The Use of Big Data for SDG monitoring and official statistics in Asia and the Pacific

21 May 2025

ESCAP
Economic and Social Commission
for Asia and the Pacific

Context

- ESCAP Committee on Statistics called for the facilitation of **capacity development on the use of new data sources and methods for official statistics,** and the conducive institutional setup and development of **tools for the integration of traditional and non-traditional data sources**

ESCAP project on Big Data for Official Statistics

- 'The 2030 Data Decade - Strengthening the institutional capacity of national statistical offices in Asia and the Pacific to use innovative, new and big data sources for official statistics in support of the 2030 Agenda for Sustainable Development'

  - Funded by UN Peace & Development Trust Fund (UNPDF)

  - 1st April 2023 - 30st June 2025

  - 20 beneficiary countries across the region, with in-depth technical assistance in 4 target countries

Beneficiary countries

- North & Central Asia:
  - Azerbaijan
  - Kazakhstan
  - Kyrgyzstan
  - Tajikistan
  - Uzbekistan
- East & North-East Asia:
  - Mongolia
- South & South-West Asia:
  - Bangladesh
  - Maldives
  - Nepal
- South-East Asia:
  - Cambodia
  - Indonesia
  - Lao PDR
  - Malaysia
  - Thailand
  - Timor-Leste
- Pacific:
  - Fiji
  - Kiribati
  - Papua New Guinea
  - Tonga
  - Vanuatu

Direct support to target countries

Focus on usable solutions for Official Statistics

Priorities identified by national users

Partnership based working

- Support for up to 4 countries in incorporating use of big data into their official statistics
  - Identified through Expression of Interest process
  - Focus on moving beyond experimentation towards regular statistical production
- Combination of in-person workshops and technical assistance, along with ongoing remote support

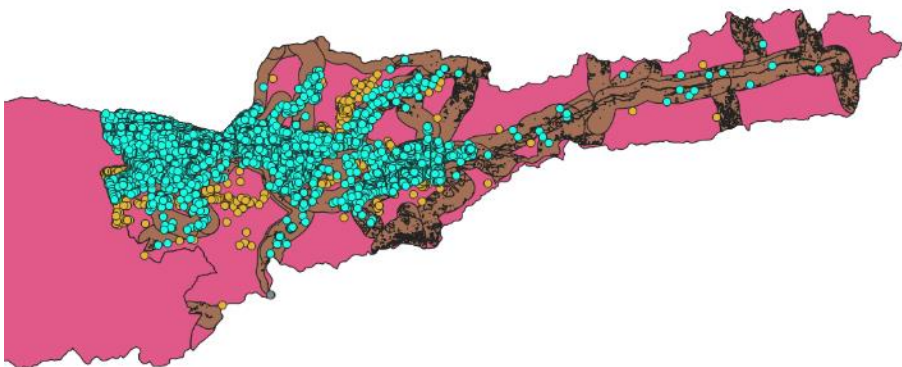# Indonesia – Rice estimation using Earth Observation data

- Accurate, timely data on rice production essential for effective planning
  - Current statistics based solely on field data collection
  - Aim to incorporate satellite imagery to improve quality and reduce costs
- ESCAP & FAO supported BPS – Statistics Indonesia on activities including:
  - Access to new data sources
  - Machine learning model development
  - Field validation/ground truthing in 10 provinces
  - System development
- BPS now working to expand activities to national scale
- *Partners*: BPS – Statistics Indonesia, National Research & Innovation Agency (BRIN), Ministry of Agriculture, Ministry of National Development Planning (Bappenas) and Geospatial Information Agency (BIG), FAO, UNSW, UN Big Data Hub in China
- Case study video: https://www.youtube.com/watch?v=qndP6dzm9KE

Rice Estimation using Earth Observation Data in Indonesia

Kyrgyzstan – Integration of statistical and geospatial information for priority SDG indicators

- Kyrgyzstan wanted to use geospatial data to address priority data gaps for the SDGs

- ESCAP supported through activities including:
  - Technical training workshops in geospatial analysis, using tools including QGIS, and geocoding of latest population census
  - Development of a national geostatistical data inventory & metadata
  - Enhanced cross-government data sharing
  - Development of national roadmap for further geostatistical integration

- Through this support, Kyrgyzstan has developed and implemented methods for estimating indicators including:
  - SDG 9.1.1: Proportion of rural population living near all-season roads
  - SDG 11.3.1: Ration of land consumption to population growth

