretests and the available budget for the pilot project.

Table 1: Distribution of Meshes in the Sampling Frame					
Stratum	Sample Meshes Selected	Replacement Meshes Selected	Number of Meshes in Frame		
			Savannakhet	Ang Thong	Thai Binh
IRRI+GlobCover	80	5	80,839	22,105	36,376
IRRI	20	10	4,650	280	589
GlobCover	15	10	154,227	2,777	4,846
Others	5	5	322,391	34	1,815
Total	120	30	562,107	25,196	43,626

IRRI = International Rice Research Institute

Source: Asian Development Bank estimates using field survey data from Savannakhet (Lao People's Democratic Republic), Ang Thong (Thailand), and Thai Binh (Viet Nam).

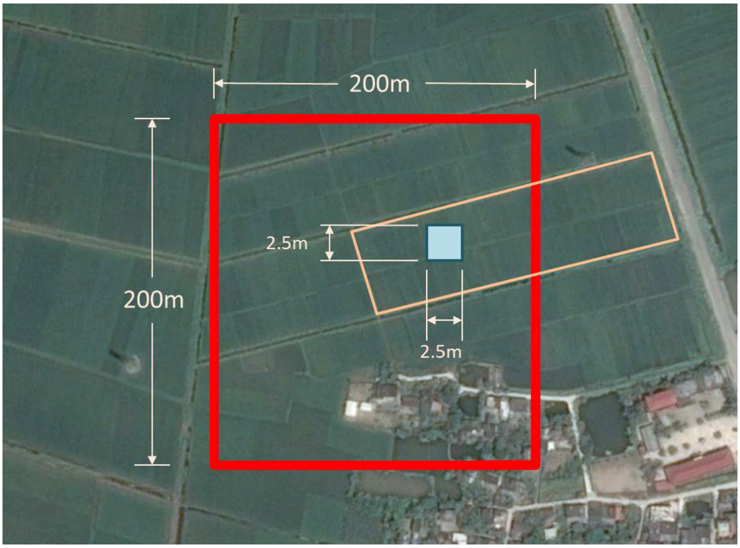
Within each of the 120 sample meshes, a listing of all rice plots was conducted. Plot boundaries were defined based on the definition adopted by the Living Standards Measurement Study Group of the World Bank, where a plot is "a continuous piece of land on which a unique crop or a mixture of crops is grown, under a uniform, consistent crop management system, not split by a path of more than one meter in width, and with boundaries defined in accordance with the crops grown and the operator" (Kilic et al. 2017). Systematic random sampling was then used to select a sample of four plots per mesh from the list of plots that met the selection criterion.

Finally, a random point was selected within each sample plot to identify a 2.5 m x 2.5 m crop-cutting subplot.[19] The rice found in this subplot was harvested, threshed, dried, and weighed to obtain objective yield estimates (Figure 14). This was compared with official estimates for rice-planted areas obtained from administrative records in each country.

Significant differences were observed when comparing the study's estimates with official area estimates (Figures 15, 16, and 17). For Savannakhet, the difference was nearly 25%, while for Ang Thong and Thai Binh, the differences were about 39% and 40%, respectively, albeit in the opposite direction to the Lao PDR. Although it is difficult to pinpoint the underlying reasons behind these deviations as the microdata from the administrative records were not made available, literature suggests likely reasons as being the presence of nonsampling errors, subjective intervention, and political leadership at the local government levels involving subsequent revisions in the administrative data collection methods (ADB 2016).

[19] In this study, the selection criterion was whether rice was planted on a plot with the intention of harvesting in the rainy season of 2015, since the objective was to obtain estimates for only one season using crop-cutting techniques.

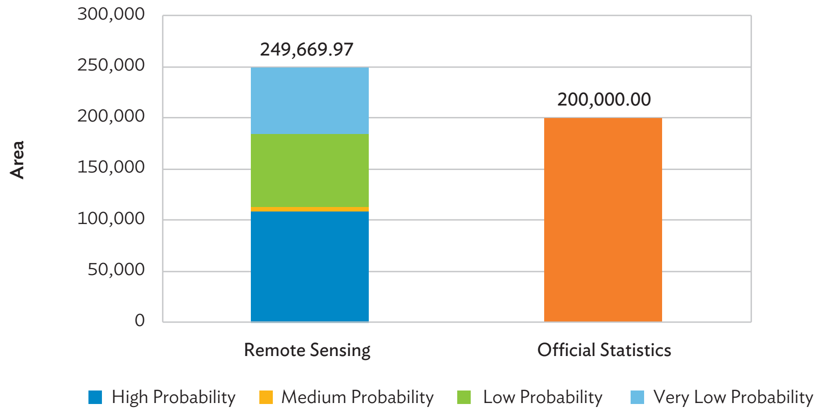
Figure 14: Crop-Cutting on a Subplot
(2.5 m x 2.5 m)



m = meter
Source: Asian Development Bank depiction using Google Earth. http://www.earth.google.com (accessed 13 October 2016).