- *Partners*: National Statistical Committee, State Agency on Land Resources (Kyrgyzgiprozem), Ministry of Transport & Communications, DESA/SD
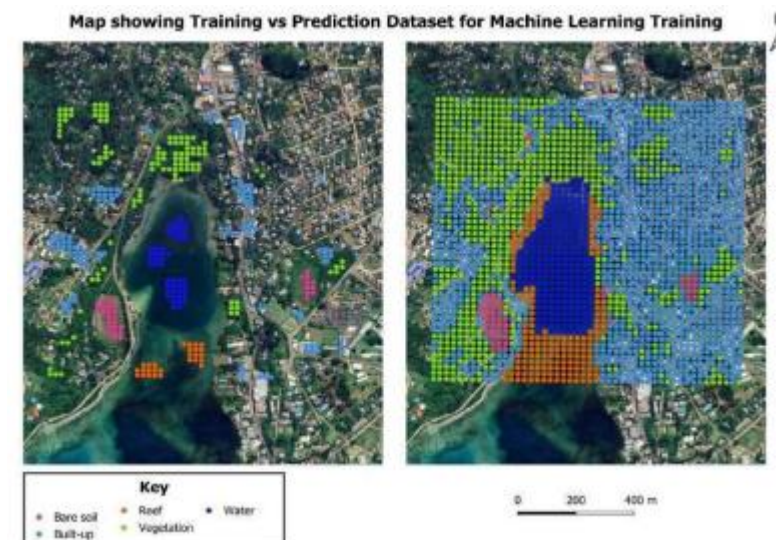
Uzbekistan – use of alternative data sources for price statistics

- Uzbekistan wanted to use alternative data sources for price statistics to increase accuracy, timeliness & granularity and reduce collection costs & response burden

- ESCAP supported activities including:

    - Processing and classification of electronic point of sale (EPOS) data collected by the tax authority and development of research/testing programme for its integration into existing statistics

    - Development of targeted webscraping programme using Python and establishment of MoUs with key data suppliers

    - Establishment of target system and data architecture, utilizing open-source solutions

- Significant achievements in reducing cost and increasing speed of data collection, with further benefits to come as outcomes of research implemented

- *Partners:* Statistics Agency, Tax Agency, Members of UN CEBD Task Team on Scanner data

- Experiences, achievements & general lessons learned shared in ESCAP Stats Brief: https://www.unescap.org/kp/2025/alternative-data-sources-price-statistics-uzbekistans-experience-modernizing-consumer-price

Vanuatu – production of land cover maps & land accounts

- Government of Vanuatu wants improved methods to monitor land cover and environmental degradation
  - Aim to establish robust machine-learning driven land accounting system, based on UN-SEEA framework, using EO data
- ESCAP provided support through combination of in-person workshops and remote follow-up
  - Training in data science and machine learning
  - Geospatial data processing, analysis and interpretation
  - Development and management of automated data pipelines
- Production framework developed and associated technical capacity within government enhanced.
  - Vanuatu working to finalise methodology and systems over next 3-6 months
- *Partners:* Vanuatu Bureau of Statistics, Department of Lands, Department of Forests, SIAP, SPC/Digital Earth Pacific

Map showing Training vs Prediction Dataset for Machine Learning Training

Key
- Bare soil
- Built-up
- Reef
- Vegetation
- Water

0   200   400 m

Support for beneficiary countries

Focus on usable solutions for Official Statistics

Priorities identified by national users

Share knowledge & resources as widely & openly as possible

- Project offered support in the following areas:
  - Integration of statistical and geospatial data
  - Use of earth observation data for small area estimation
  - Use of web-scraping for price statistics
  - Use of mobile positioning data (in partnership with the new UN Regional Hub for Big Data and Data Science, hosted by BPS-Statistics Indonesia).
- Support included remote training and mentoring, supported by in-person workshops
  - All 20 beneficiary countries received support in at least one of the above areas, based on expressions of interest

High-Level Seminar on Integration of Statistical and Geospatial Information

- **28-30 November 2023 in Bangkok**
  - **Brought together NSOs and NGIAs from 20 countries (12 from Asia-Pacific)**
  - **Organised in collaboration with UNSD and Statistics Norway**

- **Objectives:**
  - **Share global resources and frameworks and highlight how they can be used to support geo-statistical integration**
  - **Support countries in assessing current capacity & identifying priority actions**

- **Information and resources:**
  https://www.unescap.org/events/2023/high-level-seminar-integration-geospatial-and-statistical-information