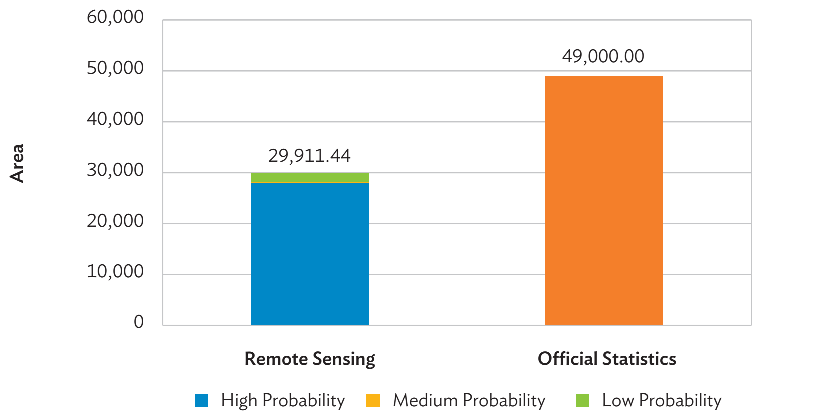
In principle, meshes in the high-probability stratum are expected to have the highest concentration of rice. Yet, even with the best-resolution freely available satellite data used for stratification at the time of this project, some meshes in the high-probability stratum in the three provinces were found to have no rice planted. There are two possible explanations for this: (i) the power of discrimination in the satellite imagery and stratification might not be sufficient; or (ii) field teams might not have accurately reported the status of all meshes, thereby systematically excluding some rice-growing meshes from the survey. For this reason, it will be necessary to improve the land-use stratification of the frame by using higher-resolution satellite images and a greater power of discrimination in the models used for defining the strata. Sentinel 1 and Sentinel 2, which have better resolution than existing freely available satellite images, are likely candidates for future research.

Figure 15: Area Estimate Comparison from Remote Sensing and Official Statistics for Savannakhet, Lao People's Democratic Republic
(ha)



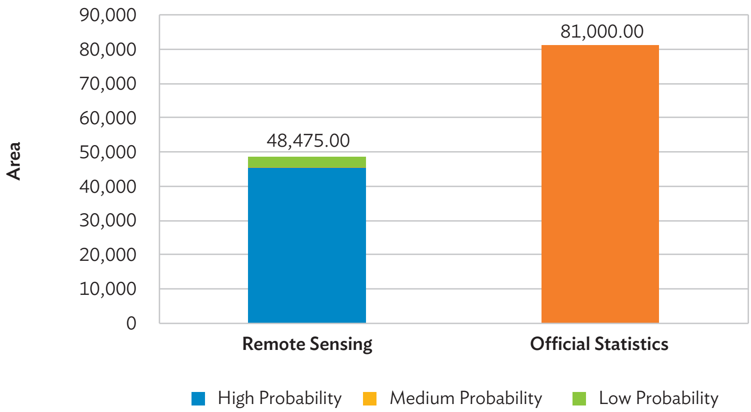
ha = hectare
Source: Asian Development Bank estimates using field survey data from Savannakhet, Lao PDR.

Figure 16: Area Estimate Comparison from Remote Sensing and Official Statistics for Ang Thong, Thailand
(ha)



ha = hectare
Source: Asian Development Bank estimates using field survey data from Ang Thong, Thailand.

Figure 17: Area Estimate Comparison from Remote Sensing and Official Statistics for Thai Binh, Viet Nam
(ha)



ha = hectare
Source: Asian Development Bank estimates using field survey data from Thai Binh, Viet Nam.

Other Technological Innovations for Agricultural and Rural Statistics

The ADB studies have provided a glimpse of the benefits of utilizing satellite-based technology for collecting quality agricultural data. However, such data are not a panacea and have their own limitations. For example, satellite data may not be useful for predicting crop health as very high-resolution images of the crops are needed. Yet, addressing crop health is important from a policy perspective as land is a finite resource and addressing pre-harvest crop damage is likely to lead to gains in crop yield. This does not discount the contribution of satellite-based technologies in improving the quality of data, but rather highlights opportunities for complementary innovations to fill in the gaps and, in turn, create a more holistic picture of agriculture.

Considering how fast the pace of technological progress has been, there exist a plethora of innovations with the potential to improve data quality and, consequently, influence agricultural policy. We live in an age where gigabytes of information and data can be stored in a thumb-sized drive. Further, the ease with which information can be disseminated and downloaded, to and from any part of the world, has spawned unprecedented progress in data science, big data, machine learning, and artificial intelligence. Such unprecedented technological advancement can only help propel innovation in agricultural statistics.

Drones

A noteworthy innovation that has been making waves in agriculture is drone technology. Similar to satellite-based technologies, drones, also referred to as unmanned aerial vehicles, are capable of extensively and rapidly collecting a vast amount of information over large plots of land, especially in remote or hard-to-reach areas. Drones can provide aerial maps that give more vivid snapshots of plots, providing a more concrete gauge of land-use and circumventing issues of cloud cover evident with optical satellite data. What also sets drone technology apart is its ability to assess crop health, a feature that satellite-based technologies currently do not permit. Using relevant camera accessories, drones can monitor the health of crops based on variables such as temperature, chlorophyll levels, and contaminants (Reinecke and Prinsloo 2017).

In Sri Lanka, the International Water Management Institute has been using drones to monitor rice crops in Anuradhapura district (Siddiqui 2016). The institute's study points towards the usefulness of drones to undertake field mapping in hard-to-reach areas in a short timespan. It also indicates that drone technology can help farmers detect when rice fields are under stress and identify areas prone to flooding.

In 2015, a pilot test was conducted in Poland by the Central Statistics Office and the Institute of Geodesy and Cartography, to assess the production of grasslands on 21 plots covering about 460 ha (Milewski 2016). Evaluating the drone's impact, they lauded its advantage in capturing pictures from lower heights to obtain images of higher spatial resolution, recognizing its value in providing additional significant material for estimating grassland production.

A research project[20] in Ghana, using aerial imagery provided by drones, confronted the challenge of understanding crop productivity over space and

---

[20] The research project is called Yieldgap and involves researchers from the University of Lund and the Swedish University of Agricultural Sciences.

time. A major consideration in doing so was that the lower spatial resolution of typical remote sensing technology, coupled with the vegetation index, was not fit for complex agrarian landscapes and systems such as those in sub-Saharan Africa. Given that traditional vegetation indices based on satellite data have been devised for structured agricultural practices such as larger and fixed plots with monocropping, they may not capture the different agricultural practices that characterize developing countries, where crop rotation and intercropping is prevalent and plots are significantly smaller and uneven. Moreover, the study areas in question were rarely revisited by satellites and often covered by cloud, rendering access to satellite imagery for time-series measurement very difficult.

To address the limitations of traditional remotely sensed data, drones were used in the Ghana project to capture higher-resolution aerial images and data. The resulting drone images were said to exhibit spatial resolution of 3 to 4 centimeters, which is significantly higher than satellite-generated images. The final images were also crisp enough to show crop details in smaller fields. Ultimately, the methodology proved appropriate and useful for delineating and classifying maize crops and calculating the vegetation fraction, which is important in estimating yields (Hall et al. 2018).

Apart from generating high-resolution images for aerial mapping and potentially providing a faster and more accurate way of measuring land and estimating crop yields, drones equipped with multispectral sensors can get detailed images of plants that is useful for crop health analysis. Consequently, a series of pilot projects were conducted in Uganda to assess the potential for drone technology to add value to agricultural development in the country.[21] The benefit of drone technology was seen in its ability to assess the health of live vegetation, which provided warning signals or early information on plant or crop health issues. This information allowed for early identification of appropriate treatment or solutions, and proved to be helpful in accurately estimating the amount of fertilizer, pesticide, or other inputs required to optimize yields.

[21] Technoserve is an international nonprofit that promotes business solutions to reduce poverty in developing countries. Its Innovations in Outcome Measurement arm initiated the series of drone pilot projects in Uganda. http://www.technoserve.org/files/downloads/case-study_eyes-in-the-sky-for-african-agriculture-water-resources-and-urban-planning.pdf.

Figure 18: Drone Image of Small-Scale Farms in Ghana

Note: The blue areas depict bare soil or soil covered by debris.
Source: Yield Gap Project Group. https://publications.cta.int/media/publications/downloads/ICT082E_PDF.pdf.

While the benefits of drones are plentiful, the technology has its own set of limitations. First, drones can be expensive if different types of sensors are needed to collect detailed information across a multitude of topics. Next, a study into the perceptions of agricultural stakeholders on the application of drone technology to agriculture found that, apart from the high costs, the complexity of the technology is a potential barrier to adoption (Soesilo and Rambaldi 2018). Individuals need to configure the drones for operation and possess knowledge on converting the images into practical or useful information. This entails significant learning costs and human capital investment (Kipkemoi 2018). Third, drones are vulnerable to weather conditions, and it might not be possible to fly them when there are very strong winds or rains. Finally, flying drones may require clearances from the relevant aerospace or aviation authorities in each country, which may be difficult to obtain.[22]

Like any innovation or technology, drones are a work in progress. With the influx of drone manufacturers, costs are likely to go down, and massive improvements will address the technology's current limitations and improve its

[22] As suggested by the following online resource, which compiles and links to many of the world's drone laws and regulations: https://uavcoach.com/drone-laws/.

user-friendliness. That said, there is an opportunity for the conduct of systematic research studies comparing the costs and benefits of using satellite images versus drone derived images within the context of agricultural statistics.

Figure 19: Drone Image of Small-Scale Farms in Uganda Indicating Crop Health

Source: Technoserve's Innovation in Outcome Measurement. http://www.technoserve.org/files/downloads/case-study_eyes-in-the-sky-for-african-agriculture-water-resources-and-urban-planning.pdf.

Computer-assisted personal interviewing

Another innovation that aims to revolutionize agricultural statistics is computer-assisted personal interviewing (CAPI). This process allows the reporting and recording of responses on handheld devices, instead of paper, and builds in checks and balances to improve the flow of questionnaires and minimize data errors. CAPI also facilitates easy data transfer through cloud or internet-based methods, thereby reducing the time and costs of conducting surveys and managing data.

Literature has established some of the advantages that CAPI has over traditional paper-based methods (Caeyers et al. 2012, Rahija et al. 2016, Zhang et al. 2012, King et al. 2013). First, CAPI is seen to eliminate variable costs such as the printing, storage, and transportation of paper questionnaires. There are, however, higher startup costs with the switch to CAPI, such as the investment in equipment such as tablets and cloud-based computer servers. Second, with a cloud-based server and faster internet, CAPI would allow collected data to be immediately transmitted to the headquarters of an NSO, significantly reducing the time from data collection to analysis. Finally, CAPI reinforces higher quality data, as questionnaires are typically programmed with skip, validation, and consistency conditions that check for errors as data are being entered by the enumerator.

Recently, CAPI platforms have started allowing innovative questions to be embedded. For example, questions that ask for a geo-reference tag or location (e.g. GPS), an image (as might be the case for price surveys), a voice recording (for consent from survey respondents), or even a scan of a barcode are now incorporated into CAPI. Other advantages of CAPI include closer field monitoring, collection of time stamps for every question, and a simplified process of remotely executing changes or corrections to the questionnaire (without having to send paper questionnaires to and from the field).