Use of mobile positioning data for official statistics

- **Mobile Positioning Data can be used to provide timely granular data on topics including tourism and internal migration**

- **29 July 29 – 2 August 2024 in Jakarta**
  - **Workshop organized with new Regional Hub on Big Data and Data Science for Asia & the Pacific, hosted by BPS – Statistics Indonesia**
  - **https://hub.bps.go.id/workshop-mpd-processing-and-analytics-for-official-statistics/**

- **Objectives:**
  - **Increasing awareness of approaches to agreeing access to MPD and help equip participants to engage with national MNOs.**
  - **Developing technical skills for the processing & analysis of MPD for official statistics.**
  - **Discussing opportunities for leveraging MPD in participants' countries**

Webscraping for CPI

- **12 online training sessions from May – Sep 2024**
  - **Developing Python coding**
  - **Showcasing country experiences**
  - **Developing country plans**
  - **Building a webscraper**

- **Mentoring from experts from UN Regional Hub for Big Data and Data Science in Indonesia & UN Committee of Experts on Big Data Task Team**

- **In-person workshop 16-20 September 2024, Bangkok**
  - **Further develop skills in webscraping in Python**
  - **Complete webscraping for chosen websites**
  - **Developing a Reproduceable Analytical Pipeline (RAP)**
  - **Use of webscraped data within CPI**

- **Knowledge product based on resources developed for workstream –due April/May 2025**
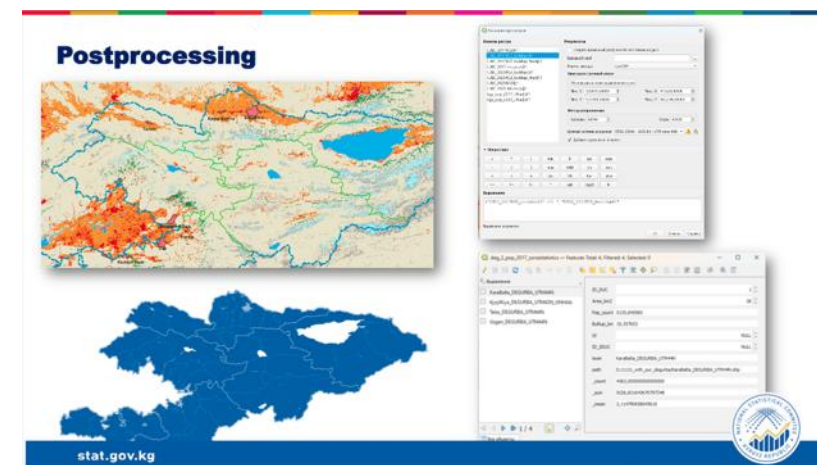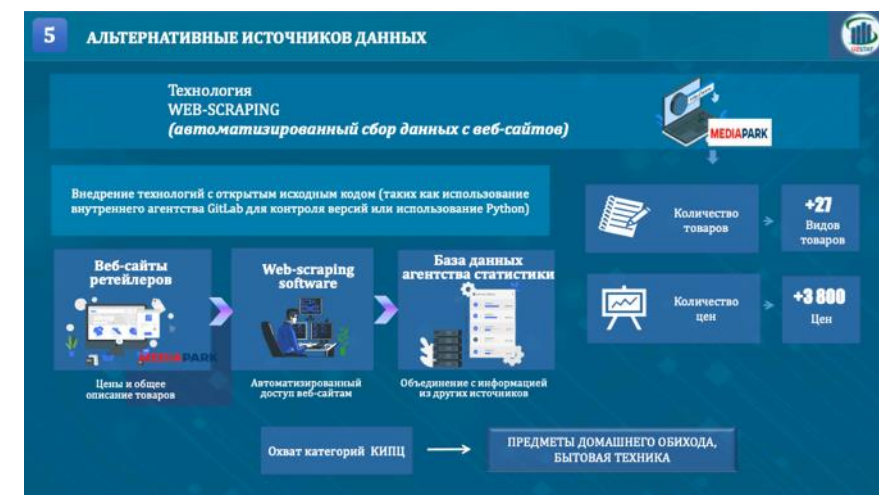
Use of geospatial data for small area estimation

- **Initial online R training session followed by 7-week guided e-learning on Small Area Estimation in R**

  - Videos, quizzes, assignments & weekly meetings

  - SAE e-learning: **https://siap-elearning.org/course/view.php?id=221**

- **In-person workshop focused on use of geospatial data – 21-25 October 2024**

  - Jointly organised with UNICEF and UNSD

  - Further training in R; use and acquisition of EO and other geospatial data; SAE using geospatial data; work on national data