There have been numerous efforts across Asia and the Pacific to incorporate CAPI into agricultural surveys. For example, the Badan Pusat Statistik in Indonesia piloted CAPI with Survey Solutions[23] in 2015. The aim was to speed up results from their crop-cutting survey, improve the quality of data collected, and foster efficient communication between enumerators, supervisors, and headquarters.[24] ADB has also been working with NSOs in three Asian countries[25] to introduce CAPI technology to nationally representative surveys[26]. Most notably, CAPI is being implemented for an agricultural survey in Sri Lanka, covering a nationally representative sample of 25,000 households over two rounds of data collection corresponding to the harvesting seasons in the country. Further, a randomized experiment to provide rigorous evidence on the benefits of transitioning from paper-based surveys to CAPI has been incorporated.

---

23 A noteworthy example of a CAPI product is Survey Solutions, which integrates management tools, and unlike the other products developed, has no associated programming costs and requires little to no programming knowledge.

24 GSARS. 2016. Adopting CAPI Technology for Agricultural Surveys in Indonesia. http://gsars.org/en/adopting-capi-technology-for-agricultural-surveys-in-indonesia/

25 The Lao PDR, Sri Lanka, and Viet Nam.

26 ADB. Implementing Information and Communication Technology Tools to Improve Data Collection and Management of National Surveys in Support of the Sustainable Development Goals. https://www.adb.org/projects/49342-001/main.

Figure 20: Computer-Assisted Personal Interviewing as Part of Project Activities

Image (1) shows data collection using CAPI during crop-cutting exercises. Image (2) shows a farmer in Thai Binh, Viet Nam being interviewed by an enumerator using CAPI. Image (3) shows plot navigation and area estimation using CAPI.

Source: Asian Development Bank.

Artificial intelligence

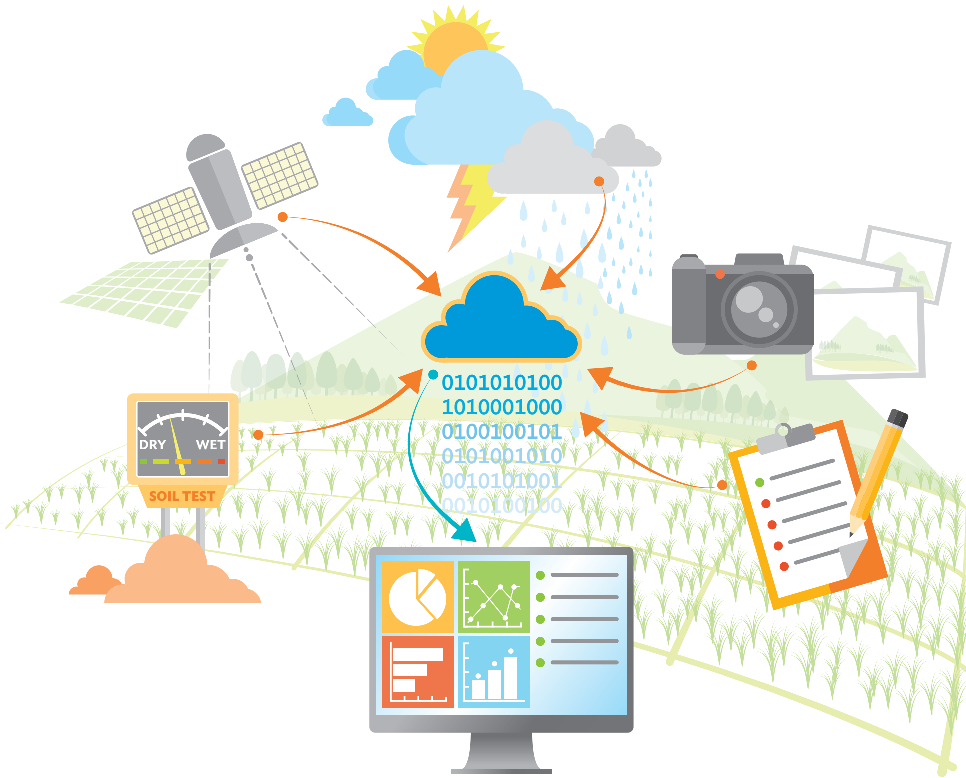
Artificial intelligence is another area that can potentially contribute to the improvement of agricultural statistics. Artificial intelligence refers to the creation of machines that can simulate activities and processes that are typically performed by humans, such as learning (acquiring information along with the rules associated with such information), reasoning, and problem-solving (using a set of rules and boundaries to come up with conclusions). The major fields or parts of artificial intelligence include knowledge engineering, machine learning, machine perception, and robotics. Knowledge engineering involves introducing abundant information relating to the world to machines so that they can act and react like humans. Machine learning, on the other hand, involves introducing learning algorithms (from numerical regressions and logic) into machines so that they can, without supervision, learn and identify patterns from a permutation of inputs or information. Machine perception has to do with honing a machine's ability to interact with, and infer, different aspects of the world, similar to the way humans use senses to relate to the world. Finally, robotics involves developing machines that can mimic human behavior and perform tasks typically performed by humans. Robotics is typically tapped for jobs that are considered hazardous to humans, such as defusing bombs and exploring hard-to-reach places.

Why is artificial intelligence important? First, it can automate a high volume of tasks reliably and without the human limitations. Second, it constantly adapts and learns with built-in learning algorithms. Artificial intelligence sifts through the patterns and structures in data so that the algorithm acquires skill. In other words, algorithms can teach themselves and can further adapt when fed with new data. Because of this artificial intelligence analyzes deeper data, gets the most of data, and attains more accuracy as you feed more data.

A potential application of artificial intelligence for agricultural statistics is the estimation of potential yields during the crop-growing stages using machine learning algorithms (Figure 21). Different types of information collected from the field—such as soil moisture from ground-based sensors, area maps from satellites or drones, weather-related information from meteorological stations, crop characteristics and crop health data from geocoded photographs, and historical data from field surveys—could be fed into a machine learning algorithm programmed to estimate potential yields. This is valuable information that could give a sense of post-harvest losses, as it would provide information on both potential and actual yield. Artificial intelligence can therefore be leveraged to predict conditions to advise sowing, pest control, commodity pricing, and trade. This can also raise incomes of small-scale farmers and lower risks in agriculture. CropIn, for instance, uses multiple data sources and a complex algorithm to help predict yields for different types of crops in many Asian countries[27]. Tellus Labs has similar experience in the United States, wherein their artificial intelligence predictions of corn yields were within only 1% of actual corn yields based on figures from the United States Department of Agriculture (Potter 2017).

[27] CropIn is an Indian agritech startup known to employ data-mining and artificial intelligence to improve crop yield and make farming profitable. http://www.cropin.com/about/

Figure 21: Artificial Intelligence for Compiling Potential and Actual Crop Yield



Source: Asian Development Bank.

Several other applications of artificial intelligence to agriculture have been explored. For example, a team of researchers developed an artificial intelligence that can assess and detect diseases in cassava plants, using the Google open source library to access more than 2,000 images of cassava leaves from Tanzania (Simon 2017). The information from the library was fed into an artificial intelligence architecture, training the computer to identify and diagnose crop diseases and pest damage. The technology was able to identify diseases with close to 100% accuracy.

However, artificial intelligence also has its own set of challenges. First, considering the complexity of the algorithms of artificial intelligence, it can quickly become expensive to program, build, and improve upon. Next, because these technologies learn from existing data, the resulting intelligence and machine learning will be only as effective and as reliable as the data that feeds into the algorithm creation.

Artificial intelligence is the future of agricultural statistics and, more broadly, the agriculture sector. While artificial intelligence might be difficult for countries to take on given its complexity, novelty, and unpredictability, strategic partnerships between the private sector and governments, facilitated by international development organizations such as ADB, will allow such technology to improve and flourish for agricultural statistics.

Conclusion

Agricultural development is essential to poverty reduction and its relevance has been strongly emphasized in the SDGs. Evidence-based policymaking for the agriculture sector has the potential to affect the lives of millions around the world. However, good policies rely on the availability of high-quality, timely, and disaggregated agricultural data, which are sparse in countries with weak statistical systems.

To improve the quality and quantity of agricultural and rural statistics across the world, the Global Strategy to Improve Agricultural and Rural Statistics (GSARS) was endorsed by the United Nations Statistical Commission in February 2010. ADB joined this global initiative as an implementing partner of the Steering Group for Agricultural Statistics for Asia and the Pacific. Through a carefully designed technical assistance project, ADB piloted innovative data collection methods, such as remote sensing and CAPI, to address measurement challenges associated with collecting paddy rice area, yield, and production statistics. The technical assistance was implemented through strategic partnerships with NSOs and MoAs in four countries: the Lao PDR, the Philippines, Thailand, and Viet Nam.[28]

By comparing farmer self-reports for paddy rice area, yield, and production to existing objective methods such as GPS-based area mapping and

[28] Results from the Philippines are not presented in this report as not all field activities could be completed due to the occurrence of Typhoon Koppu.

crop-cutting-based yield and production estimation, this report found evidence of significant measurement errors for all three variables. ADB researchers also noted that these differences are nonlinear across the land size distribution and nonuniform in the direction of the reporting bias. These biases are likely to have significant implications for estimating agricultural production functions, which are the basis for numerous policies such as land redistribution, subsidies for agricultural inputs, commodity pricing, and trade.

Given that the existing objective methods of collecting and compiling data on area measurement, production, and yields are expensive and time-consuming, the project explored the use of remote sensing and GIS techniques as an alternative through three methodological studies. In the first study, researchers compared plot area estimates derived by tracing plot boundaries on high-resolution Google Earth maps to the current gold standard of traversing plot boundaries on foot using GPS instruments. The study results show that the two methods provide statistically equivalent results, but the Google Earth method reduced observation costs by nearly 38%. The second study explored the use of a data fusion technique that combines two different freely available satellite images to improve the overall spatial and temporal resolution of the fused data product. This helps in estimating a stronger relationship between crop-cutting derived yield and the vegetation index derived by remote sensing. This relationship is used to prepare a spatially explicit crop yield map. The third study presented a methodology to construct a sampling frame using land-use maps and GIS techniques and compared rice area estimates derived from the implementation of this sampling frame to official estimates provided by the governments of the relevant countries. Given that this method eliminates potential under-coverage from an outdated population-based frame, and allows for the estimation of measures of precision, the study makes the case that area estimates from this method are more reliable than official estimates.

This report also concludes that there are some limitations in using satellite-based technology, and other technological innovations in agricultural statistics, which are in nascent stages of development, are likely to bring dividends in the long run. For example, drones have made it possible to gather data on plots down to the crop level at a moment's notice, which can potentially contribute to more effective policymaking and provide information to help farmers achieve higher yields. CAPI is expected to bolster administrative and survey data collection efforts. Finally, more complex and cutting-edge artificial intelligence has the potential to revolutionize agriculture with intelligent machinery that can provide optimal information, advisories, and solutions. The ultimate vision is for these innovative data and information sources to come together in support of better-quality and real-time data that facilitate evidence-based policymaking for the agriculture sector.