  - Workshop materials: **https://github.com/ESCAP-SD/bangkokgeospatialsae**

  - Follow-up sessions with experts during Nov/Dec 2024

- **Practical guidance on SAE using geospatial data being developed with UNSD – available May 2025**

Regional workshop

- Virtual regional workshop on Big Data and Data Science for Official Statistics
  - 11-12 March 2025
  - Country presentations from Bangladesh, Cambodia, Indonesia, Kyrgyzstan, Malaysia, Thailand, Uzbekistan and Vanuatu
  - All materials and video available from event webpage: https://www.unescap.org/events/2025/big-data-and-data-science-official-statistics-virtual-workshop-asia-and-pacific

Resources: Geo-statistical integration

Key resources to support the integration of statistical and geospatial information:

➢ The Global Statistical Geospatial Framework (GSGF): The_GSGF.pdf

➢ GSGF Implementation Guide: https://unstats.un.org/unsd/statcom/53rd-session/documents/BG-3x-EG-ISGI-GSGF-Implementation-Guide-E.pdf

➢ The SDG Geospatial Roadmap: ggim.un.org/documents/SDGs-Geospatial-Roadmap.pdf

➢ The GSGF Self-Assessment Tool: https://docs.google.com/spreadsheets/d/1YSh-2Ipr_82qpgJRF41OqXpeVzwHzR6w/edit?usp=sharing&ouid=107089501156493068680&rtpof=true&sd=true

➢ Model agreement between NSO and NGIA: https://ggim.un.org/meetings/GGIM-committee/14th-Session/documents/Background_document_Institutional_agreement_for_GSGF.pdf

➢ Materials from High-Level Seminar on Integration of Statistical & Geospatial Information (UNSD, ESCAP & Statistics Norway): High-level seminar on integration of geospatial and statistical information | ESCAP

Resources: Geo-statistical integration

Practical guides and other resources from ESCAP:

➢ Beginners' Handbook: Geographic Information System and QGIS software: [Beginner's handbook : geographic information system and QGIS software | ESCAP](#)

➢ Producing Land Cover Change Maps and Statistics: Step by Step guide: [Producing land cover change maps and statistics : step by step guide on the use of QGIS and RStudio | ESCAP](#)

➢ Producing Land Cover Change Maps and Statistics: Advanced Use of QGIS and RStudio: [Producing land cover change maps and statistics : guide on advanced use of QGIS and RStudio | ESCAP](#)

Coming soon:
➢ Estimation of the rural access index (SDG 9.1.1) using QGIS: Step by step guide
➢ Estimation of land consumption and SDG 11.3.1 using QGIS: Step by step guide
➢ Applying the degree of urbanisation using QGIS: Step by step guide
➢ Data infrastructure for the application of big data (working paper)

Resources: Geo-statistical integration

Resources for small area estimation using geospatial data:

➢ E-learning: Disaggregated Estimates Using Small Area Estimation in R: Asia and the Pacific: Course: Disaggregated Estimates Using Small Area Estimation Methods in R: Asia and the Pacific

➢ Toolkit on Using Small Area Estimation for SDGs: SAE4SDG - SAE4SDG - UN Statistics Wiki

➢ Small Area Estimation with Geospatial Data: A primer: unstats.un.org/UNSDWebsite/statcom/session_56/documents/BG-3p-Geospatial_SAE_Primer-E.pdf

Coming soon:
➢ Geospatial Small Area Estimation: How-to guide

Resources: Alternative data sources for price statistics

- UN Task Team on Scanner data, web scraped data and admin data for CPI:
  - ➤ Handbook - [UN-CEBD Task Team on Scanner Data - UN-CEBD - Scanner Data Wiki - UN Statistics Wiki](#)
  - ➤ UN Learning Management System - online learning materials for Big Data with new courses coming soon)- https://learning.officialstatistics.org/

- ESCAP Stats Briefs:
  - ➤ Work with Uzbekistan to modernise the CPI: https://www.unescap.org/kp/2025/alternative-data-sources-price-statistics-uzbekistans-experience-modernizing-consumer-price
  - ➤ Big Data for Economic Statistics - https://repository.unescap.org/handle/20.500.12870/3501

Coming soon
  - ➤ Learning Resources from the ESCAP Web Scraping for CPI training programme will be available later this month

THANK YOU

Follow us:

unescap

unitednationsescap

www.unescap.org

unescap

unescap

united-nations-escap