Although modern technology has its benefits in terms of improved data quality, greater efficiency, and cost-savings, it is important to highlight some of the challenges associated with the adoption of such tools by government agencies. Firstly, there is a need to strengthen the statistical system in terms of human capital. Well-designed learning programs are crucial for imparting new skills and bringing people with different areas of expertise together such as data producers, users, policymakers, international experts, and technicians to share technical knowledge. These could be facilitated through a combination of in-class and e-learning courses and on the practical training. Next, the use of remote sensing technology requires strategic investments and upkeep of physical infrastructure including modern computers with fast processors, access to secure database and servers, reliable internet network, and necessary software to process satellite data into statistical results. Third, adequate financial resources are needed to implement such tools, which requires institutional support and political buy-in. Legal and/or planning documents such as the Statistics Law, and the National Strategy for Development of Statistics (NSDS) should emphasize the need to invest in such technologies.

Multilateral organizations like the ADB can be the starting point for countries interested in exploring new technologies for agricultural statistics. Through carefully planned and well implemented technical assistance projects, multilaterals can leverage across various networks to identify appropriate technologies, and optimally design capacity building activities to produce high-quality, cost-effective, and timely data. This would serve as an impetus for piloting these novel tools to provide proof of concept in terms of costs and benefits to facilitate subsequent scale-up. South-South and triangular cooperation mechanisms may also serve useful in promoting information sharing related to implementation experiences across countries. That said, the value added of multilateral organizations hinges on a critical requirement of the willingness of countries to mainstreaming such tools into their statistical systems by committing adequate resources in the long run to ensure sustainability.

References

K. Abay, et al. 2018. Correlated Non-Classical Measurement Errors, 'Second Best' Policy Inference and the Inverse Size-Productivity Relationship in Agriculture. *IFPRI Discussion Paper 1710*. Washington D.C. http://www.ifpri.org/publication/correlated-non-classical-measurement-errors-%E2%80%98second-best%E2%80%99-policy-inference-and-inverse

Asian Development Bank (ADB). 2013. *Key Indicators for Asia and the Pacific (KI) 2013: Asia's Economic Transformation: Where to, How, and How Fast?* Manila. https://www.adb.org/publications/key-indicators-asia-and-pacific-2013

ADB. 2016. *Results of the Methodological Studies for Agricultural and Rural Statistics*. Manila. https://www.adb.org/publications/results-methodological-studies-agricultural-and-rural-statistics.

ADB. 2018. *Asian Development Outlook (ADO) 2018: How Technology Affects Jobs*. Manila. https://www.adb.org/publications/asian-development-outlook-2018-how-technology-affects-jobs

J. Assunação and M. Ghatak. 2003. Can Unobserved Heterogeneity in Farmer Ability Explain the Inverse Relationship between Farm Size and Productivity. *Economic Letters*. 80(2). pp. 182–194. https://www.researchgate.net/publication/222298710_Can_unobserved_heterogeneity_in_farmer_ability_explain_the_inverse_relationship_between_farm_size_and_productivity

I. Barchia and J. Cooper. 1996. Edge Row Effects on Yields in an Oat Trial. *Australian Journal of Experimental Agriculture*. 36(4). pp. 459–461. http://www.publish.csiro.au/an/EA9960459

C. Barrett. 1996. On Price Risk and the Inverse Farm Size-Productivity Relationship. *Journal of Development Economics*. 51(2). pp. 193–215. https://www.sciencedirect.com/science/article/pii/S0304387896004129

C. Barrett, M. Bellemare, and J. Hou. 2010. Reconsidering Conventional Explanations of the Inverse Productivity-Size Relationship. *World Development*. 38(1). pp. 88–97. https://www.sciencedirect.com/science/article/pii/S0305750X09001168

L. Bevis and C. Barrett. 2017. Close to the Edge: Do Behavioral Explanations Account for the Inverse Productivity Relationship? Mimeo: Cornell University. http://pubdocs.worldbank.org/en/392711496254862815/pdf/D1-Bevis-Barrett-May-2017-Close-to-the-Edge.pdf

S. Bhalla and P. Roy. 1988. Mis-specification in Farm Productivity Analysis: The Role of Land Quality. *Oxford Economic Papers*. 40(1). pp. 55–73. https://academic.oup.com/oep/article-abstract/40/1/55/2361268?redirectedFrom=fulltext

L. Breiman. 2001. Random Forests. *Machine Learning*. 45(1). pp. 5–32. https://link.springer.com/article/10.1023/A:1010933404324

B. Caeyers, N. Chalmers, and J. De Weerdt. 2012. Improving consumption measurement and other survey data through CAPI: Evidence from a randomized experiment. *Journal of Development Economics*. 98(1). pp. 19–33. https://www.sciencedirect.com/science/article/pii/S0304387811001167

C. Carletto, S. Savastano, and A. Zezza. 2013. Fact or Artifact: The Impact of Measurement Errors on the Farm Size–Productivity Relationship. *Journal of Development Economics*. 103. pp. 254–261. https://www.sciencedirect.com/science/article/pii/S0304387813000345

C. Carletto, S. Gourlay, and P. Winters. 2015. From Guesstimates to GPStimates: Land Area Measurement and Implications for Agricultural Analysis. *Journal of African Economies*. 24(5). pp. 593–628. https://academic.oup.com/jae/article/24/5/593/2357539

C. Carletto, D. Jolliffe, and R. Banerjee. 2015. From Tragedy to Renaissance: Improving Agricultural Data for Better Policies. *Policy Research Working Paper No. 7150*. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/313131468194048389/From-tragedy-to-renaissance-improving-agricultural-data-for-better-policies

C. Carletto, et al. 2016. Cheaper, Faster, and More Than Good Enough: Is GPS the New Gold Standard in Land Area Measurement? *Policy Research Working Paper No. 7759*. World Bank, Washington, DC. https://openknowledge.worldbank.org/handle/10986/24840

C. Carletto, et al. 2016. Land Area Measurement in Household Surveys: Empirical Evidence & Practical Guidance for Effective Data Collection. *Living Standards Measurement Study (LSMS) Guidebook*. Washington, DC: World Bank. https://siteresources.worldbank.org/INTLSMS/Resources/3358986-1423600559701/LandGuide_web_final_b.pdf

M. Carter and K. Weibe. 1990. Access to Capital and Its Impact on Agrarian Structure and Productivity in Kenya. *American Journal of Agricultural Economics*. 72(5). pp. 1146–1150. http://academic.oup.com/ajae/article-abstract/72/5/1146/52418?redirectedFrom=fulltext

Z. Chen, W. Huffman, and S. Rozelle. 2011. Inverse Relationship between Productivity and Farm Size: The Case of China. *Contemporary Economic Policy*. 29(4). pp. 580–592. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1465-7287.2010.00236.x

J. Cotter and J. Nealon. 1987. Area Frame Design for Agricultural Surveys. *Staff Report*. Washington, D.C.: U. S. Department of Agriculture, National Agricultural Statistics Service. https://www.nass.usda.gov/Education_and_Outreach/Reports,_Presentations_and_Conferences/GIS_Reports/Area%20Frame%20Design%20for%20Agricultural%20Surveys.pdf

D. Cervantes-Godoy and J. Dewbre. 2010. Economic Importance of Agriculture for Poverty Reduction. *OECD Food, Agriculture and Fisheries Papers, no. 23*. Paris: OECD Publishing. https://www.oecd.org/tad/44804637.pdf

L. Christiaensen, L. Demery, and J. Kuhl. 2011. The (Evolving) Role of Agriculture in Poverty Reduction—An Empirical Perspective. *Journal of Development Economics*. 96(2). pp. 239–254. https://www.sciencedirect.com/science/article/pii/S0304387810001252

S. Desiere and D. Jolliffe. 2018. Land Productivity and Plot Size: Is Measuring Error Driving the Inverse Relationship? *Journal of Development Economics*. 130. pp. 84–98. http://agri.ckcest.cn/ass/608f46dc-945d-4e86-82c4-6fcefe13bf17.pdf

A. Dillon, et al. 2017. Land Measurement Bias and its Empirical Implications: Evidence from a Validation Exercise. *Economic Development and Cultural Change*. https://www.journals.uchicago.edu/doi/abs/10.1086/698309

A. Dillon and L.N. Rao. 2018. Land Measurement Bias: Comparisons from Global Positioning System, Self-Reports, and Satellite Data. *ADB Economics Working Paper Series*. No. 540. Manila: Asian Development Bank. https://www.adb.org/publications/land-measurement-bias-gps-satellite-data

M. Eswaran and A. Kotwal. 1985. A Theory of Contractual Structure in Agriculture. *American Economic Review*. 75(3). pp. 352–367. https://www.jstor.org/stable/1914805

A. Fermont and T. Benson. 2011. Estimating Yield of Food Crops Grown by Smallholder Farmers. *IFPRI Discussion Paper 1097*. Washington, D.C. http://www.ifpri.org/publication/estimating-yield-food-crops-grown-smallholder-farmers

M. Friedl, et al. 2002. Global Land Cover Mapping from MODIS: Algorithms and Early Results. *Remote Sensing of Environment*. 833. pp. 287–302. https://

www.researchgate.net/publication/261707258_Global_land_cover_mapping_from_MODIS_Algorithms_and_early_results

F. Gao, et al. 2006. On the Blending of the Landsat and MODIS Surface Reflectance: Predicting Daily Landsat Surface Reflectance. *IEEE Transactions on Geoscience and Remote Sensing*. 44(8). pp. 2207–2218. https://ieeexplore.ieee.org/document/1661809/

M. Goldstein and C. Udry. 1999. Agricultural Innovation and Resource Management in Ghana. *Final Report to IFPRI Under MP17.* https://www.researchgate.net/publication/253839564_Agricultural_Innovation_and_Resource_Management_in_Ghana_Final_Report_to_IFPRI_under_MP17

S. Gourlay, T. Kilic, and D. Lobell. 2017. Could the Debate Be Over? Errors in Farmer-Reported Production and their Implications for the Inverse Scale-Productivity Relationship in Uganda. *Policy Research Working Paper. No. 8192*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/28369

R. Griffin. 2014. Potential Uses of Administrative Records for Triple System Modeling for Estimation of Census Coverage Error in 2020. *Journal of Official Statistics.* 30(1). pp. 177–189. https://doi.org/10.2478/jos-2014-0012.

K. Guan, et al. 2018. Measuring Rice Yield from Space: The Case of Thai Binh Province, Viet Nam. *ADB Economics Working Paper Series*. No. 541. Manila Asian Development Bank. https://www.adb.org/publications/measuring-rice-yield-space-viet-nam

O. Hall, et al. 2018. Classification of Maize in Complex Smallholder Farming Systems Using UAV Imagery. *Drones*. 2(22). pp. 1–8. http://mdpi.com/2504-446X/2/3/22

K. Himelein, S. Eckman, and S. Murray. 2014. Sampling Nomads: A New Technique for Remote, Hard-to-Reach, and Mobile Populations. *Journal of Official Statistics*. 30(2). pp. 191–213. https://content.sciendo.com/view/journals/jos/30/2/article-p191.xml

B. Hofmann-Wellenhof, H. Lichtenegger, and E. Wasle. 2008. GNSS – Global Navigation Satellite Systems. New York: Springer-Verlag. https://www.springer.com/gp/book/9783211730126

E. Holman and C. Bednarz. 2001. Alley Effect on Several Cotton Cultivars in Small Plot Research. *Communications in Soil Science and Plant Analysis*. 32(1-2). pp. 119–126. https://www.tandfonline.com/doi/abs/10.1081/CSS-100102997

International Rice Research Institute (IRRI). 2013. Rice Almanac. Los Banos. http://books.irri.org/9789712203008_content.pdf

T. Kilic, I. Yacoubou Djima, and C. Carletto. 2017. Mission Impossible? Exploring the Promise of Multiple Imputation for Predicting Missing GPS-Based Land Area Measures in Household Surveys. *Policy Research Working Paper 8138*. World Bank. http://documents.worldbank.org/curated/en/668211499349698549/Mission-impossible-exploring-the-promise-of-multiple-imputation-for-predicting-missing-GPS-based-land-area-measures-in-household-surveys

P. Kipkemoi. 2018. The Pros and Cons of Drones in Agriculture. *Drone Guru*. http://www.droneguru.net/the-pros-and-cons-of-drones-in-agriculture/

S. Klasen and M. Reimers. 2016. Looking at Pro-Poor Growth from an Agricultural Perspective. *Courant Research Centre: Poverty, Equity and Growth (PEG) Discussion Papers, no. 201*. Göttingen: Courant Research Centre Poverty, Equity and Growth. https://www.econstor.eu/bitstream/10419/130592/1/857486284.pdf

J. King, et al. 2013. A Novel Electronic Data Collection System for Large-Scale Surveys of Neglected Tropical Diseases. *PLoS ONE*. 8(9). http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0074570

T. Little and F. Hills. 1978. *Agricultural Experimentation: Design and Analysis*. NY, USA: John Wiley and Sons. https://www.wiley.com/en-us/Agricultural+Experimentation%3A+Design+and+Analysis-p-9780471023524

T. Milewski. 2016. The Use of Satellite Pictures and Data Provided by Drones for the Purpose of Identification of Crops and Assessment of Plant Production. *Prepared for the ICAS VII Seventh International Conference on Agricultural Statistics*. Italy, 24–26 October. https://www.istat.it/storage/icas2016/f30-milewski.pdf

B. Potter. 2017. Corn Could be on Track for 166.2 Bu. https://www.agweb.com/article/corn-could-be-on-track-for-1662-bu-naa-ben-potter/

M. Rahija, et al. 2016. Are CAPI Based Surveys a Cost-effective and Viable Alternative to PAPI surveys? Evidence from Agricultural Surveys in Tanzania and Uganda. *Prepared for the ICAS VII Seventh International Conference on Agricultural Statistics*. Italy, 24–26 October. https://www.istat.it/storage/icas2016/f39-rahija.pdf

M. Reinecke and T. Prinsloo. 2017. The Influence of Drone Monitoring on Crop Health and Harvest Size. *In Proceedings of the 2017 1st International Conference on Next Generation Computing Applications*. Mauritius, 19–21 July 2017. NJ, USA. https://ieeexplore.ieee.org/document/8016168/?part=1

S. Siddiqui. 2016. Sri Lanka's Drone Pioneers. *ICT Update: A Current Awareness Bulletin for ACP Agriculture*. https://publications.cta.int/media/publications/downloads/ICT082E_PDF.pdf.

M. Simon. 2017. Phone-powered AI Spots Sick Plants with Remarkable Accuracy. https://www.wired.com/story/plant-ai/

D. Soesilo and G. Rambaldi. 2018. Drones in Agriculture in Africa and other ACP Countries: A Survey on Perceptions and Applications. *CTA Working Paper 18/02*. https://publications.cta.int/media/publications/downloads/2026_PDF.pdf

B. Ward, B. Roe, and M. Batte. 2016. Potential Profitability of Strip Intercropping with Corn and Soybeans. *Journal of the American Society of Farm Managers and Rural Appraisers*. pp. 149–166. https://ageconsearch.umn.edu/bitstream/236661/2/440-Ward.pdf

S. Zhang, et. al. 2012. Smartphone Versus Pen and Paper Data Collection of Infant Feeding Practices in Rural China. *Journal of Medical Internet Research*. 14(5). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3510690/

Technological Innovation for Agricultural Statistics
Key Indicators for Asia and the Pacific 2018 Special Supplement

This special supplement to the *Key Indicators for Asia and the Pacific 2018* showcases the role that technology can play in improving the quality, timeliness, and frequency of agricultural statistics. The first part presents a summary of existing methods for collecting land area, production, and yield data in the region. The second part discusses measurement errors associated with the existing data collection methods. The third part presents ways to address these measurement errors using remote sensing technology by showcasing results from three methodological research activities undertaken by the Asian Development Bank in three countries: the Lao People's Democratic Republic, Thailand, and Viet Nam. This report concludes with a summary of how other innovations, such as drones, computer-assisted personal interviewing, and artificial intelligence hold promise in transforming the field of agricultural statistics.

About the Asian Development Bank

ADB is committed to achieving a prosperous, inclusive, resilient, and sustainable Asia and the Pacific, while sustaining its efforts to eradicate extreme poverty. Established in 1966, it is owned by 67 members—48 from the region. Its main instruments for helping its developing member countries are policy dialogue, loans, equity investments, guarantees, grants, and technical assistance.

ADB

ASIAN DEVELOPMENT BANK
6 ADB Avenue, Mandaluyong City
1550 Metro Manila, Philippines
www.adb.org